Demand for Food Commodities by Income Groups in Indonesia

Helen H. Jensen and Justo Manrique

Working Paper 96-WP 166
October 1996

Demand for Food Commodities by Income Groups in Indonesia

Helen H. Jensen and Justo Manrique

Working Paper 96-WP 166
October 1996

Center for Agricultural and Rural Development
Iowa State University
Ames, IA 50011

Helen H. Jensen is a professor of economics and head of the Food and Nutrition Policy Division, CARD; and Justo Manrique is an assistant professor of economics, Saint Louis University, Madrid Campus, Spain.

Journal Paper No. J-15007 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 2851.

Abstract

An analysis of the structure of demand was performed on household data classified into income groups for urban Indonesia. A demographically augmented Linearized Almost Ideal Demand System was used to estimate the structural parameters of the demand equations. Endogenous switching regression techniques yielded unbiased and consistent demand parameter estimates for the low income group, which had a large number of zeros for some food groups. Standard, seemingly unrelated, equation techniques were used to estimate the demand parameters for the other income groups. The results showed demands for the medium-high and high income households to be responsive to prices, income, and demographic variables. Demands for the medium-low income households were responsive to income and prices only. Demands for low income households were responsive to income and prices of rice and fish only.

Demand for Food Commodities by Income Groups in Indonesia

The process of liberalizing the agricultural sector is under way in many countries. Budget problems, macroeconomic imbalances, the high costs of the agricultural support programs and agreements under the General Agreement on Tariffs and Trade (GATT) are the main reasons for this change in policy. These reforms are likely to lead to food price adjustments.

Such price changes, however, can have different effects on consumers' well-being. The fact that consumption patterns vary by income level means that welfare effects also vary for different income groups when commodity prices change (Pinstrup-Andersen and Caicedo, 1978). Under these conditions, aggregate demand analysis is not very useful and may be misleading if policymakers are concerned with the effects of these adjustments on the well-being of specific target groups. Specific demand parameters by income group can be used not only to measure accurately the welfare effects caused by given price policies but also to allow the design of compensation schemes for the poor based on specific commodities (Pinstrup-Andersen et al., 1976; Pinstrup-Andersen and Caicedo, 1978; Savadogo and Brandt, 1988; Burney and Akmal, 1991).

In addition to the fact that people from different income groups have different consumer behavior, there are other reasons to estimate demand systems for different income groups instead of in aggregate.

First, it is difficult to incorporate the effects of income distribution into an aggregate demand analysis. Researchers often use average expenditure as a representative level of income and assume that the approximation error is small. This error, however, is minimized only if the expenditure distribution and the demographic composition remain relatively constant (Deaton and Muellbauer, 1980b). These assumptions generally do not hold.

In addition, income group specific demand parameters capture income class specific substitution effects that should not be ignored in policy formulation. Because consumption patterns for low-income consumers are generally less varied and hence contain fewer food items consumed than for others, approaches to estimating demand parameters that do not account for these zeros will lead to biased and inconsistent estimated demand parameters and elasticities. Conclusions based on such estimates would be erroneous and misleading.

In addition, the Linearized Almost Ideal Demand System (LA/AIDS), which was used in this paper as the specification for the demand system, does not allow for switching from necessity to normal goods. This provides a technical reason for estimating demands by income classes.

A few studies have estimated demand elasticities by income group, for example, Teklu and Johnson (1988), Jarque (1987), Savadogo and Brandt (1988), Burney and Akmal (1991), and Jones and Mustiful (1996). Most of these studies, however, do not follow a formal treatment of the household classification problem and instead take pre-established (often government defined) income groupings or classify households on an ad-hoc basis. One exception is Jarque (1987), who presented a clustering procedure to treat the classification problem more formally. Although this procedure has good theoretical basis, its practical use and its importance for policymaking are limited. It requires very precise and specific information and can produce too many socioeconomic groups (raising the question of the relevance of many of them) and groups defined in terms of several qualitative variables. Finally Jarque's procedure requires a large number of observations.

The analysis presented in this paper is based on classification of households into income groups with different consumption behavior. Households showing similar consumption behavior are classified in the same group. The procedure proposed here is easy to implement, does not need a great deal of specific information, has good statistical foundation and sets specific income boundaries for the groups.

This paper has two basic objectives: (1) to develop a procedure to classify households into income groups and (2) to analyze expenditure patterns and the structure of demand for different income groups using data for Indonesia. The plan of this paper is as follows. The second section discusses data issues and the methodology to classify households into income groups and includes a brief analysis of patterns of consumption of these newly formed income groups. We then present the almost ideal demand system (AIDS) model, which incorporates demographic variables. The econometric methodology and the problems found in the empirical estimation of demand systems for different income groups are addressed. Empirical findings are discussed and summarized.

The Data and Classification of Households in Income Groups

Data Issues

Data from the National Social and Economic Surveys (SUSENAS) of households in Indonesia were used in this study. The government of Indonesia periodically conducts these surveys to collect data related to expenditure and socioeconomic characteristics of Indonesian households. The surveys from 1981 (subround 1), 1984, and 1987 provide the basis of this study.

SUSENAS used a proportional random sample of households within a primary sampling unit (PSU); PSUs are subunits of census area segments. The selection of PSUs for these surveys was based upon a stratified sample design established for the Indonesian Census. The unit of observation for this study was a "representative" PSU household, hereafter referred to as the *household*, which was constructed by dividing the aggregate levels of selected variables (demographic and expenditures) by the number of households in that PSU. For this study, only observations belonging to the urban regions both on and off Java were used. In total, there were 3,705 "representative household" observations for urban areas on and off Java for the three time periods.

Eight commodity groups formed the basis of the analysis: rice, meats, dairy, fish, palawija products (e.g., soybeans, corn, and cassava) wheat, fruits, and other foods and nonfoods. These commodity groups had similar nutritional components or source, were important to food policy concerns, were used in past studies of the Indonesian food sector, and met the need for a parsimonious model. In this study, we used unit values (expenditures divided by quantities) as "prices" because actual prices paid were not reported in the surveys. Commodity group prices were obtained from the sum of prices of component food items that were weighted by respective budget shares.

Missing or unreported prices, required for estimating the demand system, were estimated by regressing observed prices on regional dummies and household total expenditures (see, for instance Heien and Pompelli, 1989). The estimated prices replaced these missing in the estimation of the demand system. Dagenais (1973) and Gourieroux and Monfort (1981) discuss the properties of the parameter estimates found by using data obtained in this way.

For nonfoods, quantities were not defined. Therefore, price indexes for housing, clothing, and other nonfood consumption were those computed by the Central Bureau of Statistics for the province's most important regional cities. The aggregate price for the nonfood commodity group was computed by using an average of the price indexes for housing, clothing, and other nonfood consumption.

Total expenditures, the sum of expenditures on all commodities, were used as a measure of income for classifying households into income groups and for estimating the demand system.

Classification of Households by Income Groups

Differences in household behavior, as expressed by differences in income and household characteristics, in the acquisition of goods was the fundamental criterion behind this classification. Households showing similar consumption behaviors were classified as belonging to the same income group.

For low income households, food expenditures are almost completely explained by income. For high income households, food expenditures also depend upon other factors such as household demographic characteristics (ages of household members, race, religion, education, health, employment status, geographic location, etc). For these households, the part of expenditures not explained by income is more likely to vary. In other words, when estimating food expenditures explained by income and some of these household characteristics (Engel relations), the values of the disturbances are likely to be small for low income households and large for high-income households.

The method for classifying households into income groups was based on an analysis of homogeneity of variances of residuals from these Engel regressions. The procedure has two basic steps: estimation of Engel relations and tests for homoskedasticity of variances.

Estimation of Engel relations. The objective of the estimation was to obtain residuals of sample observations from Engel regressions. First, an Engel function of the form

$$
\begin{aligned}
E_i &= \alpha_{i0}\ \mathrm{REGION} + \alpha_{i1}\ \mathrm{AS1} + \alpha_{i2}\ \mathrm{AS2} + \alpha_{i3}\ \mathrm{AS3} + \alpha_{i4}\ \mathrm{AS4} \\
&\quad + \alpha_{i5}\ \mathrm{AS5} + \alpha_{i6}\ \mathrm{AS6} + \alpha_{i7}\ \mathrm{TOTEXP} + \mu_i \qquad (1) \\
&\quad i = \text{foods, nonfoods, fish, fruits, vegetables, eggs} \\
&\quad \mu_i \sim \mathrm{iid}\ (0, \nu_i^2)
\end{aligned}
$$

was estimated for 1981, 1984, and 1987, independently, where
E_i is expenditures in commodity group i;
REGION is a dummy variable (Java = 1, Off Java = 0);
AS1 is the average number of children 1-5 years of age per household;
AS2 is the average number of children 5-10 per household;
AS3 is the average number of males 10-20 per household;
AS4 is the average number of females 10-20 per household;
AS5 is the average number of males 20 years and older per household;
AS6 is the average number of females 20 years and older per household; and
TOTEXP is the total expenditure per household.

Then, for each regression, these parameter estimates were used to obtain the corresponding residuals.

Tests for Homoskedasticity of Variances. Successive Goldfeld-Quandt tests using the residuals the Engel estimation 1 were performed in order to classify the household observations into groups with different variances. Classification of households into income groups was determined by setting income boundaries for groups of residuals.

The Goldfeld-Quandt test is based on the idea that if sample observations have been generated under the conditions of homoskedasticity, or if the null hypothesis

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_m^2 \qquad (m \le n),$$

is true (where n is the number of observations and m is the number of groups), then the variance of the disturbances of one part of the sample observations is the same as the variance of the disturbances of another part of the observations. Thus a test for homoskedasticity becomes simply a test for the equality of two variances. Moreover, because under H_0 of the two each sample variances has a chi-square distribution divided by the number of degrees of freedom, their ratio has an F distribution, provided the two sample variances are independent. The requirement that the two sample variances be independent means that two separate regression equations must be estimated, one for each part of the sample observations. Then, the test statistic is

$$s_2^2 / s_1^2 \sim F\ (n_2 - 2,\ n_1 - 2),$$

where s_i^2 is the variance for sample i, and where n_i is the number of observations in sample i.

Equation (1) was reestimated independently for each group of observations identified as having homogeneous variance. The tests were performed to see if the variances of the residuals of each adjacent pair of groups of observations were the same. If they were, then the observations in both groups were said to belong to the same income group. If they were not the same (i.e., statistically different at $\alpha = .05$), then the observations in each group were said to belong to different income groups.

Precise final boundaries were determined for every income group by repeating the Goldfeld-Quandt tests for smaller groups of residuals in the neighbourhood of two adjacent groups. This process was repeated for each survey. Then, the income groups were reconciled so that the same number of groups existed for every year. Final income groups were found by grouping the corresponding yearly income classes.

The 3,705 observations for urban zones reported in the 1981, 1984, and 1987 SUSENAS surveys were distributed by following this methodology into four income groups: low, medium-low, medium-high, and high.

Food Participation Rates

The percentage of households that reported expenditures on food groups within the PSU, defined as "participation rates," provides a good indication of expenditure patterns and is important for understanding the extent of the problem of zero expenditures for subsequent econometric analysis. Food group participation rates for urban Indonesia are presented in Table 1 for all three years.

Low participation rates in meats, dairy products and some palawija products were present for low income groups and high participation rates in all commodity groups were observed for the high income groups (Table 1). Rice was consumed by nearly all households regardless of income level.

The Basic Model

The linearized Almost Ideal Demand System (LA/AIDS) was used to estimate the structural parameters of the demand equations. Detailed derivations of the AIDS model are available in Deaton and Muellbauer (1980a,b). The general form of the derived share equation was

$$w_i = \alpha_i + \Sigma_j \gamma_{ij} \ln p_j + \beta_i \ln (X / P), \tag{2}$$

where w_i was the expenditure share of the i^{th} commodity, p_j is the price of the j^{th} commodity, X is total expenditures, and P is a price index such that

$$\ln P = \alpha_0 + \Sigma_i \alpha_i \ln p_i + \tfrac{1}{2} \Sigma_i \Sigma_j \gamma_{ij} \ln p_i \ln p_j \tag{3}$$

for the i,j=1,..., n commodities. The basic demand restrictions were expressed in terms of the model's coefficients

$$\Sigma_i \alpha_i = 1; \quad \Sigma_i \gamma_{ij} = 0; \; \Sigma_i \beta_i = 0 \qquad \text{(adding up)}$$
$$\Sigma_j \gamma_{ij} = 0 \qquad \text{(homogeneity)}$$
$$\gamma_{ij} = \gamma_{ji} \text{ for all } i \quad (i \neq j). \qquad \text{(symmetry)} \tag{4}$$

Differences in household behavior depend not only on prices and income but also on household characteristics and demographic factors. These relationships were estimated by adding parameters to the demand system; only these additional parameters depended on the demographic variables (Pollak and Wales, 1980, 1981). This demographic translating was used to incorporate demographic variables into the model so that

$$\alpha_i = \rho_{i0} + \Sigma_s \rho_{is} N_s, \tag{5}$$

where the N_s are the demographic variables (s = 1, ..., d).

For estimation purposes, the price index P was approximated using Stone's index,

$$\ln P^* = \Sigma_i \bar{w}_i \ln p_i, \tag{6}$$

where $\bar{w}_i$ is the mean of the budget share.

The resulting system is

$$w_i = \rho_{i0} + \Sigma_s \rho_{is} N_s + \Sigma_j \gamma_{ij} \ln p_j + \beta_i \ln (X / P^*), \tag{7}$$

where i = 1, ..., n and the adding-up restriction is now

$$\Sigma_i \rho_{i0} = 1; \ \Sigma_i \rho_{is} = 0; \ \Sigma_i \gamma_{ij} = 0; \ \Sigma_i \beta_i = 0 \tag{8}$$

where i = 1, ..., n and j=1,n and s = 1, ... d.

The uncompensated own, cross-price, and income elasticities for this system are

$$e_{ii} = [\gamma_{ii} - \beta_i w_i + \beta_i^2 \ln (\bar{X})] / w_i - 1, \tag{9}$$

$$e_{ij} = [\gamma_{ij} - \beta_i w_j + \beta_i \beta_j \ln (\bar{X})], \tag{10}$$

$$e_i = \beta_i / w_i + 1. \tag{11}$$

Estimation of Demand Systems per Income Group

The existence of a problem of zero expenditures for some of the commodities conditioned the methodology for the estimation of the demand systems. For the three higher income groups, standard estimation techniques were used because zero expenditures were not an important problem. For the low income group, a limited dependent variable model was used.

The Zero Expenditure Problem

As shown in Table 1, almost all households in the low income group had expenditures for rice, fruits, palawija crops, fish, other foods and non foods. The low income households often did not purchase dairy or meat commodities within the survey period. In addition, participation rates for the medium low, medium high, and high income households were generally 90 percent for all commodity groups. These facts conditioned the econometric methodology for the demand system estimation.

From a statistical viewpoint, a large number of observations at the zero expenditure share boundary causes a nonzero mean for the disturbances and a probability of zero expenditures that is not negligible. Under these conditions, standard estimation methods yield biased and inconsistent estimates of the parameters because they do not take account the nonzero mean of the disturbances (Wales and Woodland, 1983; Maddala, 1983). The problem of zero expenditures is quite frequent whenever disaggregated cross sectional data on commodity consumption are used to estimate demand systems (Wales and Woodland, 1983; Yen and Roe, 1989).

The traditional method to deal with the limited dependent variable problem has been standard Tobit analysis. The Tobit method assumes that the decision to consume a given food item is determined by the same factors that determine the amount of food to be consumed (Lin and Schmidt, 1983). But, factors explaining the probability of consuming a food commodity may differ from those that explain the quantities.

An alternative approach to deal with the zero consumption problem is a two-step decision process in which individuals first decide to consume some nonzero amount of a particular good and then, conditional on this decision, they choose the amount. This approach allows different sets of factors to explain expenditures on each outcome and different demand functions for the set of commodities when some of them are not consumed.

Along this line, Buse and Cox (1986) applied a Heckman-type sample selectivity correction to the analysis of consumption decisions as a two-step process. Others have also used the Heckman's two-step procedure to analyze demand and expenditure patterns (e.g., Cheng and Caps, 1988, for disaggregated fresh and frozen finfish and shellfish species in the United States). Nayga

(1995 a,b) also used Heckman's procedure in his studies of household expenditures on fruits and vegetables and dissagregated meat products.

Heien and Wessells (1990) used Amemiya's two-step estimator for a household-level demand system for the United States. First, they estimated a probit regression to compute the inverse Mills ratio for each household and then they used this ratio as an instrument that incorporates the censoring latent variables in the second-stage estimation of the demand relations. Wales and Woodland (1983), Lee and Pitt (1986) and Yen and Roe (1989) used econometric models based on the Kuhn-Tucker conditions for nonnegativity (or virtual prices, which are dual to the Kuhn-Tucker conditions) to estimate a two-level demand system. Recently, Burton et al. (1994) concluded that a single step approach was inappropriate in modeling the demand for meat in the United Kingdom (rejecting the Tobit approach) and that Cragg and Heckman's models provided a better representation of the factors that influenced the separate decisions of participation and expenditure level.

Another well-known alternative two-step model used in demand analysis is that Cragg (1971). Many studies on analysis of consumer behavior have been based on this model (Lin and Schmidt, 1983; Haines et al. 1988). Deaton and Irish (1984) and Keen (1986) consider other approaches to handle the zero expenditures based on the discrepancy between observed expenditure and actual consumption.

Switching regression models also provides a way to model the consumption decision as a two-step decision process (Maddala, 1983). Lee and Brown (1986), for example, used a two-stage switching regression type model to examine food expenditures at home and away from home in the United States as individuals choose to belong to one group or another; i.e., by individual self-selection. This approach is followed here.

Estimation of a Demand System for the Low Income Group

Relatively low participation rates for the dairy and meats commodity groups presented a problem of estimation for low income households. Therefore, low income households were divided into four groups or regimes: those consuming (i) all commodities; (ii) all except meat; (iii) all except dairy; and (iv) all except meat and dairy.

Four alternative regimes were identified, based upon the outcomes of the discrete choices of consumption of meats and dairy products. Decisions regarding membership in one regime or another were the result of optimizing behavior. Endogenous switching among the four regimes can occur when individuals are not randomly assigned to each regime (Maddala, 1983; Huffman, 1988; Lee and Brown, 1986).

By letting w_7 and w_8 be the share equations for meats and dairy products, all N households were classified into four mutually exclusive subsamples (S_1, S_2, S_3, and S_4) according to the discrete choices on w_7 and w_8: S_1: households in which all w's have nonzero values,

S_2: households in which $w_8 = 0$, $w_i \neq 0$, and $i=1,\ldots,7$;

S_3: households in which $w_7 = 0$, $w_i \neq 0$, and $i=1,\ldots,6,8$;

S_4: households in which $w_7 = 0$, $w_8 = 0$, and $i=1,\ldots,6$.

All observations have a nonzero probability of being assigned to one of the four subsamples or regimes. This probability is determined by evaluating the following bivariate probabilities:

$$M_{11} \equiv P(S_1) = P(w_1,\ldots,w_8 \neq 0)$$
$$= P\,[\; w_7^* = \delta_7'Z_7 + \eta_7 > 0 \;\;,\; w_8^* = \delta_8'Z_8 + \eta_8 > 0 \;],\qquad (12)$$

$$M_{10} \equiv P(S_2) = P(w_1,\ldots,w_7 \neq 0,\; w_8 = 0)$$
$$= P\,[\; w_7^* = \delta_7'Z_7 + \eta_7 > 0 \;\;,\; w_8^* = \delta_8'Z_8 + \eta_8 \leq 0 \;],\qquad (13)$$

$$M_{01} \equiv P(S_3) = P(w_1,\ldots,w_6 \neq 0,\; w_7 = 0,\; w_8 \neq 0)$$
$$= P\,[\; w_7^* = \delta_7'Z_7 + \eta_7 \leq 0 \;\;,\; w_8^* = \delta_8'Z_8 + \eta_8 > 0 \;],\text{ and}\qquad (14)$$

$$M_{00} \equiv P(S_4) = P(w_1,\ldots,w_6 \neq 0,\; w_7 = w_8 = 0)$$
$$= P\,[\; w_7^* = \delta_7'Z_7 + \eta_7 \leq 0 \;\;,\; w_8^* = \delta_8'Z_8 + \eta_8 \leq 0 \;].\qquad (15)$$

In this context w_7^* and w_8^* are unobservable variables. But, we can observe two dummy variables, $\tilde{w}_7$ and $\tilde{w}_8$, such that $\tilde{w}_7 = 1$ if $w_7^* > 0$, $\tilde{w}_7 = 0$

(otherwise), and $\tilde{w}_9 = 1$ if $\tilde{w}_9^* > 0$ or else $\tilde{w}_9 = 0$. Z_8 and Z_9 are vectors of explanatory variables, β_8 and β_9 are parameter vectors, and η_8 and η_9 are disturbance terms. Bivariate probit regressions can be used to obtain estimates of β_8 and β_9. These estimates, in turn, yield probabilities (12) through (15).

The disturbance terms of the conditional demands estimated without taking account of the probability of selection do not have a zero mean, and direct application of the standard estimation techniques will produce biased and inconsistent estimates. Adding a correction term for self-selectivity bias to each demand equation yields a new disturbance term, which has a zero mean. Probabilities (12) through (15) were used to construct estimates of selection terms for the demand equations and give the conditional demand systems corrected for selectivity bias.

For Regime 1 (S_1):

$$w_i = \alpha_{i1} + \sum_{j=1}^{8} \gamma_{ij1} \ln p_j + \beta_{i1} \ln (X/P^*) + \Sigma_s \rho_{is1} N_s$$

$$+ \tau_{i11} A_{11} / M_{11} + \tau_{i21} A_{21} / M_{21} + \varepsilon_{i1} \qquad \text{if} \quad i = 1, \ldots, 6;$$

or

$$+ A_{11} / M_{11} + \varepsilon_{i1} \qquad \text{if} \quad i = 7;$$

or

$$+ A_{21} / M_{11} + \varepsilon_{i1} \qquad \text{if} \quad i = 8; \qquad (16)$$

For Regime 2 (S_2):

$$w_i = \alpha_{i2} + \sum_{j=1}^{7} \gamma_{ij2} \ln p_j + \beta_{i2} \ln (X/P^*) + \Sigma_s \rho_{is2} N_s$$

$$+ \tau_{1i2} A_{12} / M_{12} + \tau_{2i2} A_{22} / M_{12} + e_{i2} \qquad \text{if } i = 1, \ldots, 6;$$

or

$$+ A_{72} / M_{12} + e_{i2} \qquad \text{if } i = 7; \tag{17}$$

For Regime 3 (S_3):

$$w_i = \alpha_{i3} + \sum_{j=1}^{6} \gamma_{ij3} \ln p_j + \gamma_{i83} \ln p_8 + \beta_{i3} \ln (X/P^*) + \Sigma_s \alpha_{is3} N_s$$

$$+ \tau_{1i3} A_{13} / M_{13} + \tau_{2i3} A_{23} / M_{13} + e_{i3} \qquad \text{if } i=1, \ldots, 6;$$

or

$$+ A_{83} / M_{13} + e_{i3} \qquad \text{if } i = 8; \tag{18}$$

For Regime 4 (S_4):

$$w_i = \alpha_{i4} + \sum_{j=1}^{6} \gamma_{ij4} \ln p_j + \beta_{i4} \ln (X/P^*) + \Sigma_s \alpha_{is4} N_s$$

$$+ \tau_{1i4} A_{14} / M_{14} + \tau_{2i4} A_{24} / M_{14} + e_{i4} \qquad \text{if } i=1, \ldots, 6 \tag{19}$$

where τ_k and τ_j are parameter vectors conformable to A^k and $Ee_{ik} = 0$, $i = 1, \ldots, 8$, $j = 1, \ldots, 8$, $k = 1, \ldots, 4$, and $s = d, \ldots, d$.

Finally, the adding up, homogeneity and symmetry restrictions were imposed on the system of equations (16) through (19). The restrictions now are:

$$\Sigma_i \alpha_{ik} = 1;$$

$$\Sigma_i \alpha_{iks} = 0;$$

$\Sigma_i \beta_{ik} = 0;$

for i = 1, ..., 8; k = 1, ..., 4; and s = 1, ..., d

$\Sigma_i \gamma_{ijk} = 0; \qquad \Sigma_j \gamma_{ijk} = 0$

for i = 1, ...,6 and k = 1, ..., 4

$\Sigma_j \gamma_{ijk} = 0$ where (j = 1, ..., 8 if k = 1)
(j = 1, ..., 7 if k = 2)
(j = 1, ..., 6, 8 if k = 3)
(j = 1 , ..., 6 if k = 4);

and i = 1, ..., 8.

Estimation of a Demand System for the Higher Income Groups

For the three higher income groups, the demand system represented by equation (7) subject to restrictions (4) and (8) was estimated by using Standard Iterative Seemingly Unrelated Regression Equation (ITSURE) techniques. This procedure produces maximum likelihood estimates for linear equation systems and produces parameter estimates invariant to the choice of the deleted equation. The omitted equation is the budget share of nonfood commodities.

The estimation of demand systems for the higher income groups used observations having positive expenditures only with the standard assumption of a multivariate normal disturbance distribution of the errors.

Empirical Results

To get the parameter estimates, a set of linearized AIDS models was estimated for the eight commodity groups. The variables included time and regional dummies, average number of people per age group, prices, and total expenditures. SAS was used to estimate the demand systems for the high, medium high, and medium-low income groups, and LIMDEP was used to estimate the demand systems for the low income group.

High Income Group

Most estimated parameters of the LA/AIDS were statistically significant (71 out of 112). The statistical significance of these coefficients suggests that demands were responsive to prices, income, and demographic variables. A large number of the estimated cross-price coefficients, γ_{ij}, had t-values absolutely larger than 2. All estimated own-price coefficients, γ_{ii}, were positive. All estimated β_i coefficients for the food groups were statistically significant and negative; the one for the nonfood group was positive. This implies that all food groups were classified as necessities and the nonfoods were classified as luxuries. Most demographic coefficients (36 out of 49) were statistically different from zero.

Table 2 presents the matrix of uncompensated own-price and total expenditure elasticities of demand for all income groups.[illegible] All the own-price elasticities were negative. Strictly speaking, only nonfoods seemed price elastic. The estimated elasticity for palawija foods was very close to one. Rice, the staple food, was the least responsive to own-price changes. All commodity groups were more responsive to other food and nonfood prices than to rice prices. Changes in the prices of fruits, milk, fish, and palawija crops had little effect on any demand. In general, high income households were more responsive to own-price changes than to cross-price changes.

All food demands had total expenditure elasticities less than unity (necessities). Nonfoods were the only commodity group with a total expenditure elasticity of greater than one (luxury).

Medium-high Income Group and Medium-low Income Groups

Most estimated parameters of the AIDS for these two income groups were statistically significant. Cross-price effects were statistically significant for about half of the commodity groups. For both income groups, all β_i coefficients were statistically significant and negative; thus, all the commodity groups were classified as necessities.

Most demographic coefficients were statistically different from zero. For the medium-high income group, almost all demands, except nonfoods, were affected positively by the number of children, teenagers, and adults. Fewer demographic variables were statistically significant for the medium-low income group. For this group, additional family members affected the value shares of specific food groups, principally the ones that they consumed the most (rice and fish).

In general, households in these income groups showed greater own-price elasticities and stronger cross-price effects than households in the high income group. Although the own-price elasticities were negative and generally less than one, for the medium-high income group, nonfoods, meats, and palawija crops were greater than, or nearly unity. Other foods and rice were the least responsive to own-price changes. Estimated price elasticities for the medium-low income group were generally greater (in absolute value) and showed stronger cross-price effects compared to high-income groups.

Most food demands had income elasticities smaller than unity (necessities). The nonfood group was the only commodity group having an expenditure elasticity of greater than one (a luxury good) for both income groups. Fish had a negative expenditure elasticity (an inferior good) for both groups; meat had a negative expenditure elasticity for the medium-low-income group. As observed in Table 1, as total expenditure increases, fish consumption and (especially) dry fish consumption decrease.

Low-Income Group

A bivariate probit analysis was performed to construct estimates of the correction terms for self-selectivity bias and to better understand the meat and dairy product consumption decisions of the low income households. The variables included in the bivariate probit estimated equations explaining meat and dairy consumption probabilities were number of children, teenagers, and adults; time dummies; and six regional dummies. The parameter estimates are reported in Table 3. The estimated correlation coefficient of the disturbances in the share of meat and share of dairy product equations turned out to be positive and statistically different from zero. The correlation implies that both equations were not statistically independent and that the disturbance terms were affected similarly by random shocks. Thus, the bivariate probit estimation of the participation equations is appropriate.

The results also showed that the presence of teenagers in the household increased the probability of consuming meats only. The presence of the children and adults did not have a statistically significant effect on the probability of consuming either of the commodities. This result confirmed the tendency observed for the medium-low income group: additional family members increase only the demands of those food groups that they consume the most.

Conditional demand systems including demographic characteristics, prices, income, and the correction terms were estimated for each subsample (regime) of low income households. In preliminary analysis, most demographic variables were not statistically significant and generated only small improvement in goodness of fit by including these variables (evaluated by using root mean square error and R^2s). To obtain a relatively parsimonious model, we estimated an alternative demand system, which included prices, income, and the correction terms only for the low income households.

Four conditional demand systems for the eight commodity groups with sample selection terms included were fit by SURE to each one of the four subgroups of low income households. The included variables were logs of prices, log of real income, and the correction terms.

Most of the γ_{ij} parameters were not significantly different from zero. This finding is consistent with the observed trend that shows a decreasing number of the cross-price effects to be significantly different from zero as estimation proceeds from the highest to the lowest income group. It is also interesting to note that, in most instances, the prices of fish and rice had statistically significant effects on the value shares of some other commodity groups. Fish and rice are the food groups consumed mostly by low income households.

Most β_i coefficients had t-values of greater than 2. The estimated parameters indicated foods to be necessities and nonfoods to be luxuries. For the subsample conditional on Smeat=0 and Smilk>0, fruits were also classified as luxuries. Most correction terms were significantly different from zero. This finding indicates the need to correct for the presence of self-selection bias.

In general, households in this income group showed greater own-price elasticities and stronger cross-price effects than households in the two highest income groups but smaller than those for the medium-low income group. For the four subsamples, all but one of the own-price elasticities were negative. In general, demand was price inelastic. Nevertheless, two well-defined commodity groups showed price elastic demand: palawija crops and nonfoods. And, for three of the subsamples, rice was also own price elastic

(Smeat>0 and Smilk=0, Smeat=0 and Smilk>0, Smeat=0 and Smilk=0). Some commodity demands were responsive to other commodity prices. They were affected first by rice and nonfood prices and second by other foods. Changes in the prices of fruits, dairy products, and fish had little effect on most commodity demands.

Most food demands had income elasticities less than unity. Nonfoods was the only commodity group with an income elasticity of greater than one for all subsamples. The fruit group was income elastic for the subsample conditional on Smeat=0 and Smilk>0. Palawija crops had a negative income elasticity in the subsamples conditional on Smeat=0 (both for Smilk>0 and Smilk=0).

Summary and Conclusions

Differences in consumption behavior and demand for food among income groups show the importance of estimating separate food demand parameters for income groups in Indonesia. In the first part of this paper, we presented a methodology to classify households in income groups based on the behavior of households in acquisition of goods. The methodology is based on an analysis of homoskedasticity of variances of residuals from regressions of Engel relations. Indonesian data were used to regress total expenditures and household characteristics on total food expenditures, nonfood expenditures and food group expenditures. A tabular analysis of food participation rates showed that, for Indonesia, meats and dairy products were almost exclusively consumed by high income households and that rice was consumed by nearly all households regardless of their income level. Meat and dairy product consumption patterns were used to differentiate consumption for the low income households.

Demand system parameters were estimated for each of four income groups. Household characteristics, incorporated into the basic AIDS models by demographic translating techniques, explained differences in the households' preferences for all except the lowest income group. Endogenous switching regression techniques were used to obtain unbiased and consistent AIDS demand parameter estimates for the low income group.

The results confirmed that the demand structure and the corresponding elasticities varied for different income groups. Demands for the high income households were very responsive to prices, income, and demographic variables, whereas demands for the medium-low-income households were responsive mainly to income and prices. Demands of low-income households were most responsive to income and prices of rice and fish and not responsive to the demographic variables.

In general, the estimated price and income elasticities for all income groups looked quite reasonable. The own price elasticities of demand become more price elastic (larger in absolute value) in moving from the high- to the low income groups. For all income groups, there were two price-elastic food groups: nonfoods and palawija crops. Rice was also price elastic for most subsamples of low income households. Cross-price elasticities were greater in absolute value for the low income groups. Consistently, the price of nonfoods affected all demands. Rice prices also affected all demands but especially the demands of the low income households. Nonfoods were a luxury for all income groups.

Such results have important consequences for food policy formulation and welfare analysis, particularly when income differences lead to markedly different food consumption patterns. Income group specific demand parameters can be used to make more accurate evaluation of the effects of alternative price policies on the well being of the different income groups, to design any specific target group compensation schemes based on specific food items (such as a food price subsidies, food cards), and to design policies to improve the adequacy of diets for groups at risk of nutritional deficiencies. The price sensitivity of low income households in Indonesia to rice prices both in own-commodity and cross-commodity demand suggests that increases in rice prices are likely to shift consumption of low income households toward other secondary food crops more than for high-income groups. Although some other foods may be nutritionally superior to rice, welfare losses of such price increases may be relatively large for low income households.

Table 1. Household participation rates for food expenditures by income group, urban Indonesia, all years

Food group	Low	Medium Low	Medium High	High	Total
			Percent		
Meat	68.1	90.1	95.2	98.5	90.0
Dairy	48.0	77.6	89.5	94.7	80.3
Rice	99.5	99.9	100.0	100.0	99.9
Fruits	94.5	98.6	99.3	99.7	98.4
Fish	97.2	99.7	99.7	99.5	99.3
Fresh fish	87.2	96.7	98.5	98.8	96.2
Dry fish	89.8	92.5	93.0	89.6	91.7
Palawija	98.4	99.2	99.7	99.7	99.4
Cassava	73.8	75.0	76.1	74.5	75.1
Corn	38.0	35.5	36.0	37.7	36.4
Nuts	66.6	79.5	86.1	91.7	82.1
Wheat	22.7	38.2	48.0	54.4	42.2

Note: Includes data from 1981, 1984, and 1987 SUSENAS.

Table 2. Unconditional own-price and income elasticities of demand all groups

	High	Medium High	Medium Low	Low: Smeat>0 Smilk>0	Low: Smeat>0 Smilk>0	Low: Smeat=0 Smilk=0	Low: Smeat=0 Smilk=0
Meats							
Price	-.89	-.91	-.81	-.53	-.91		
Total exp.	.69	.25	-.85	.39	.65		
Rice							
Price	-.42	-.58	-.87	-.71	-1.59	-1.67	-.98
Total exp.	.26	.10	.15	.34	.10	.71	.31
Fruits							
Price	-.59	-.77	-.83	-.75	-.73	-1.14	-.87
Total exp.	.56	.43	.45	.54	.64	1.32	.94
Milk							
Price	-.74	-.64	-.55	-.29		.33	
Total exp.	.70	.71	.23	.84		.34	
Fish							
Price	-.50	-.66	-.63	-.84	-.53	-.63	-.48
Total exp.	.22	-.82	-.34	.16	.70	.98	.56
Palawija							
Price	-.97	-1.03	-1.02	-1.09	-1.45	-2.06	-1.62
Total exp.	.65	.44	.54	.91	.42	-.40	.09
Other Food							
Price	-.88	-.51	-.83	-.97	-.77	-.79	-.58
Total exp.	.74	.86	.95	.98	.59	.63	.61
Nonfood							
Price	-1.05	-1.11	-1.26	-1.38	-1.53	-1.28	-1.46
Total exp.	1.07	1.14	1.19	1.19	1.31	1.16	1.28

Table 3. Bivariate probit explanation of participation in meats and dairy consumption

Variables	Consumption of meats	Consumption of dairy
Intercept	-.133 (-.3)[a]	-.184 (-.5)
Region3	.374 (2.2)	-.368 (-2.3)
Region5	.535 (2.7)	-.812 (-4.1)
Region6	.088 (.3)	-.105 (-.4)
Region7	-.302 (-1.3)	.978 (.4)
Region8	-.594 (-.6)	.100 (.1)
T84	.129 (.9)	.332 (2.3)
T87	.450 (3.0)	.428 (3.0)
Demo1	-.066 (-.6)	.078 (.7)
Demo2	.206 (2.0)	-.072 (-.7)
Demo3	.697 (.5)	.113 (.9)
Rho (correlation coefficient)	.434	(6.6)

[a] The numbers in parentheses are asymptotic t-ratios.

ENDNOTES

1. This aggregation was done in part to handle the large size of survey data, nearly 58,000 households for a single survey.

2. Deaton (1988) has reviewed the limitations of working with unit values instead of market prices.

3. In this paper, we used the AIDS elasticity formulas. Green and Alston (1990, 1991) derived the specific elasticity formulas for the LAIDS. However, in their empirical work, the two sets of formulas led to essentially identical elasticity estimates.

4. The values of the A's are available to interested readers from the authors.

5. The values of the cross-price elasticities are available to interested readers from the author.

REFERENCES

Burney, N. and Akmal, M. (1991) Food Demand in Pakistan: An Application of the Extended Linear Expenditure System, *Journal of Agricultural Economics*, **42**, 185-195.

Burton, M., Tomlinson, M. and Young, T. (1994) Consumers' Decisions Whether or not to Purchase Meat: A Double Hurdle Analysis of Single Adult Households, *Journal of Agricultural Economics*, **45**, 202-212.

Buse, R. and Cox, T. (1986) A Conceptual Model for Changing Food Demand, mimeo, University of Wisconsin, Madison.

Cheng, H. and Capps, O. (1988) Demand Analysis of Fresh and Frozen Finfish and Shellfish in the United States, *American Journal of Agricultural Economics*, **70**, 533-542.

Cragg, J. C. (1971) Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods, *Econometrica*, **39**, 829-844.

Dagenais, M. G. (1973) The Use of Incomplete Observations in Multiple Regression Analysis, *Journal of Econometrics*, **1**, 317-328.

Deaton, A. (1988) Quality, Quantity and the Spatial Variation in Price, *American Economic Review*, **78**, 418-430.

Deaton, A. and Irish, M. (1984) Statistical Models for Zero Expenditures in Household Budgets, *Journal of Public Economics*, **23**, 59-80.

Deaton, A. and Muellbauer, J. (1980a) An Almost Ideal Demand System, *American Economic Review*, **70**, 312-326.

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

Green, R. and Alston, J. (1990) Elasticities in AIDS models, *American Journal of Agricultural Economics*, **72**, 442-445.

Green, R. and Alston, J. (1991) Elasticities in AIDS models: A clarification and Extension, *American Journal of Agricultural Economics*, **73**, 874-875.

Gourieroux, C. Monfort, A. (1981) On the Problem of Missing Data in Linear Models, *Review of Economic Studies*, **48**, 579-586.

Haines, P. S., Guilkey, D. K., and Popkin, B. M. (1988) Modelling Food Consumption Decisions as a Two-Step Process, *American Journal of Agricultural Economics*, **70**, 543-552.

Heien, D. and Wessells, R. (1988) The Demand for Dairy Products: Structure, Prediction, and Decomposition, *American Journal of Agricultural Economics*, **70**, 219-228.

Heien, D. and Pompelli, G. (1989) The Demand for Alcoholic Beverages: Economic and Demographic Effects, *Southern Economic Journal*, **55**, 759-770.

Huffman, W. E. (1988) An Econometric Methodology for Multiple-Output Agricultural Technology: An Application of Endogenous Switching Models, in *Agricultural Productivity Measurement and Explanation*. (Eds.) M. Capalbo and J. M. Antle, Resources for the Future, Washington D. C.

Jarque, M. C. (1987) An Application of Limited Dependent Variable Models to Household Expenditure Analysis in Mexico, *Journal of Econometrics*, **36**, 31-53.

Jones, E. and Mustiful, B. (1996) Purchasing behaviour of higher-and lower-income shoppers: a look at breakfast cereals, *Applied Economics*, **28**, 131-137.

Keen, M. (1986) Zero Expenditures and the Estimation of Engel Curves, *Journal of Applied Econometrics*, **1**, 277-286.

Lee, J. and Brown, M. G. (1986) Food Expenditures at Home and Away from Home in the United States - A Switching Regression Analysis, *The Review of Economics and Statistics*, **68**, 142-147.

Lee, L. and Pitt, M. M. (1986) Microeconomic Demand Systems with Binding Nonnegative Constraints: The Dual Approach, *Econometrica*, **54**, 1237-1242.

Lin, T. and Schmidt, P. (1983) A Test of the Tobit Specification Against an Alternative Suggested by Cragg, *The Review of Economics and Statistics*, **65**, 174-177.

Maddala, G. S. (1983) *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

Nayga, R. (1995a) Determinants of U.S. Household Expenditures on Fruit and Vegetables: A Note and Update, *Journal of Agricultural and Applied Economics*, **27**, 588-594.

Nayga, R. (1995b) Microdata Expenditure Analysis of Disaggregate Meat Products, *Review of Agricultural Economics*, **17**, 275-285.

Pinstrup-Andersen, P., N. Londoño, and Hoover, E. (1976) The Impact of Increasing Food Supply on Human Nutrition: Implications for Commodity Priorities in Agricultural Research and policy, *American Journal of Agricultural Economics*, **58**, 131-142.

Pinstrup-Andersen, P. and Caicedo, E. (1978) The Potential Impact of Changes in Income Distribution on Food Demand and Human Nutrition, *American Journal of Agricultural Economics*, **60**, 402-415.

Pollak, R. A. and Wales, T. J. (1980) Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects, *Econometrica*, **48**, 595-612.

Pollak, R. A. and Wales, T. J. (1981) Demographic Variables in Demand Analysis, *Econometrica*, **49**, 1533-1558.

Savadogu, K. and Brandt, J. (1988) Household Food Demand in Burkina Faso: Implications for Food Policy, *Agricultural Economics*, **2**, 345-364.

Teklu, T. and Johnson, S. R. (1988) Demand Systems from Cross-Section Data: An Application to Indonesia, *Canadian Journal of Agricultural Economics*, **36**, 83-101.

Wales, T. J. and Woodland, A. D. (1983) Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints, *Journal of Econometrics*, **21**, 263-285.

Yen, S. T., and Roe, T. L. (1989) Estimation of a Two-Level Demand System with Limited Dependent Variables, *American Journal of Agricultural Economics*, **71**, 85-98.