Journal of Agricultural and Applied Economics, 46,4(November 2014):615–634

Household Expenditures on Vegetables in Malaysia

Andrew K.G. Tan, Steven T. Yen, Abdul Rahman Hasan, and Kamarudin Muhamed

Factors associated with purchase likelihoods and amounts spent on fresh and preserved vegetables in Malaysia are investigated. A sample selection system is applied to data from a national survey in Malaysia. Statistical test supports separate analyses for urban and rural households. Income and sociodemographic characteristics such as ethnicity, location of residence, household size, and education are closely associated with household expenditure patterns of fresh and preserved vegetables. Urban–rural differences are also uncovered in these effects. Several observations are noted vis-à-vis the expenditure patterns for fresh and preserved vegetables in Malaysia.

Key Words: censoring, Malaysia, sample selection system, vegetables

JEL Classifications: D10, D12

The World Health Organization (WHO, 2003) recommends an optimal diet of five servings of fruits and vegetables (at least 400 g) daily to prevent diet and nutrition-related chronic diseases. This recommendation consists of at least two daily servings of fruits (160 g) and three daily servings of vegetables (240 g) with at least one serving of nutrient-rich vegetables containing dark green and leafy or orange vegetables (Striegel-Moore et al., 2006).

In Malaysia, statistics from the Food and Agriculture Organization show that average daily per-capita consumption of vegetables increased steadily by approximately 70% between the periods of 1980s (66 g) and 2000s (112 g) (FAOSTAT–Food and Agriculture Organization Statistics Division, 2013). However, the amounts consumed are still far below the WHO-recommended dietary guideline of 240 g (or three daily servings). Underconsumption of vegetables among Malaysians is highlighted by the 2003 Malaysian Adult Nutrition Survey, whereby only 40% of adult Malaysians consumed one cup (approximately 96 g) of green leafy vegetables daily (Norimah et al., 2008). Data from various Malaysian Household Expenditure Surveys (1993–2010) confirm that although average household expenditures on food products have increased gradually over the years, the rise in spending on vegetables (42%) is lagging behind those of other higher value products such as fish (85%) and meat (56%) (Department of Statistics Malaysia, 2011). It is therefore surmised that Malaysians are not eating adequate amounts of vegetables relative to other staple foods such as fish, meat, or even rice.

Andrew K.G. Tan is an associate professor, School of Social Sciences, Universiti Sains Malaysia, Putrajaya, Malaysia. Steven T. Yen is a senior economist, STY Health Econometrics, Knoxville, Tennessee. Abdul Rahman Hasan is from the Department of Statistics Malaysia, Putrajaya, Malaysia. Kamarudin Muhamed is from the Department of Statistics Malaysia, Putrajaya, Malaysia.

Research support from the Universiti Sains Malaysia Research University grant (Grant No. 1001/PSOSIAL/816238) is acknowledged.

Although studies investigating the demand for vegetables are prevalent (Gustavsen and Rickertsen, 2006; Kasteridis and Yen, 2012; Mutuc, Pan, and Rejesus, 2007; Niu and Wohlgenant, 2013), few have examined the sociodemographic factors associated with purchase likelihoods and expenditure levels for vegetables, particularly in the context of a newly industrialized country such as Malaysia. Although Hadi et al. (2010) and Nurul Izzah et al. (2012) examined the demand for vegetables in Malaysia, these studies are limited in scope as a result of sectoral limitations of the data. Yen and Tan (2012) and Yen, Tan, and Nayga (2011) examined vegetable consumption in terms of weekly servings and frequency counts but did not take into consideration the expenditure patterns of Malaysian households. Tey et al. (2009) used nationwide household expenditures data to investigate the demand for various types of vegetables, focusing on expenditure and own-price elasticities. Tey et al. (2008) examined expenditure, quantity, and quality elasticities of vegetables with scant focus on the roles of sociodemographic characteristics.

Besides differences in consumer tastes and preferences for the specific types of vegetables, the market for vegetables itself is also distinctive. For example, vegetables can be purchased either fresh or preserved/canned for at-home consumption. This entails preconsumption preparation time and may therefore be directly related with the demand for convenience, a preference typically identified among younger, all-working, and professional households (Newman, Henchion, and Matthews, 2001). Moreover, although there exist numerous studies on vegetable consumption in the literature, differences in fresh and preserved consumption behaviors between urban and rural households have rarely been investigated with the exception of a few notable studies (Ettienne-Gittens et al., 2013; Mutuc, Pan, and Rejesus, 2007). Such analysis is relevant because urban and rural households may have different vegetable consumption patterns, perhaps as a result of differences in household composition, lifestyles, locations, and relative prices. Although vegetable prices may vary among locations, the standard of living (and hence prices) in metropolitan locales with higher population densities would be invariably higher than that of nonmetropolitan neighborhoods, like it would in urban than rural areas. Such regional price variations are hypothesized to affect vegetable consumption patterns in terms of purchase likelihoods and amounts spent. Moreover, it is expected that urban households with busy lifestyles and possessing demand attributes valuing convenience, easy storage, and long shelf life may prefer preserved instead of fresh vegetables. Rural households may possess wide-open spaces and fertile farmland for planting fresh produce. On the other hand, availability of grocery stores in more populated areas may provide easier access among urban households to acquire their supply of fresh vegetables.

From a health perspective, although vegetables are admittedly a large group and preferences may differ depending on vegetable type, there have also been growing health concerns that certain preserved food items may contain high contents of salts, sodium nitrites, nitrates, and other chemical preservatives or additives to prevent decay but are detrimental to health or may cause environmental pollution (Jian et al., 2004; White and McFadden, 2008). In this aspect, households with certain health-related demand attributes may favor fresh instead of preserved vegetables. Hence, information on whether there exists differential vegetable demand behavior among urban and rural populations may be important as a result of rising concerns of the growing urban–rural divide in targeting domestic food policies (Mutuc, Pan, and Rejesus, 2007). In summary, the sociodemographic characteristics associated with vegetable demand may be interpreted through a range of household demand attributes, including convenience, perceived healthiness, storage amenities, and even location.

The current analysis augments existing studies by examining the sociodemographic characteristics associated with purchase likelihoods and expenditure levels of vegetables in Malaysia. Based on a large household data set, a sample selection system procedure is used to account for censoring (observed zeroes) in household expenditures of two (fresh and preserved) types of vegetable products in Malaysia.

This system procedure is relevant because demand for fresh and preserved vegetables may conceivably be interdependent. The econometric model also features separate parameterization of the stochastic processes, which govern the discrete (zero-positive) and level outcomes, which are more flexible than the Tobit system, in which effects of explanatory variables on the probabilities and levels of consumption are restricted (Lin and Schmidt, 1984). Finally, on statistical grounds, the system estimation also produces more efficient empirical estimates relative to existing two-step estimates (e.g., Shonkwiler and Yen, 1999). Given the underconsumption of vegetables among Malaysians, it is therefore imperative to gain a better understanding of the factors associated with the likelihoods of purchasing and amounts spent on vegetables in Malaysia. A good understanding of the sociodemographic characteristics associated with household expenditures on vegetables is important to public health authorities concerned with the dietary patterns and health of Malaysians and to consumer insight analysts' intent on understanding their target markets.

Empirical Literature

Studies on the demand for vegetables and its correlation to sociodemographic characteristics have gained prominence in the literature. Dong and Stewart (2008) and Reynolds (1990) find household size to be positively associated with vegetable demand as households with more family members incurring higher expenses on a varied array of vegetables. Nonetheless, Kasteridis and Yen (2012) emphasize that this positive relationship occurs for certain vegetable types only and may have an opposite effect on organic vegetables. Reynolds (1990) and Tey et al. (2009) also note that economies of scale in vegetable consumption may occur, whereby household size has a positive impact on purchase decisions but a negative impact on amounts purchased.

Researchers have consistently found income levels positively associated with vegetable demand (Blisard, Stewart, and Jolliffe, 2004; Reynolds, 1990). Household income was found to contribute to household spending patterns on vegetables with the largest effect for low-consuming households (Gustavsen and Rickertsen, 2006). Additionally, affluent individuals demand higher quality vegetables and spend more per unit on the items, especially on nutrient-dense leafy and salad vegetables (Dong and Stewart, 2008; Tey et al., 2009).

Regional differences and urbanization are associated with consumption patterns for specific types of vegetable products. Kasteridis and Yen (2012) and Smith, Huang, and Lin (2009) note that relative to households in the western United States, those in the eastern, central, and southern regions eat less organic vegetables. However, regional differences are less obvious because eastern U.S. households consume more peppers and potatoes but less of other conventional vegetables. Meanwhile, Gustavsen and Rickertsen (2006) attribute higher vegetable purchases in major cities and lower purchases in rural areas to limited availability of fresh vegetables in rural areas. However, Tey et al. (2009) suggest that vegetable demand in Malaysia's urban markets may have reached saturation levels given that, although urban consumers demand more high-quality vegetables than rural consumers, expenditures are higher in rural than urban areas.

Existing studies on vegetable demand consider whites, blacks, Hispanics, and Asians in ethnic classifications. Striegel-Moore et al. (2006) suggest that although vegetable intake varies across ethnic/racial groups, children of white households consume smaller amounts of vegetables than those of blacks. Reynolds (1990) finds that nonwhite/nonblack households spend more on fresh vegetables than white/black households. Stewart and Harris (2005) note that Asians and Hispanics are more diversified in vegetable purchases than others. However, although black households purchase fewer types of vegetables than white and non-Hispanic households, they spread out purchases more evenly across what they do buy. Dong and Stewart (2008) further suggest that if vegetable choices are limited in stores, whites will buy more often, whereas if greater varieties are available, Asian and Hispanic households will purchase more instead.

Few studies have examined the association between genders of household heads to vegetable consumption patterns. Although Malaysian males have a greater degree of decision-making power and responsibilities in socioeconomic aspects like in most Asian countries, females play a very integral role in household food and family decisions (Tan, 2010). Meanwhile, previous studies have consistently shown a positive association between age and vegetable consumption (Gustavsen and Rickertsen, 2006; Reynolds, 1990). Although Blisard, Variyam, and Cromartie (2003) find households with mature family members incurring higher vegetable expenditures, Stewart and Harris (2005) note a tendency of younger households to purchase fewer varieties instead. One reason suggested by Tey et al. (2009) is that older persons are generally more health-conscious and possess greater nutritional and dietary needs. In terms of employment status, Dong and Stewart (2008) find that households headed by a working female spend more per unit on vegetables. In contrast, individuals working full-time may not have sufficient time for preparing meals at home, including those with vegetable dishes, thus accounting for their lower expenditures on vegetable products (Thiele and Weiss, 2003).

Education of the household head is often associated with patterns of vegetable consumption. Dong and Stewart (2008) found that households headed by females with post-high school education tend to purchase vegetables more frequently. Blisard, Variyam, and Cromartie (2003) posit that better educated household heads are more cognizant of the benefits of healthy diets and consequently spend more on vegetables. These authors also present the notion of education as a form of investment, whereby those who invest longer periods of time and effort in education pursuit value the future more highly than lower educated individuals. Therefore, because better educated individuals acquire higher discount rates of time or time preference, they may consider healthier diets containing vegetables as drivers in achieving future economic or social goals. In contrast, persons with more future uncertainties have lower discount rates of time because they value the present more and are less perturbed by the effects of present-day unhealthy diets on future health (Yen, Tan, and Nayga, 2011).

Theoretical Framework and Empirical Specification

Microdata allow in-depth investigation of the effects of economic and sociodemographic characteristics with sufficient degrees of freedom to estimate a large number of parameters. However, zero expenditures (i.e., the decision not to purchase) for vegetable products need to be accommodated to produce consistent estimates of the parameters. Statistical efficiency is also compromised in a single equation model. System estimation accommodates interactions of unobserved household characteristics among equations and improves statistical efficiency.

We begin with a utility maximization framework to motivate the empirical specifications for fresh and preserved vegetable demands. Households are assumed to maximize utility from consumption of consumer goods subject to a fixed budget:

$$(1) \qquad \max_{q,c} \{U(Dq, c; s) | \; p'q + c = m\},$$

where $q = [q_1,..., q_n]'$ is the quantity vector with positive prices $p = [p_1,..., p_n]'$, c is a composite commodity for other goods with price normalized at unity, s is a vector of demographic variables, m is budget, and $D = \text{diag}(d_1,..., d_n)$ is a diagonal matrix with elements d_i such that $d_i = 1$ if a household is a potential consumer of q_i and $d_i = 0$ otherwise. Solution to this optimization problem gives the demand functions for q_i among the consumers. Expressed in expenditure forms, zero observation in each expenditure $(y_i = p_i q_i)$ is governed by a sample selection mechanism. We draw on the multivariate sample selection system of Yen (2005) and consider a much simplified two-good system for fresh vegetable ($i = 1$) and preserved vegetable ($i = 2$), which is algebraically more transparent (observation subscript suppressed for brevity):

$$(2) \qquad \begin{array}{ll} \log y_i = x'\beta_i + v_i & \text{if} \quad z'\alpha_i + u_i > 0 \\ y_i = 0 & \text{if} \quad z'\alpha_i + u_i \le 0, \; i = 1, 2. \end{array}$$

In equation (2), x and z are vectors of explanatory variables with conformable parameters β_i

and α_i, the error terms $[u', v']' = [u_1, u_2, v_1, v_2]'$ are distributed as four-dimensional normal with zero means, variances $[1, 1, \sigma_1^2, \sigma_2^2]'$, correlations $R = [\rho_{ij}^{k\ell}]$, and covariance matrix $\Sigma = [\sigma_i \sigma_j \rho_{ij}^{k\ell}]$, where $i, j = 1, 2$ and $k, \ell = u, v$.[1] The log transformation for positive values of each dependent variable y_i ameliorates nonnormality and heteroscedasticity in the error term (Yen and Rosinski, 2008). The model represents a multiequation extension of the bivariate sample selection model (Heckman, 1979) and the single-hurdle model (Cragg, 1971; equations [12] and [13]); it is identical to that proposed by Shonkwiler and Yen (1999), but instead of two-step estimation, we use a more efficient maximum-likelihood procedure suggested by Yen (2005).

To construct the sample likelihood function, denote the k-dimensional standard normal probability density function as ϕ_k and cumulative distribution function (cdf) as Φ_k for $k = 1, 2$. For a sample regime in which expenditures for both goods are zeroes, the sample likelihood contribution is the bivariate standard normal probability, identical to that of a two-zero regime in bivariate probit:

$$L = \Phi_2(-z'\alpha_1, -z'\alpha_2; \rho_{21}^{uu}). \tag{3}$$

Define $h_i = (\log y_i - x'\beta_i)/\sigma_i$ for $i = 1, 2$. The likelihood contribution for a sample regime with $y_1 > 0$ and $y_2 = 0$ is

$$L = y_1^{-1}\sigma_1^{-1}\phi_1(h_1)\,\Phi_2(r_1, -r_2; -\tau_{12}), \tag{4}$$

where $r_1 = (z'\alpha_1 + \rho_{11}^{vu}h_1)/[1-(\rho_{11}^{vu})^2]^{1/2}$, $r_2 = (z'\alpha_2 + \rho_{12}^{vu}h_1)/[1-(\rho_{12}^{vu})^2]^{1/2}$, and $\tau_{12} = (\rho_{21}^{uu} - \rho_{11}^{vu}\rho_{12}^{vu})/\{[1-(\rho_{11}^{vu})^2]^{1/2}\ [1-(\rho_{12}^{vu})^2]^{1/2}\}$. Likelihood contribution for a regime with $y_1 = 0$ and $y_2 > 0$ is obtained by reciprocity.

For a sample regime with both goods positive, the error covariance matrix Σ is partitioned into submatrices $\Sigma_{11}, \Sigma_{12}, \Sigma_{21}$ and Σ_{22} of order two. Then, $(u_1, u_2 |\ v_1, v_2)$ is distributed as bivariate normal with mean vector $\Sigma_{12}\Sigma_{22}^{-1}v = [\mu_1, \mu_2]'$ and covariance matrix $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ with diagonal elements (variances) denoted as $\omega = [\omega_1^2, \omega_2^2]'$, covariance (off-diagonal element) as ω_{12}, and correlation as $\tau_{12} = \omega_{12}/(\omega_1, \omega_2)$ (Kotz, Balakrishnan, and Johnson, 2000). Then, the likelihood contribution is

$$L = y_1^{-1}y_2^{-1}\sigma_1^{-1}\sigma_2^{-1}\phi_2(h_1, h_2; \rho_{21}^{vv}) \times \Phi_2(q_1, q_2; \tau_{12}), \tag{5}$$

where $q_1 = (z'\alpha_1 + \mu_1)/\omega_1$, $q_2 = (z'\alpha_2 + \mu_2)/\omega_2$, and $(y_1^{-1}y_2^{-1})$ is the Jacobian of the transformation from (v_1, v_2) to (y_1, y_2). The unknown parameters are vectors from the selection equations (α_1, α_2), level equations (β_1, β_2), error standard deviations (σ_1, σ_2), and unique error correlations $(\rho_{21}^{uu}, \rho_{11}^{vu}, \rho_{12}^{vu}, \rho_{21}^{vu}, \rho_{22}^{vu}, \rho_{21}^{vv})$.

The sample selection system contains two nested specifications. First, imposing zero error correlations $\rho_{21}^{uu} = \rho_{12}^{vu} = \rho_{21}^{vu} = \rho_{21}^{vv} = 0$ (viz., all but $\rho_{11}^{vu} \neq 0$ and $\rho_{22}^{vu} \neq 0$) reduces the model to a "pairwise selection system," which corresponds to two sets of bivariate sample selection models (Heckman, 1979) stacked together. Furthermore, restricting all six unique error correlations to zero ($\rho_{ij}^{k\ell} = 0\ \forall\ i \neq j$ or $k \neq \ell$) produces the "independent system," which corresponds to two sets of two-part models (Duan et al., 1984) with separable parameter sets α_i and (β_i, σ_i) for $i = 1,2$, each of which can be estimated as probit for using the whole sample and ordinary least-squares for $\log y_i$ using the truncated sample conditional on $y_i > 0$. Tests against the two nested specifications can be accomplished by Wald, likelihood ratio, or Lagrange multiplier tests (Engle, 1984).

Marginal (discrete) effects of explanatory variables can be calculated by differentiating (differencing) the probability of a positive observation

$$\Pr(y_i > 0) = \Phi_1(z'\alpha_i), \tag{6}$$

the conditional mean of the dependent variable y_i (Yen and Rosinski, 2008, p. 5)

$$E(y_i\ |\ y_i > 0) = \exp(x'\beta_i + \sigma_i^2/2) \times \Phi_1(z'\alpha_i + \rho_{ii}^{vu}\sigma_i)\ /\Phi_1(z'\alpha_i), \tag{7}$$

and the unconditional mean of y_i (using equations [6] and [7])

$$E(y_i) = \exp(x'\beta_i + \sigma_i^2/2)\,\Phi_1(z'\alpha_i + \rho_{ii}^{vu}\sigma_i)\,. \tag{8}$$

Marginal effects are calculated for all observations and averaged over the sample. For

[1] Prices are constant in a single cross section and are absorbed in the constant terms in z and x.

statistical inference, standard errors are calculated by an approximation procedure known as the delta method.

Data and Variables

The Survey

Data came from the Malaysian Household Expenditure Survey (MHES) 2009/2010. This data set was collected by the Department of Statistics Malaysia and is the most recent of the national household expenditure surveys. The sample was designed based on a stratified multistage, area probability sampling method to ensure sociodemographic and geographic representativeness of the Malaysian population.

In the survey, respondents were asked to record their monthly (nominal) expenditures on fresh and preserved vegetable products (see definitions in Table 1). Sociodemographic information of the respondents was also collected. A total number of 21,641 households responded to the survey, but, after removing households with incomplete information on important variables, 21,542 observations were retained. From this sample, 20,507 (95.2%) and 11,015 (51.1%) households reported fresh and preserved vegetable expenditures, respectively, during the survey period. The final sample contains 14,917 urban (69.2%) and 6625 (30.8%) rural households.[2]

Variables

Selection of variables hypothesized to be associated with household vegetable demand relies on previous studies by Gustavsen and Rickertsen (2006), Reynolds (1990), Stewart and Blisard (2008), Stewart and Harris (2005), Tey et al. (2009), Yen and Tan (2012), and Yen, Tan, and Nayga (2011), among others. The sociodemographic characteristics hypothesized to influence the probabilities and amounts of expenditures on fresh and preserved vegetables include: 1) household size, 2) ethnicity, and 3) monthly household income brackets; and 4) gender, 5) age group, 6) occupation type, and 7) education level of the household head (Table 1).

Ethnicity of household head was represented by a set of categorical variables composed of Malay (base group), Chinese, Indian, and others. Age of the household head is coded into four categories (18–29, 30–45, 46–59, ≥60 years) to denote younger, middle-age younger (base group), middle-age older, and retiree household heads, respectively. The highest level of the household head's formal education is segregated into three categories (none/primary, secondary, tertiary). Monthly household income, measured in Ringgit Malaysia (RM), is coded into categories to represent poverty-low (RM 0–999, U.S. $0–296),[3] lower-middle (RM 1000–3999, U.S. $297–1186), upper-middle (RM 4000–7999, U.S. $1187–2373), and high (RM ≥8000, U.S. ≥$2374) income brackets. Also included are dummy variables indicating gender (male) and occupation type (white collar) of the household head (Table 1).

Table 1 also presents the descriptive statistics. Among consuming households, urban households on average spend more in both fresh (RM 43.46) and preserved (RM 7.14) vegetables than rural households (RM 41.40 and RM 6.02, respectively). This could be attributed to the higher average monthly household incomes of urban (RM 4251.21) compared with rural (RM 2573.03) households. Ethnic composition of households in urban (rural) areas consists of 58% (79%) Malays, 30% (9%) Chinese, 7% (4%) Indians, and 5% (7%) of other ethnic backgrounds, suggesting higher proportions of Malays but lower proportion of Chinese in rural than urban areas. The proportions of household heads with white collar occupations are higher among urban (16%) than rural (7%) surroundings. Similarly, the proportion of household heads that are uneducated or with primary level education living

[2] The Department of Statistics Malaysia (2011) classifies urban areas as those with 10,000 or more in population and rural areas as those with less than 10,000 inhabitants.

[3] As of January 1, 2010, exchange rate was approximately U.S. $1.00 = RM 3.37 or RM 1.00 = U.S. $0.30.

Table 1. Definitions and Sample Statistics of Variables in the Statistical Model (n = 21,542)[a]

Variable	Definition	Urban	Rural	Total
Dependent variables				
Fresh vegetables	Monthly household expenditures (in RM) on fresh vegetables (e.g., tomatoes/green beans/green spinach/ lettuce/ cauliflower/spring onions/white cabbage/celery/brinjals/ fresh mushrooms/asparagus/broccoli, etc.)	40.76 (32.65)	40.29 (35.34)	40.61 (33.50)
	Among the consuming households (HHs)	43.46 (31.93)	41.40 (35.17)	42.66 (33.04)
	Percent consuming	64.94%	29.92%	95.20%
Preserved vegetables	Monthly household expenditures (in RM) on preserved vegetables (e.g., preserved leeks/salted cabbage/dried or canned mushrooms/soya beans/bean curd/other canned or preserved vegetables, etc.)	3.43 (6.55)	2.65 (5.16)	3.19 (6.17)
	Among the consuming HHs	7.14 (8.01)	6.02 (6.45)	6.24 (7.44)
	Percent consuming	32.88%	13.24%	51.13%
Continuous explanatory variables				
HH size	Number of family members in HH	4.10 (2.13)	4.28 (2.21)	4.15 (2.16)
HH income	Monthly HH income (in RM)	4251.21 (4356.42)	2573.03 (2138.85)	3735.11 (3892.05)
Binary explanatory variables (yes = 1; no = 0)				
Malay	Ethnicity of HH head is Malay (reference)	0.58	0.79	0.64
Chinese	Ethnicity of HH head is Chinese	0.30	0.09	0.24
Indian	Ethnicity of HH head is Indian	0.07	0.04	0.06
Others	Ethnicity of HH head is others	0.05	0.07	0.06
Male	HH head is male	0.83	0.85	0.83
Age 18–29	Age of HH head 18–29 years (younger)	0.13	0.10	0.12
Age 30–45	Age of HH head 30–45 years (middle-age younger, reference)	0.39	0.38	0.39
Age 46–59	Age of HH head 46–59 years (middle-age older)	0.32	0.32	0.32
Age ≥ 60	Age of HH head 60 years and older (retiree)	0.15	0.21	0.17
White collar	HH head has white collar occupation (e.g., legislators/senior officials/managers/professionals)	0.16	0.07	0.13

Table 1. Continued

Variable	Definition	Urban	Rural	Total
None/primary	HH head is not educated or has primary school as highest level	0.25	0.45	0.31
Secondary	Highest level of education for HH head is secondary/high school (reference)	0.55	0.47	0.53
Tertiary	Highest level of education for HH head is tertiary	0.20	0.08	0.17
Poverty-low[b]	Monthly HH income bracket RM 0–999[c]	0.06	0.15	0.09
Lower-middle	Monthly HH income bracket RM 1000–3999 (reference)	0.57	0.68	0.60
Upper-middle	Monthly HH income bracket RM 4000–7999	0.26	0.13	0.22
High income	Monthly HH income bracket RM 8000 and above	0.11	0.03	0.09
Sample size		14,917	6,625	21,542

Source: Department of Statistics Malaysia (2011).

Note: RM, Ringgit Malaysia.

[a] Values for continuous variables are sample means and standard deviations in parentheses.

[b] The four income categories correspond to poverty-low (U.S. $0–296), lower-middle (U.S. $297–1186), upper-middle (U.S. $1187–2373), and high (U.S. ≥$2374) income brackets.

[c] As of January 1, 2010, exchange rate was approximately U.S. $1.00 = RM 3.37 or RM 1.00 = U.S. $0.30.

in rural areas (45%) exceeds that in urban areas (24%), whereas the breakdown of those with tertiary education favors urbanites (20%) compared with rural residents (8%). Last, the percentage of households listed in the poverty-low income bracket living in rural areas (15%) is higher than those in the urban areas (6%), whereas the proportion of high-income households is higher in urban (11%) than rural (3%) areas.

Results

We first investigate poolability of the sample by testing for equality of all "slope" parameters between urban and rural households. Denote the maximum log likelihood values of the urban sample, rural sample, and pooled sample (with an urban dummy variable in all equations) as $\log L_1$, $\log L_2$, and $\log L_p$, with corresponding numbers of parameters k_1,k_2, and k_p. Then, the likelihood ratio (LR) statistic $2(\log L_1 + \log L_2 - \log L_p)$ is χ^2 distributed with $(k_1 + k_2 - k_p)$ degrees of freedom (df). Based on maximum log likelihood values of the three samples, the hypothesis of equal slope coefficients between urban and rural samples is rejected (LR = 217.39, df = 62, $p < 0.0001$), which suggests separate analyses for urban and rural households.

We then test for superiority of the sample selection system over the two restricted specifications discussed previously. Based on maximum likelihood (ML) estimates, Wald test results suggest that the sample selection system performs better than the pairwise selection system (Heckman, 1979) ($\chi^2 = 3385.23$, df = 4) and independent system ($\chi^2 = 4860.55$, df = 6) for the urban sample. The corresponding Wald statistics are 667.01 and 1986.94 for the rural sample. All tests are significant with a p value < 0.0001. Likelihood ratio and Lagrange multiplier tests suggest similar findings.[4]

[4] Vuong's (1989) nonnested tests suggest the sample selection system performs better than the bivariate Tobit model in fitting the data for both the urban (standard normal statistic $z = 8.04$) and rural ($z = 4.20$) samples, both with a p value < 0.0001.

Parameter Estimates

ML estimates of the sample selection system are presented in the appendix (Table A1). Results for the pooled sample are available on request. We briefly summarize the ML estimates. All six error correlation coefficients are significant at the 1% level for the urban sample, and all but one coefficient are significant at the 5% level or lower for the rural sample. Significance of these error correlations corroborates results of the Wald tests and justifies accommodation of endogenous sample selection and estimation of all equations in a system to improve statistical efficiency relative to the pairwise selection system and independent system.

There are notable differences in the parameter estimates between the urban and rural samples. For instance, whereas tertiary education is significant and negative in all selection and level equations for urban households, the variable is never significant for the rural sample. The variable is significant in all equations as well in the pooled sample dictated by the much larger urban sample. Urban–rural differences such as this are not likely to be found in the use of a pooled sample and highlights the importance of segmented sample analyses.

Marginal Effects of Explanatory Variables

Marginal effects of explanatory variables on the probabilities, conditional levels, and unconditional levels of expenditures are presented in Table 2 for the urban sample and Table 3 for the rural sample. In what follows, the marginal effect of an explanatory variable on the probability reflects contribution of the variable on the likelihood of consuming (i.e., a positive expenditure outcome) for each unit increase in the variable. The marginal effect on the conditional level indicates effect of a unit increase in the variable on the level of monthly expenditure, conditional on expending (or among the consuming). Finally, the marginal effect on the unconditional level signifies effect on the monthly level unconditionally (viz., overall, among the whole population of interest).

Our pooled sample estimates (table not shown) suggest that relative to rural households, urbanites display lower purchase likelihoods (2.13 percentage points, henceforth, %) and expenditure levels (RM 1.65 conditional and RM 2.28 unconditional) on fresh vegetables. Although urban households display lower propensities (2.79%) to purchase preserved vegetables than rural households, the effects on conditional and unconditional spending levels are not statistically significant. Our test result, presented previously, suggests that use of an urban dummy variable does not adequately account for the urban–rural differences in consumption. Such differences are more easily seen in the marginal effects by samples, on which we focus next.

Urban Sample. Many of the variables (e.g., ethnicity and income-level dummies) have different effects in terms of signs and statistical significance on the probabilities and conditional levels of consumption (this is also the case for rural households). These different effects would have been masked by the Tobit parameterization and highlight one important advantage of the sample selection system vis-à-vis the Tobit system.

Household size is significantly and positively associated with purchase likelihoods and expenditure levels of vegetables in Malaysia (Table 2). An additional family member contributes to higher probabilities of fresh (by 3.9%) and preserved (2.5%) vegetable purchases among urban households. Among consuming households, an additional family member increases conditional (RM 5.31; U.S. $1.58) and unconditional (RM 6.36; U.S. $1.89) fresh vegetable expenditures, whereas preserved vegetable expenditures increase by RM 0.40 (U.S. $0.12) and RM 0.39 (U.S. $0.11), respectively.[5]

Chinese are 2.0% less likely to expend on fresh vegetables but 21.7% more likely to

[5] To illustrate, the prevailing (May 13, 2014) market prices of a standard can of canned baby corn, straw mushrooms, and pickled lettuce sells are RM 4.59 (U.S. $1.37), RM 4.25 (U.S. $1.27), and RM 2.29 (U.S. $0.68), respectively. The approximate market price ranges of a kilogram of fresh tomatoes (RM 2.503.00; U.S. $0.740.89), kacang buncis or green beans (RM 7.008.00; U.S. $2.082.37), and sawi hijau or green mustard (RM 3.504.00; U.S. $1.041.19) are noted (Ministry of Agriculture and Agro-Based Industry Malaysia, 2014).

Table 2. Marginal Effects of Explanatory Variables on Probabilities, Conditional Levels, and Unconditional Levels of Expenditures: Urban Households

	Fresh Vegetables			Preserved Vegetables		
Variable	Probability (×100)	Conditional Level (RM)	Unconditional Level (RM)	Probability (×100)	Conditional Level (RM)	Unconditional Level (RM)
Continuous explanatory variables						
HH size	3.935***	5.305***	6.357***	2.525***	0.404***	0.385***
	(0.137)	(0.167)	(0.166)	(0.196)	(0.036)	(0.025)
HH income/ 100		0.252***	0.238***		0.046***	0.024***
		(0.020)	(0.019)		(0.005)	(0.003)
Binary explanatory variables						
Chinese	−2.047***	11.030***	9.938***	21.710***	3.310***	3.439***
	(0.413)	(0.748)	(0.731)	(0.897)	(0.191)	(0.153)
Indian	−1.141	11.014***	10.161***	23.879***	4.546***	4.926***
	(0.828)	(1.488)	(1.458)	(1.360)	(0.423)	(0.376)
Others	1.724***	1.491	1.991	3.576**	1.506***	1.120***
	(0.553)	(1.402)	(1.374)	(1.804)	(0.402)	(0.289)
Male	−2.941***	0.935	−0.003	−1.105	−0.019	−0.080
	(0.347)	(0.767)	(0.753)	(1.080)	(0.173)	(0.123)
Age 18–29	−4.234***	−7.257***	−8.245***	−8.799***	−0.406*	−0.755***
	(0.570)	(0.777)	(0.754)	(1.296)	(0.219)	(0.136)
Age 46–59	2.846***	8.147***	8.826***	4.775***	0.607***	0.649***
	(0.385)	(0.775)	(0.765)	(0.947)	(0.152)	(0.111)
Age ≥ 60	3.571***	11.600***	12.469***	7.716***	1.144***	1.185***
	(0.428)	(1.181)	(1.162)	(1.294)	(0.232)	(0.177)
White collar	0.535	−0.902	−0.700	−1.102	−0.076	−0.112
	(0.497)	(0.912)	(0.893)	(1.284)	(0.200)	(0.140)
None/ primary	0.018	2.291***	2.211**	−0.663	0.462***	0.215*
	(0.507)	(0.884)	(0.866)	(1.068)	(0.170)	(0.121)
Tertiary	−1.535***	−5.627***	−5.878	−5.358***	−0.193	−0.441***

Table 2. Continued

Variable	Fresh Vegetables			Preserved Vegetables		
	Probability (×100)	Conditional Level (RM)	Unconditional Level (RM)	Probability (×100)	Conditional Level (RM)	Unconditional Level (RM)
	(0.476)	(0.766)	(0.748)	(1.168)	(0.182)	(0.124)
Poverty-low	−3.285***	−0.361***	−1.407***	−11.314***	0.730***	−0.390***
	(0.884)	(0.122)	(0.386)	(1.659)	(0.132)	(0.068)
Upper-middle	−0.584	−0.064	−0.247	3.796***	−0.213***	0.116***
	(0.456)	(0.052)	(0.193)	(0.922)	(0.053)	(0.029)
High income	−3.116***	−0.337***	−1.316***	2.150	−0.121	0.066
	(0.815)	(0.111)	(0.347)	(1.451)	(0.081)	(0.044)

Note: Asymptotic standard errors in parentheses. Asterisks *** indicate statistical significance at 1% level, ** at 5% level, and * at 10% level. RM, Ringgit Malaysia; HH, household.

acquire preserved vegetables than Malays. Meanwhile, Indian (other ethnic) households are 23.9% (3.6%) more likely to purchase preserved vegetables than Malays with other factors remaining constant. Among consuming households, Chinese (RM 11.03) and Indians (RM 11.01) spend more on fresh vegetables than Malay households. Similarly, Chinese (RM 3.31), Indian (RM 4.55) and other (RM 1.51) ethnic households who consume preserved vegetables also expend more than Malay households. Overall, Chinese (RM 9.94) and Indian (RM 10.16) households contribute higher outlays on fresh vegetables while spending RM 3.44 and RM 4.93 more, respectively, on preserved vegetables. Overall, households of other ethnicity spend RM 1.12 more on preserved vegetables than Malays.

Urban households managed by males display a 2.9% lower likelihood to purchase fresh vegetables than female-managed households. Younger households (heads age 18–29 years) are less likely to spend on fresh (4.2%) and preserved (8.8%) vegetables than younger-middle age (30–45 years) households. Instead, older-middle age (46–59 years) households exhibit higher propensities to procure fresh (2.8%) and preserved (4.8%) vegetables than their younger-middle age cohorts. Similarly, households with a retired household head (≥60 years) display higher propensities to purchase fresh (3.6%) and preserved (7.7%) vegetables than younger-middle age household heads. Among consuming households, those with a younger household head spend RM 7.26 (RM 0.41) less on fresh (preserved) vegetables than their younger-middle age counterparts. Overall, households with a younger household manager spend less on fresh (RM 8.25) and preserved (RM 0.76) vegetables than those with a younger-middle age household head. Within consuming households, those headed by older-middle age (RM 8.15) and retiree (RM 11.60) persons spend significantly more on fresh vegetables while spending RM 0.61 and RM 1.14 more on preserved vegetables, respectively, than their younger-middle age peers. Overall, families with an older-middle age (RM 8.83) and retiree (RM 12.47) household head pay more for fresh vegetables than

Table 3. Marginal Effects of Explanatory Variables on Probabilities, Conditional Levels, and Unconditional Levels of Expenditures: Rural Households

	Fresh Vegetables			Preserved Vegetables		
Variable	Probability (× 100)	Conditional Level (RM)	Unconditional Level (RM)	Probability (× 100)	Conditional Level (RM)	Unconditional Level (RM)
Continuous explanatory variables						
HH size	2.261***	3.342***	3.981***	0.779***	0.264***	0.175***
	(0.194)	(0.204)	(0.211)	(0.298)	(0.039)	(0.025)
HH income/ 100		0.678***	0.659***		0.039***	0.018**
		(0.034)	(0.034)		(0.012)	(0.005)
Binary explanatory variables						
Chinese	−1.239	13.504***	12.827***	22.203***	3.646***	3.589***
	(0.798)	(1.567)	(1.573)	(1.991)	(0.350)	(0.289)
Indian	−0.889	14.994***	14.393***	30.299***	3.872***	4.495***
	(1.242)	(2.920)	(2.886)	(2.466)	(0.514)	(0.457)
Others	1.516***	−2.511*	−2.040	−5.750**	2.858***	1.074***
	(0.407)	(1.421)	(1.408)	(2.525)	(0.507)	(0.305)
Male	−1.763***	1.663	1.121	−2.746	0.033	−0.117
	(0.362)	(1.073)	(1.065)	(1.773)	(0.223)	(0.149)
Age 18–29	−2.171***	−4.939***	−5.492***	−8.832***	0.261	−0.319*
	(0.778)	(1.181)	(1.169)	(2.185)	(0.317)	(0.181)
Age 46–59	0.581	8.123***	8.186***	5.098***	0.685***	0.617
	(0.481)	(1.034)	(1.032)	(1.527)	(0.205)	(0.137)
Age ≥ 60	0.604	9.416***	9.469***	10.735***	0.993***	1.103***
	(0.605)	(1.452)	(1.437)	(1.928)	(0.267)	(0.187)
White collar	−0.553	0.750	0.566	0.786	−0.123	−0.026
	(0.728)	(1.646)	(1.627)	(2.539)	(0.289)	(0.198)
None/ primary	0.870	3.319***	3.533***	−3.326**	0.090	−0.117
	(0.549)	(0.969)	(0.967)	(1.486)	(0.186)	(0.120)
Tertiary	0.230	−2.549*	−2.441*	−2.849	0.092	−0.094
	(0.517)	(1.449)	(1.434)	(2.559)	(0.327)	(0.209)
Poverty-low	−1.003*	−0.151	−0.463*	−14.008***	−0.388**	−0.853***
	(0.609)	(0.094)	(0.279)	(1.748)	(0.161)	(0.118)

Table 3. Continued

Variable	Fresh Vegetables			Preserved Vegetables		
	Probability (× 100)	Conditional Level (RM)	Unconditional Level (RM)	Probability (× 100)	Conditional Level (RM)	Unconditional Level (RM)
Upper-middle	−0.304	−0.045	−0.137	3.509*	0.095	0.223*
	(0.758)	(0.112)	(0.341)	(1.823)	(0.062)	(0.118)
High income	−1.356	−0.196	−0.608	−4.287	−0.118	−0.267
	(1.420)	(0.207)	(0.633)	(3.639)	(0.111)	(0.225)

Note: Asymptotic standard errors in parentheses. Asterisks *** indicate statistical significance at 1% level, ** at 5% level, and * at 10% level. RM, Ringgit Malaysia; HH, household.

younger-middle age household heads. Similarly, higher overall expenditures are observed for preserved vegetables among households helmed by older-middle age (RM 0.65) and retiree (RM 1.19) persons than younger-middle age individuals. These results suggest a pattern of healthier food consumption habits by households with mature-age decision-makers given their elevated spending propensities and expenditures on fresh and preserved vegetables.

Compared with those with secondary education, household heads with no or primary education spend more on fresh vegetables conditional (RM 2.29) and unconditional (RM 2.21) on spending. These lower-educated households also display higher conditional (RM 0.46) and unconditional (RM 0.22) expenditures on preserved vegetables than secondary-educated households. However, tertiary-educated households display lower consumption likelihoods (1.5%) and conditional (RM 5.63) expenditures on fresh vegetables than their secondary-educated counterparts. Similarly, households with tertiary-educated household heads display lower consumption likelihoods (5.4%) and unconditional (RM 0.44) expenditures on preserved vegetables than those with secondary education.

An additional RM 100 in household income per month increases fresh vegetable expenditures by RM 0.25 conditional on purchase and RM 0.24 unconditional. The effects on preserved vegetables are smaller—RM 0.05 and RM 0.02, respectively. Relative to lower-middle income families, poverty-low income households are less likely to purchase fresh (3.3%) and preserved (11.3%) vegetables. Interestingly, high-income households are 3.1% less likely to purchase fresh vegetables than lower-middle income households. Households in the poverty-low bracket spend lower amounts in conditional (RM 0.36) and unconditional (RM 1.41) expenditures compared with lower-middle income households. Furthermore, households in the high income bracket expend less in conditional (RM 0.34) and unconditional (RM 1.32) expenditures on fresh vegetables. For preserved vegetables, poverty-low households spend RM 0.39 less unconditional amount of expenditures than lower-middle income households.

Meanwhile, upper-middle income households are 3.8% more likely to purchase preserved vegetables while spending RM 0.12 more in unconditional expenditures than their lower-middle income cohorts.

Rural Sample. An additional family member increases the likelihoods of fresh (2.3%) and preserved (0.8%) vegetables among rural households (Table 3). Each additional family member increases conditional (RM 3.34; U.S. $0.99) and unconditional (RM 3.98; U.S. $1.18) fresh vegetable expenditures, whereas preserved vegetable expenditures increase by RM 0.26 (U.S. $0.08) and RM 0.18 (U.S. $0.05), respectively. These effects on probabilities and levels are uniformly smaller than the urban households.

Within rural households, Chinese (22.2%) and Indians (30.3%) are more likely to acquire preserved vegetables than Malays. Although households of other ethnic background are 1.5% more likely to purchase fresh vegetables, they display 5.8% lower purchase likelihoods on preserved vegetables than Malays with all else constant. Chinese spend RM 13.50 (RM 12.83) more in conditional (unconditional) expenditures on fresh vegetables than Malays. Indian households exhibit RM 14.99 (RM 14.39) higher conditional (unconditional) outlays on fresh vegetables than Malays. Chinese households also spend more in terms of conditional (RM 3.65) and unconditional (RM 3.59) expenditures on preserved vegetables than Malays. Indian households display higher conditional (RM 3.87) and unconditional (RM 4.50) preserved vegetable purchases than Malays. Likewise, households of other ethnic background exhibit higher conditional (RM 2.86) and unconditional (RM 1.07) expenditures on preserved vegetables compared with Malays.

Rural households headed by males are 1.8% less likely to spend on fresh vegetables than those led by females. Households headed by individuals aged 18–29 years are less likely to spend on fresh (2.2%) and preserved (8.8%) vegetables than younger-middle age (30–45 years) individuals. Meanwhile, older-middle age (46–59 years) and retiree (≥60 years) households exhibit 5.1% and 10.7%, respectively, higher propensities to procure preserved vegetables than their younger-middle age cohorts. Within consuming households of fresh vegetables, those with a younger household head spend RM 4.94 less, whereas households led by older-middle age persons (46–59 years) (RM 8.12) and retirees (≥60 years) (RM 9.42) spend more than their younger-middle age counterparts. Household managed by older-middle age persons (46–59 years) (RM 0.69) and retirees (≥60 years) (RM 0.99) spend more in terms of conditional expenditures on preserved vegetables than their younger-middle age counterparts. Overall, households with a younger household manager spend less on fresh (RM 5.49) and preserved (RM 0.32) vegetables than those with a younger-middle age household head. Within consuming households, those headed by older-middle age (RM 8.12) and retiree (RM 9.42) persons spend significantly more on fresh vegetables while spending RM 0.69 and RM 0.99 more on preserved vegetables, respectively, than their younger-middle age peers. Overall, older-middle age (RM 8.19) and retiree (RM 9.47) households show higher purchases of fresh vegetables than their younger-middle age cohorts. Families led by a retiree also spend RM 1.10 more in unconditional expenditures on preserved vegetables than those led by younger persons. Again, these results suggest a pattern of healthier food consumption habits by households with mature-age decision-makers given their elevated spending propensities and expenditures on fresh and preserved vegetables among rural households.

Compared with those with secondary education, household heads with no or primary level education spend more on conditional (RM 3.32) and unconditional (RM 3.53) expenditures on fresh vegetables. However, tertiary-educated households display lower conditional (RM 2.55) and unconditional (RM 2.44) expenditures on fresh vegetables than those with secondary education. Households led by individuals with no or primary education are 3.3% less likely to procure preserved vegetables than those led by secondary educated individuals.

The effects of household income differ notably from those for urban households—higher on fresh vegetables but lower on preserved

vegetables. An additional RM 100 in household income per month increases fresh vegetable expenditures by RM 0.68 conditional on purchase and RM 0.66 unconditional. The effects on preserved vegetables are smaller than those for the urban households—RM 0.04 conditional on purchase and RM 0.02 unconditional. Poverty-low income households are 1.0% less likely to expend on fresh vegetables while spending RM 0.46 less unconditional expenditures than lower-middle income households. Relative to lower-middle income families, poverty-low income households are 14.0% less likely to purchase preserved vegetables while spending less in conditional (RM 0.39) and unconditional (RM 0.85) expenditures. Meanwhile, upper-middle income households exhibit 3.5% higher propensity to purchase and spend RM 0.22 more in unconditional expenditures on preserved vegetables than households in the lower-middle income bracket.

Discussion and Concluding Remarks

The sample selection system allows an investigation of the sociodemographic determinants of the probabilities of purchasing and the levels of spending on fresh and preserved vegetables among urban and rural households in Malaysia. In accommodating the censored expenditures, the sample selection parameterization ameliorates shortcomings of the Tobit system and the system estimation also improves statistical efficiency of the empirical estimates relative to two-step estimation. The sample selection system is also found to perform better than the pairwise selection system, independent system, and Tobit system in fitting the data.

Results indicate that household (income bracket, family size, ethnicity) and household head (age, education level) characteristics play key roles in determining consumption likelihoods and expenditure patterns of fresh and preserved vegetables among urban and rural households in Malaysia. Segmented-sample analysis also uncovers major urban–rural differences in the effects of income and many sociodemographic variables. Several observations are noted. First, poverty-low and high-income households in urban areas exhibit lower propensities to purchase and spend less on fresh vegetables than their middle income cohorts. This can be attributed to the wider array of food types and eating establishments in urban compared with rural areas as well as the notion that affluent urban households may better afford pricier consumption alternatives such as fish and meat products (Lee and Tan, 2007). Meanwhile, poverty-low income urban households may not be able to afford fresh vegetables given their tighter budget constraints amid higher costs of living in the cities. For these low-income households, healthy fresh vegetable consumption may not be considered as a basic necessity because health consciousness may set in only once all basic needs (e.g., housing payments/rentals, school fees, medical expenses, clothing) are met.

Second, whereas tertiary education lowers fresh vegetable consumption propensities and expenditure levels for urban households, it does not play a role for rural households. Besides the higher proportion of college-educated household heads in urban (20.0%) compared with rural (8.0%) areas, these seemingly unlikely results can be explained by the fact that higher education levels may translate to better paying occupations and thus result in more hectic working schedules among urbanites. This phenomenon may be indicative of changing lifestyles and cooking habits in the urban areas, whereby households with time constraints may be more accustomed to eating out instead of cooking from scratch and having their meals at home. Because fresh vegetables constitute one of the primary ingredients in meals prepared from scratch (Stewart and Blisard, 2008), consumption likelihoods and levels for fresh vegetables may decline among educated urban households with time constraints. Policy-wise, it may be fruitful to promote simple and varied cooking recipes among urbanites who may face time constraints in fresh meal preparations.

Third, findings from the current study using household expenditures data corroborate those of Yen and Tan (2012) and Yen, Tan, and Nayga (2011) based on individual consumption data that age is a significant and positive contributor

to the demand for vegetables in Malaysia. Because households managed by older and more matured household decision-makers display higher purchase likelihoods and expenditures on both vegetable types than households led by younger individuals, these effects are consistent among urban and rural households, although the magnitude of conditional and unconditional level expenditures are higher among urban than rural households. According to Guthrie and Lin (2002) and Nayga (1995), food consumption patterns of older individuals are more likely to be based on health considerations. In essence, individuals become more cognizant of their health and dietary habits and become "healthy eaters" as they age. This results in higher consumption likelihoods and expenditures on healthful foods such as vegetables, because older household heads influence their household's consumption patterns for health reasons. In terms of policy implications, it would be prudent for the Malaysian health authorities to continue promoting and creating awareness about the beneficial attributes of vegetable consumption, particularly among those in the younger age groups. At present, the probability of consumption and expenditure levels on fresh and preserved vegetables by younger households, either in urban or rural settings, are still lagging compared with the older age groups.

Fourth, the positive relationship between household size and demand for vegetables among urban and rural households contrasts that of Tey et al. (2009) in which a negative relationship is found between vegetable expenditures and household size. Instead, the positive association established in the present analysis suggests that both fresh and preserved vegetables, when viewed from the perspective of complementary consumption, may be necessity provisions in Malaysian diets. Another reason can be attributed to the economies of scales in meal preparation because both vegetable types are used as primary ingredients for cooking for more individuals in the family (Lee and Tan, 2007). As such, increases in household size induce higher purchase likelihoods and proportional increases in expenditures of basic household necessities such as fresh and preserved vegetables irrespective of urbanicity considerations.

Last, in terms of marketing implications, market penetration strategies may target urban Chinese households to boost consumption frequency of fresh vegetables. Although urban Chinese households may have lower propensities to purchase fresh vegetables, once in the market, both urban and rural Chinese households will spend higher expenditures on the item. Market expansion plans to develop sales through new consumers, particularly among rural households of Indian ethnic descent, could be introduced given their lack of significance in terms of vegetable expenditures. These efforts could include the use of language-based media advertisements (e.g., newspapers, TV programs, radio channels) because ethnicity is found to be a significant factor in vegetable demand.

Although the current analysis augments existing studies in examining the patterns and urban–rural differences of household expenditures on vegetables in Malaysia, several limitations are acknowledged based on the secondary nature of the MHES 2009/2010. First, only values of total monthly household expenditures on fresh and preserved vegetable products are available, even because information on vegetables produced for self-consumption are unavailable. Therefore, the dependent variables in this study are assumed to implicitly reflect the relationship of both prices and quantities (demand) in the fresh and preserved vegetable market, although this assumption can be validated given the relatively stable prices of Malaysian vegetables in recent years. Second, in the absence of prices as a result of lack of data, regional location and urbanicity dummy variables may also reflect price variations besides (real) regional and urban–rural differences in consumption (e.g., as a result of differences in tastes and habits). Future studies may also consider estimation of utility-theoretic demand systems, when prices become available in future surveys. Such studies would provide the much needed price and income elasticities of vegetables. Third, data from the household survey do not take into account the quality of the vegetables nor the type of vegetables purchased. It is acknowledged that purchases

of higher quality produce would invariably result in greater (varying) household expenditures and may not reflect larger amounts or quantities of vegetables purchased. Finally, although other information (e.g., presence of children in the household, marital status, number of working adults, preparation knowledge, working hours, and reasons for consumption and nonconsumption) may provide a more comprehensive indication of sociodemographic effects, these data were unavailable in the survey.

[Received February 2014; Accepted June 2014.]

References

Blisard, N., H. Stewart, and D. Jolliffe. *Low-Income Households' Expenditures on Fruits and Vegetables*. Washington, DC: U.S. Department of Agriculture, Economic Research Service, Agricultural Economic Report No. 833, 2004.

Blisard, N., J. Variyam, and J. Cromartie. *Food Expenditures by U.S. Households: Looking Ahead to 2020*. Washington, DC: U.S. Department of Agriculture, Economic Research Service, Agricultural Economic Report No. 821, 2003.

Cragg, J.G. "Some Statistical Models for Limited Dependent Variables with Applications to the Demand for Durable Goods." *Econometrica* 39(1971):829–44.

Department of Statistics Malaysia. *Report on Household Expenditure Survey 2009/10*. Putrajaya, Malaysia: Department of Statistics Malaysia, 2011.

Dong, D., and H. Stewart. "Household Food Purchase Patterns: The Case of Vegetables." Paper presented at the American Agricultural Economist Association (AAEA) Meetings, Orlando, FL, July 27–29, 2008.

Duan, N., W.G. Manning, C.N. Morris, and J.P. Newhouse. "Choosing between the Sample-Selection Model and the Multi-Part Model." *Journal of Business and Economic Statistics* 2(1984):283–89.

Engle, R.F. "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics." *Handbook of Econometrics*. Z. Griliches and M.D. Intriligator, eds. Amsterdam, Holland: Elsevier Publishing, 1984, pp. 775–826.

Ettienne-Gittens, R., E.L. McKyer, M. Odum, C.S. Diep, Y. Li, A. Girimaji, and P.S. Murano. "Rural versus Urban Texas WIC Participants' Fruit and Vegetable Consumption." *American Journal of Health Behavior* 37(2013):130–41.

FAOSTAT–Food and Agriculture Organization Statistics Division. 2013. Internet site: http://faostat.fao.org (Accessed June 24, 2014).

Gustavsen, G.W., and K. Rickertsen. "A Censored Quantile Regression Analysis of Vegetable Demand: The Effects of Changes in Prices and Total Expenditure." *Canadian Journal of Agricultural Economics* 54(2006):631–45.

Guthrie, J.F., and B.H. Lin. "Older Americans Need to Make Every Calorie Count." *Food Review* 25(2002):8–13.

Hadi, A.H.I.A., J. Selamat, M.N. Shamsudin, and A. Radam. "Demand for Food Safety Attributes for Vegetables in Malaysia." *EnvironmentAsia* 3(2010):160–67.

Heckman, J.J. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1979):153–1.

Jian, L, D.H. Zhang, A.H. Lee, and C.W. Binns. "Do Preserved Foods Increase Prostate Cancer Risk?" *British Journal of Cancer* 90(2004): 1792–95.

Kasteridis, P., and S.T. Yen. "U.S. Demand for Organic and Conventional Vegetables: A Bayesian Censored System Approach." *Australian Journal of Agricultural and Resource Economics* 56(2012):405–25.

Kotz, S., N. Balakrishnan, and N.L. Johnson. *Continuous Multivariate Distributions: Models and Applications*. 2nd ed. New York, NY: John Wiley & Sons, 2000.

Lee, H.L.S., and A.K.G. Tan. "Factors Affecting Malaysian Household Purchase Decisions of Food-Away-from-Home." *Journal of International Food & Agribusiness Marketing* 19(2007): 97–115.

Lin, T., and P. Schmidt. "A Test of the Tobit Specification Against an Alternative Suggested by Cragg." *The Review of Economics and Statistics* 66(1984):174–77.

Ministry of Agriculture and Agro-Based Industry Malaysia. 2014. Official portal of the Federal Agricultural Marketing Authority: Ministry of Agriculture and Agro-Based Industry Malaysia. Internet site: http://www.fama.gov.my/en/harga-pasaran-terkini2 (Accessed June 24, 2014).

Mutuc, M.E., M. Pan, and R.M. Rejesus. "Household Vegetable Demand in the Philippines: Is There an Urban–Rural Divide?" *Agribusiness* 23(2007):511–27.

Nayga, Jr., R.M. "Determinants of U.S. Household Expenditures on Fruit and Vegetables: A

Note and Update." *Journal of Agricultural and Applied Economics* 27(1995):588–94.

Newman, C., M. Henchion, and A. Matthews. "Infrequency of Purchase and Double-Hurdle Models of Irish Households' Meat Expenditure." *European Review of Agricultural Economics* 28(2001):393–412.

Niu, L., and M. Wohlgenant. "Subsidizing Fruits and Vegetables by Income Group: A Two-stage Budgeting Approach." Paper presented at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4–6, 2013.

Norimah, A.K., M. Safiah, K. Jamal, H. Siti, H. Zuhaida, S. Rohida, et al. "Food Consumption Patterns: Findings from the Malaysian Adult Nutrition Survey (MANS)." *Malaysian Journal of Nutrition* 14(2008):25–39.

Nurul Izzah, A., A. Aminah, A. Md Pauzi, Y.H. Lee, W.M. Wan Rozita, and D. Siti Fatimah. "Patterns of Fruits and Vegetables Consumption among Adults of Different Ethnics in Selangor, Malaysia." *International Food Research Journal* 19(2012):1095–107.

Reynolds, A. "Analyzing Fresh Vegetable Consumption from Household Survey Data." Southern *Journal of Agricultural Economic*s 22(1990):31–38.

Shonkwiler, J.S., and S.T. Yen. "Two-step Estimation of a Censored System of Equations." *American Journal of Agricultural Economics* 81(1999):972–82.

Smith, T.A., C.L. Huang, and B. Lin. "Does Price or Income Affect Organic Choice? Analysis of U.S Fresh Produce Users." *Journal of Agricultural and Applied Economics* 41(2009):731–44.

Stewart, H., and N. Blisard. "Are Younger Cohorts Demanding Less Fresh Vegetables?" *Applied Economic Perspectives and Policy* 30(2008):43–60.

Stewart, H., and J.M. Harris. "Obstacles to Overcome in Promoting Dietary Variety: The Case of Vegetables." *Review of Agricultural Economics* 27(2005):21–36.

Striegel-Moore, R.H., D.R. Thompson, S.G. Affenito, D.L. Franko, B.A. Barton, G.B. Schreiber, et al. "Fruit and Vegetable Intake: Few Adolescent Girls Meet National Guidelines." *Preventive Medicine* 42(2006):223–28.

Tan, A.K.G. "Demand for Food-Away-from-Home in Malaysia: A Sample Selection Analysis by Ethnicity and Gender." *Journal of Foodservice Business Research* 13(2010): 252–67.

Tey, Y.S., S. Mad Nasir, M. Zainalabidin, S. Jinap, and R. Abdul Gariff. "Demand for Quality Vegetables in Malaysia." *International Food Research Journal* 16(2009):313–27.

Tey, Y.S., M.N. Shamsudin, Z. Mohamed, A.M. Abdullah, and A. Radam. "Analysis of Demand for Vegetable in Malaysia." Munich Personal RePEc Archive (MPRA) Paper No. 15033. 2008. Internet site: http://mpra.ub.uni-muenchen.de/15033 (Accessed June 24, 2014).

Thiele, S., and C. Weiss. "Consumer Demand for Food Diversity: Evidence for Germany." *Food Policy* 28(2003):99–115.

Vuong, Q.H. "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses." *Econometrica* 57(1989):307–33.

White, J.M.L., and J.P. McFadden. "Contact Allergens in Food Ingredients and Additives: Atopy and the Hapten-Atopy Hypothesis." *Contact Dermatitis* 58(2008):245–46.

World Health Organization (WHO). *Diet, Nutrition, and the Prevention of Chronic Diseases*. Geneva, Switzerland: World Health Organization, Report of a Joint WHO/Food and Agriculture Organization Expert Consultation, Technical Report Series No. 916, 2003.

Yen, S.T. "A Multivariate Sample-Selection Model: Estimating Cigarette and Alcohol Demands with Zero Observations." *American Journal of Agricultural Economics* 87(2005): 453–66.

Yen, S.T., and J. Rosinski. "On the Marginal Effects of Variables in the Log-Transformed Sample Selection Models." *Economics Letter*s 100(2008):4–8.

Yen, S.T., and A.K.G. Tan. "Who are Eating and not Eating Fruits and Vegetables in Malaysia?" *International Journal of Public Health* 57(2012):945–51.

Yen, S.T., A.K.G. Tan, and R.M. Nayga, Jr. "Determinants of Fruit and Vegetable Consumption in Malaysia: An Ordinal System Approach." *Australian Journal of Agricultural and Resource Economics* 55(2011):239–56.

Table A1. Maximum-likelihood Estimates of Sample Selection System of Fresh Vegetables and Preserved Vegetables: Urban and Rural Households

	Urban Households				Rural Households			
	Selection		Level (RM)		Selection		Level (RM)	
Variable	Fresh Vegetables	Preserved Vegetables	Fresh Vegetables	Preserved Vegetables	Fresh Vegetables	Preserved Vegetables	Fresh Vegetables	Preserved Vegetables
Constant	0.679***	−0.451***	2.112***	−0.507***	0.950***	−0.114*	1.434***	0.464
	(0.070)	(0.040)	(0.084)	(0.174)	(0.178)	(0.062)	(0.099)	(0.336)
HH size	0.462***	0.068***	0.108***	0.088***	0.553***	0.021***	0.071***	0.048***
	(0.012)	(0.005)	(0.003)	(0.006)	(0.027)	(0.008)	(0.005)	(0.008)
Chinese	−0.231***	0.581***	0.238***	0.677***	−0.262*	0.599***	0.287***	0.460***
	(0.045)	(0.025)	(0.015)	(0.032)	(0.147)	(0.058)	(0.029)	(0.060)
Indian	−0.126	0.681***	0.224***	0.761***	−0.191	0.864***	0.309***	0.432***
	(0.086)	(0.044)	(0.027)	(0.048)	(0.237)	(0.086)	(0.052)	(0.084)
Others	0.228***	0.097**	0.028	0.251***	0.490***	−0.152**	−0.066*	0.495***
	(0.083)	(0.049)	(0.030)	(0.057)	(0.177)	(0.067)	(0.035)	(0.071)
Male	−0.391***	−0.030	0.028	−0.013	−0.537***	−0.073	0.046*	0.021
	(0.052)	(0.029)	(0.017)	(0.030)	(0.136)	(0.047)	(0.026)	(0.046)
Age 18–29	−0.426***	−0.235***	−0.161***	−0.151***	−0.433***	−0.234***	−0.114***	0.101
	(0.050)	(0.035)	(0.020)	(0.041)	(0.128)	(0.059)	(0.031)	(0.065)
Age 46–59	0.369***	0.128***	0.168***	0.140***	0.150	0.134***	0.183***	0.106***
	(0.056)	(0.025)	(0.016)	(0.026)	(0.132)	(0.040)	(0.023)	(0.041)
Age ≥ 60	0.503***	0.208***	0.227***	0.242***	0.155	0.285***	0.205***	0.130**
	(0.073)	(0.035)	(0.022)	(0.036)	(0.162)	(0.052)	(0.030)	(0.053)
White collar	0.064	−0.030	−0.021	−0.022	−0.125	0.021	0.019	−0.029
	(0.061)	(0.035)	(0.021)	(0.035)	(0.153)	(0.067)	(0.038)	(0.060)
Non/primary	0.002	−0.018	0.050***	0.067**	0.214	−0.088**	0.075***	0.036
	(0.060)	(0.029)	(0.019)	(0.029)	(0.136)	(0.039)	(0.023)	(0.039)
Tertiary	−0.171***	−0.144***	−0.126***	−0.082**	0.058	−0.075	−0.062*	0.034
	(0.051)	(0.031)	(0.018)	(0.033)	(0.135)	(0.068)	(0.036)	(0.065)
Poverty-low	−0.329***	−0.305***			−0.225*	−0.372***		
	(0.077)	(0.045)			(0.127)	(0.048)		

Table A1. Continued

	Urban Households				Rural Households			
	Selection		Level (RM)		Selection		Level (RM)	
Variable	Fresh Vegetables	Preserved Vegetables	Fresh Vegetables	Preserved Vegetables	Fresh Vegetables	Preserved Vegetables	Fresh Vegetables	Preserved Vegetables
Upper-middle	−0.067	0.102***			−0.071	0.093*		
	(0.052)	(0.025)			(0.170)	(0.048)		
High-income	−0.316***	0.058			−0.274	−0.114		
	(0.073)	(0.039)			(0.241)	(0.097)		
Log(HH inc./100)			0.343***	0.205**			0.846***	0.307*
			(0.047)	(0.082)			(0.065)	(0.175)
[Log(HH inc./100)]^2			−0.334***	−0.021			−1.055***	−0.346
÷ 10			(0.066)	(0.111)			(0.109)	(0.273)
Error standard			0.763***	1.028***			0.711***	0.896***
deviation			(0.004)	(0.025)			(0.005)	(0.030)
Error correlations								
Preserved vegetables	0.664***				0.604**			
(selection)	(0.021)				(0.051)			
Fresh vegetables	−0.227***	0.345***			−0.306***	0.326***		
(level)	(0.043)	(0.011)			(0.068)	(0.015)		
Preserved vegetables	0.335***	0.656***	0.403***		−0.148	−0.353***	0.186***	
(level)	(0.056)	(0.042)	(0.009)		(0.128)	(0.129)	(0.042)	
Log likelihood		−96997.034					−42021.237	

Note: Asymptotic standard errors in parentheses. Asterisks *** indicate statistical significance at 1% level, ** at 5% level, and * at 10% level. RM, Ringgit Malaysia; HH, household.