Accounting for Aggregation Bias in Empirical Demand Models - The Case of Almost Ideal Demand Systems

by

Thomas I. Wahl, Ron C. Mittelhammer,

and Hongqi Shi

Suggested citation format:

Wahl, T. I., R. C. Mittelhammer, and H. Shi. 1994. "Accounting for Aggregation Bias in Empirical Demand Models - The Case of Almost Ideal Demand Systems." Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL. [http://www.farmdoc.uiuc.edu/nccc134].

Accounting for Aggregation Bias in Empirical Demand Models--The Case of Almost Ideal Demand Systems

Thomas I. Wahl, Ron C. Mittelhammer, and Hongqi Shi*

Introduction

Aggregate time series data is often used in empirical demand studies. An assumption often made is that the market demand function based on aggregate time series data is consistent with the micro demand functions of individual consumers so that the neoclassical restrictions apply to market demands. However, this assumption is only valid under restrictive conditions as pointed out by previous researchers such as Gorman (1959) and Muellbauer (1975 and 1976). According to Gorman, market demand functions expressed as a function of aggregate income are consistent with individual consumer demand functions when all consumers have identical marginal propensities to consume, which implies that the individual Engel curves are linear and parallel to each other. This is often referred to as the property of "exact linear aggregation" because the aggregate demand (per capita) function is linear in aggregate (per capita) income of all consumers. Muellbauer (1975) proposed a class of preferences called Price Independent Generalized Linearity (PIGL), and demonstrated how individual demand functions underlying this type of preference structure lead to exact nonlinear aggregation. Under PIGL preferences, demand functions need not be linear in total (or per capita) income in order to obtain consistent aggregation from individual consumer demands to market demand. A special case of PIGL preferences is its logarithmic form, called PIGLOG preferences, from which the Almost Ideal Demand System (AIDS) is derived.

One of the most important reasons for the popularity of the AIDS model among empirical demand analysts is its property of consistent aggregation across consumers. However, for AIDS demand functions to exhibit this property when demands are specified in terms of total (or per capita) income, it must be the case that income is equally distributed among consumers and the income distribution across consumers must be stable over time, i.e., no income redistribution occurs during the period of analysis. It is evident that this is a very restrictive and unrealistic assumption. However, most previous empirical demand studies have taken the aggregation property of the AIDS model for granted or simply ignored the problem entirely. It has been argued in the literature that ignoring the income distributional effect in the aggregated demand model can result in biased parameter estimates (Muellbauer 1975 and 1976, Stoker 1986, and Blundell et al. 1993). Furthermore, failure to explicitly take account of the aggregation bias can invalidate the aggregate demand model as a valid representation of the underlying individual micro-functions.

---

*Thomas I. Wahl is an Assistant Professor of Agricultural Economics, Ron C. Mittelhammer is a Professor of Agricultural Economics and Adjunct Professor of Statistics, and Hongqi Shi is a post-doctoral research associate at Washington State University.

One of the major obstacles in explicitly modelling the aggregation bias in aggregate time series analysis is the fact that most time series data sets do not contain detailed information on the distribution of micro-variables across consumers. In this paper, we propose a procedure that allows applied researchers to estimate the aggregation bias effect by utilizing accessible information on the income distribution of consumers. The goal is to obtain unbiased and/or consistent parameter estimates of the aggregate market demand functions that are valid representations of micro-functions at the individual consumer level.

Aggregation Theory

The problem of aggregation in demand analysis has received considerable attention in the economics literature. The seminal works in this area include Samuelson (1947), Theil (1954), Gorman (1953, 1959), and Green (1964). These early works focused on establishing the conditions for consistent aggregation. According to Green (1964, p.3), "Aggregation will be said to be consistent when the use of information more detailed than that contained in the aggregates would make no difference to the results of the analysis of the problem at hand." Most of the earlier studies concentrated on linear aggregation. In this case, the necessary and sufficient condition for micro demand (consumption) functions to consistently aggregate to market demand (consumption) function is that the Marginal Propensity to Consume (MPC) is identical across consumers, which leads to market demand functions that are independent of the income distribution. Under this assumption, all Engel curves are linear and have the same slope. In this case, it is sufficient to model market demand behavior by utilizing the aggregate time series data expressed in the form of total or per capita income since total or per capita income is sufficient to capture the effect of changing income levels on aggregate quantity demanded. It is evident that assuming identical MPCs is very restrictive and unrealistic and that it is more reasonable to expect that consumers in different income categories will have different MPCs for many of the goods consumed.

Muellbauer (1975, 1976) established more general conditions based on PIGL preferences for the market demand functions to be consistent with the micro functions at the individual consumer level. A major advantage of PIGL preferences is that it allows nonlinear forms of demand (and Engel) functions and yet still allows for aggregate relations at the market level to be consistent with micro relations at the individual consumer level. The PIGL preferences can be represented by the following cost function,

$$(1) \qquad c(u_o, p) = [a(p)^{\alpha}(1-u_o) + b(p)^{\alpha}u_o]^{\frac{1}{\alpha}},$$

where a(p) and b(p) are linearly homogenous functions of prices, u_o is a utility index, and α is a constant. It is evident that by choosing different α values, equation (1) encompasses a wide variety of cost structures that lead to different forms of demand function which can be derived by applying Shephard's lemma (see Muellbauer 1975, Deaton and Muellbauer 1980a). For example, when $\alpha = 1$ the Engel function takes a linear form and the linear aggregation case is obtained. When $\alpha = -1$ the Engel curves are quadratic. In this case, both the mean and variance of incomes across consumers are needed to capture the effect of income movements over time and hence both terms should be incorporated in an empirical aggregate market demand function. When $\alpha \rightarrow 0$, PIGL preferences limit to the PIGLOG form and the corresponding demand functions represent the AIDS model. In this case, the effects of income redistribution on the aggregate market demand can be captured by the mean and the entropy measure of income dispersion across consumers (Muellbauer 1975). As long as the Engel functions take nonlinear forms i.e., α is not equal to one, the simple arithmetic mean of income is not sufficient to represent the effect of changes in consumer income on aggregate quantity demanded. In this case, additional terms for capturing income effects are needed in the specification of the aggregate demand equation in order to be consistent with micro demand functions. Only when changes in incomes are equiproportional across all consumers is no aggregation error committed when estimating the market demand function utilizing the aggregate time series data (Muellbauer (1975)).

We now examine the case of PIGLOG preferences in more detail. Following Deaton and Muellbauer (1980b), a consumer-specific AIDS model derived from the PIGLOG cost function can be expressed in share form as

$$w_{ih} = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log(x_h / k_h P), \quad \forall\, i,h$$

(2) $$\log P = \alpha_o + \sum_i \alpha_i \log(p_i) + 1/2 \sum_i \sum_j \gamma_{ij} \log(p_i) \log(p_j)$$

$$where \sum_i \alpha_i = 1, \quad \sum_j \gamma_{ij} = 0, \quad \sum_i \beta_i = 0, \quad \gamma_{ij} = \gamma_{ji} \quad \forall\, i \neq j,$$

and $k_h > 0$, $\forall h$, are parameters allowing for different preference relations across consumers. The share of aggregate expenditure allocated to good i can be defined as

(3) $$\bar{w}_i = \sum_h p_i q_{ih} / \sum_h x_h = \sum_h x_h w_{ih} / \sum_h x_h,$$

where q_{ih} is the quantity of commodity i consumed by consumer h, p_i is the price of commodity i, x_h is the total expenditure of consumer h, and w_{ih} is the share of total expenditure allocated to commodity i by consumer h.

The aggregate expenditure share equation in the AIDS model can be obtained by substituting equation (2) into (3), obtaining

(4) $$\bar{w}_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i (\sum_h x_h \log(x_h / k_h P) / \sum_h x_h).$$

Letting $r_h = x_h / \sum_h x_h$ represent the h^{th} consumer's share of aggregate expenditure, equation (4) can be reexpressed as

(5) $$\bar{w}_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i [\sum_h r_h \log(x_h / k_h P)].$$

The last term in brackets in equation (5) is a function of the weighted geometric means of real expenditures and taste difference parameters of consumers using r_h as the weights, where $r_h \in [0, 1]$ and $\sum_h r_h = 1$. Letting x^* and k^* denote the geometric means of expenditures and taste difference parameters, respectively, equation (5) can be rewritten as

$$\bar{w}_i = \alpha_i + \sum_j \gamma_{ij} (\log p_j) + \beta_i \log (x^* / k^* P), \tag{6}$$

$$\textit{where } x^* = \prod_h x_h^{r_h} \textit{ and } k^* = \prod_h k_h^{r_h}.$$

Clearly the calculation of the weighted geometric mean of expenditures, x^*, in the aggregate share equation requires detailed information on the distribution of total expenditures over consumers. Unfortunately, most empirical data available for applied demand studies are measured at an aggregated level, and the information necessary for computing x^* is often not available in practice. In demand studies utilizing aggregate time series data, researchers often use the simple average of individual expenditures (i.e., per capita expenditure) to replace the geometric mean. Deaton and Muellbauer (1980a and 1980a) have shown that if the average aggregate budget share is to be specified as a function of prices and per capita expenditure, this requires the restrictive conditions of exact linear aggregation. In the case of exact (price-independent) nonlinear aggregation, such as AIDS, it is required that the aggregate budget share, $\bar{w}_i$ depend on prices and a representative level of total expenditure x_o which itself depends on the distribution of expenditures. In this case, "the market pattern of demand can be thought of as deriving from the behavior of a single representative individual endowed with total expenditure x_o and facing prices p." (Deaton and Muellbauer 1980a, pp.154). In the case of PIGLOG preferences, it is clear from (6) that the appropriate level of representative expenditure is given by $x_o = \prod_h x_h^{r_h}$. It follows that using $\bar{x}$ in place of x_o constitutes a misspecification of the AIDS model.

There is a functional relationship between the representative expenditure, x_o and the simple arithmetic mean of individual expenditures, $\bar{x}$, which can be defined based on x_o being the geometric mean of individual expenditures under PIGLOG preferences. In particular, one can write $x_o = \tau_o\bar{x}$, with τ_o being a composite measure or indicator of the expenditure distribution (dispersion) across individual consumers defined by

$$\tau_o = N / Z,$$

where N is the number of consumers and $\log Z = -\sum_h r_h \log(r_h)$ is Theil's entropy measure of the distribution (dispersion) of consumers' expenditure shares. This result can be easily demonstrated by substituting $x_h = r_h(\sum_h x_h) = r_h(N\bar{x})$ into the weighted geometric mean definition of representative expenditure, as

$$x_o = \prod_h (x_h)^{r_h} = \prod_h (r_h(N\bar{x}))^{r_h} = (\prod_h r_h^{r_h})(N\bar{x}) = \left[\frac{N}{Z}\right]\bar{x}, \tag{7}$$

$$\textit{where } Z = (\prod_h r_h^{r_h})^{-1}.$$

It can be shown that Z achieves its maximum value of N when the consumers' expenditure shares, the r_h's, are identical, i.e., $r_h = 1/N \; \forall \; h$ (Theil, 1971). In the general case where the consumer's aggregate expenditure shares are not identical, the value of Z is always smaller than N. In this case, $\tau_o = N / Z > 1$, which implies that x_o is larger than $\bar{x}$. This indicates that under PIGLOG preferences, the simple mean value of consumers' expenditures is always an under-estimate of the true value of the aggregate representative expenditure represented by the weighted geometric mean. Therefore, ceteris paribus, the parameter estimates (β_i's) associated with real expenditure are generally biased if the simple mean is used to replace the geometric mean of consumers' expenditures. Furthermore, if τ_0 is not constant over time, other parameter estimates in the AIDS model will generally be biased as well.

Regarding the geometric mean of taste change parameters, k^*, in the aggregate AIDS model (6), first note that if all consumers have the same tastes, so that $k_h = 1\ \forall h$, then $k^* = 1$ and the taste variable vanishes. Alternatively, if tastes differ across consumers, but if tastes and the distribution of income shares across consumers remain stable, then k^* is a constant that can be subsumed into the intercept term of (6), as $\alpha_i^* = \alpha_i - \beta_i \ell n\, k^*$. Finally, if tastes and/or income shares change over time, k^* will change over time, leading to the intercept α_i^* changing over time as well. Note in the latter case that even if consumers' tastes do not change over time, the fact that they are different across consumers can induce a taste effect through a changing income distribution. Modelling a changing intercept value $\alpha_i^* = \alpha_i - \beta_i \ell n k^*$ would involve incorporation of a function of time and/or socio-demographic variables in the specification of the AIDS model.

Modeling The Aggregation Bias In AIDS

From the preceding discussion, the representative expenditure in the AIDS model can be expressed as $\bar{x}^* = (N/Z)\bar{x}$, where we now use the symbol $\bar{x}^*$ to denote the geometric mean, with * emphasizing that the mean is a geometric one. By substituting this expression into equation (6), and subsuming any taste effect into the intercept term, one obtains the following aggregate AIDS model

$$\text{(8)} \qquad \bar{w}_i = \alpha_i^* + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{\bar{x}}{P}) + \beta_i[\log(N) - \log(Z)],\ \forall i$$

where log(Z) is the aforementioned Theil's entropy measure of expenditure distribution disparity, N is the size of the consumer population, and $\bar{x}$ is per capita expenditure. We refer to the entire term in brackets as the aggregation bias term, which represents an omitted variable when utilizing per capita expenditure in place of the weighted geometric mean of consumers' expenditures. From equation (8), it is evident that to calculate values of the aggregation bias term one needs time series information on the size of the consumer population and on individual consumers' shares of aggregate expenditure.

Information on the shares of aggregate expenditure across consumers is generally unavailable so the aggregation bias cannot be calculated exactly. However, time series information on the number of consumers in different income categories is readily available for most developed countries and can provide valuable information for approximating the expenditure distribution and aggregation bias term in the aggregate AIDS model. We now focus on motivating an approximation for the structure of the aggregation bias term expressed in terms of the proportions of the consumer population in different income categories. Based on this approximation, a procedure is proposed to correct or mitigate the potential aggregation bias.

First, redefine the aggregate budget share of a particular commodity for the whole consumer population in terms of a weighted average of budget shares corresponding to consumers in k different income categories as

$$(9)\qquad \bar{w}_i = \frac{\sum_k N_k p_i \bar{q}_{ki}}{\sum_k N_k \bar{x}_k} = \frac{\sum_k N_k \bar{x}_k \bar{w}_{ki}}{\sum_k N_k \bar{x}_k} = \sum_k \rho_k \bar{w}_{ki},$$

where $\bar{w}_i$ is the aggregate budget share for the i^{th} commodity for the entire consumer population, N_k and $\bar{x}_k$ are respectively the number of consumers and the per capita income in the k^{th} income category, $\bar{q}_{ki}$ is the simple average quantity of the i^{th} commodity consumed by consumers in the k^{th} income category, ρ_k is the share of aggregate income represented by the k^{th} income group, i.e., $\rho_k = N_k\bar{x}_k / \sum_k N_k\bar{x}_k$, and $\bar{w}_{ki}$ is the share of aggregate income contained in the k^{th} income group that is spent on the i^{th} commodity. The k^{th} income category budget share for the i^{th} commodity can be written as

$$(10)\qquad \bar{w}_{ki} = \frac{\sum_{h=1}^{N_k} x_{kh} w_{khi}}{\sum_{h=1}^{N_k} x_{kh}} = \sum_{h=1}^{N_k} \rho_{kh} w_{khi}$$

where x_{kh} is the income of the h^{th} consumer in the k^{th} income category, ρ_{kh} is the share of the aggregated group k income attributable to the h^{th} consumer in the k^{th} income groups, i.e., $\rho_{kh} = x_{kh}/ \sum_{h=1}^{N_k} x_{kh}$, and w_{khi} is the share of the income of consumer h in income group k devoted to good i.

Assuming temporarily that consumers have identical PIGLOG preference structures, the share of income devoted to commodity i by the h^{th} consumer in the k^{th} income category can be expressed as

$$(11)\qquad w_{khi} = \alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{x_{kh}}{P}).$$

By substituting equation (11) into equation (10), $\bar{w}_{ki}$ can be expressed as

$$(12)\qquad \bar{w}_{ki} = \alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{\bar{x}_k^*}{P}),$$

where $\bar{x}_k^*$ is the weighted geometric mean of incomes for consumers in the k^{th} income category, defined as $\log(\bar{x}_k^*) = \sum_{h=1}^{N_k} \rho_{kh}\log(x_{kh})$. Substituting equation (12) into equation (9), expresses the aggregate budget share for good i in terms of the $\bar{x}_k^*$'s as

$$\bar{w}_i = \alpha_i + \sum_j \gamma_{ij}\log(p_j) + \beta_i \log\left[\frac{\prod_{k=1}^{K}(\bar{x}_k^{*})^{\rho_k}}{P}\right] \tag{13}$$

$$= \alpha_i + \sum_j \gamma_{ij}\log(p_j) + \beta_i \log(\frac{\bar{x}^{*}}{P}),$$

where the second equality follows directly from variable definitions. Now let $z_k = \left[\prod_{h=1}^{N_k} \rho_{kh}^{\rho_{kh}}\right]^{-1}$ and $\bar{x}_k = \sum_{h=1}^{N_k} x_{kh}/N_k$, where $\log(z_k)$ is the entropy measure of the income distribution disparity for consumers in the k^{th} income category. Then the logarithm of $\bar{x}^*$ can be expressed as

$$\log(\bar{x}^{*}) = \sum_{k=1}^{K} \rho_k \log(\bar{x}_k^{*}) = \sum_{k=1}^{K} e_k \log(\prod_{h=1}^{N_k} x_{kh}^{\rho_{kh}})$$

$$= \sum_{k=1}^{K} \rho_k \log[(\prod_{h=1}^{N_k} \rho_{kh}^{\rho_{kh}})(N_k \bar{x}_k)] \tag{14}$$

$$= \sum_{k=1}^{K} \rho_k \log\left[\frac{N_k}{z_k}\bar{x}_k\right]$$

where the third equality follows from the fact that $x_{kh} = r_{kh}(N_k\bar{x}_k)$ and $\sum_{h=1}^{N_k} r_{kh} = 1$. If income is equally distributed among the consumers within each income group, then z_k achieves its maximum value of N_k so that the ratio $(N_k / z_k) = 1$. Assuming this to be the case as an approximation that improves as the number of income categories increase, equation (14) can be reexpressed in terms of $\bar{x}$ as

$$\log(\bar{x}^{*}) = \sum_{k=1}^{K} \rho_k \log(\bar{x}_k) = \log\left[\prod_{k=1}^{K} \bar{x}_k^{\rho_k}\right]$$

$$= \log\left[\prod_{k=1}^{K}\left(\frac{\rho_k(N\bar{x})}{N_k}\right)^{\rho_k}\right] = \log\left[\prod_{k=1}^{K}\left(\frac{\rho_k}{N_k/N}\right)^{\rho_k}\bar{x}\right] \tag{15}$$

$$= \log(\bar{x}) + \sum_{k=1}^{K} \rho_k[\log(\rho_k) - \log(N_k/N)]$$

$$= \log(\bar{x}) - [\log(z) + (\sum_{k=1}^{K} \rho_k \log(\lambda_k)],$$

where $\log(z) = -\sum_k \rho_k \log(\rho_k)$, is the entropy measure of income distribution disparity across the income groups and $\lambda_k = N_k/N$ is the proportion of consumer population in the k^{th} income group.

By substituting equation (15) into equation (13), one obtains the following aggregate AIDS model expressed in terms of per capita expenditure and an aggregation bias term.

$$(16) \quad \bar{w}_i = \alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{\bar{x}}{P}) - \beta_i[\log(z) + \sum_{k=1}^{K} \rho_k \log(\lambda_k)].$$

Recalling that ρ_k in equation (16) is defined as $\rho_k = N_k \bar{x}_k / \sum_k N_k \bar{x}_k$, it follows that

$$(17) \quad \rho_k = \frac{\frac{N_k}{N}\bar{x}_k}{\sum_{k=1}^{K} \frac{N_k}{N}\bar{x}_k} = \frac{\lambda_k \bar{x}_k}{\sum_{k=1}^{K} \lambda_k \bar{x}_k}.$$

Based on the above results, the aggregation bias term can be empirically modeled if information on the proportions of the consumer population and the mean income in each income group is available.

It should be noted in the above derivation that we assume total expenditure in the demand system is the same as income. In practice, researchers may be interested in a separable conditional demand system, i.e., a demand system that consists of a sub-group of goods, say food. In this case, the proportion variable ρ_k should be based on group expenditure share instead of group income share. If a priori knowledge concerning the income elasticity of expenditure is available for the different income categories, then one can re-derive (or recover) the expenditure distribution from the income distribution by scaling the mean values of income in each income categories. In recent literature, utilizing a full demand system framework in which total expenditure is equal to income has become more prevalent even if the focus of the study is on a sub-group of commodities, so that the preceding derivation of the aggregation bias term is directly applicable.

To this point we have suppressed the taste parameter k_h (recall equation (2)). Reintroducing k_h values in the preceding derivation of (16) can be viewed as altering the intercept of the share equations from α_i to $\alpha_i^* = \alpha_i - \beta_i \ell n k^*$, where k^* is the income (or expenditure) share-weighted geometric mean of the k_h's.

Empirical Application

The above theoretical framework for modeling the aggregation bias term in the case of the AIDS model is applied to Annual Japanese Family Expenditure Data and to Annual U.S. Per Capita Consumption Data. The Japanese expenditure data contains fairly detailed information on quantities and expenditures of various type of goods consumed for households in different income groups. The maximum number of income groups reported in the data is 16, which spans the period 1963 to 1979. In order to obtain as accurate an approximation to the aggregation bias as possible, the expenditure data on all 16 income groups is used in the empirical application. For Japanese data, we focus on a meat expenditure system which includes beef, pork, and chicken.

The U.S. data on per capita consumption and prices are from USDA sources spanning the years 1962 to 1989. The information on income distribution for the U.S. is obtained from Current Population Reports: Current Income. For the U.S. data we also focus on a meat expenditure system. Since time series on meat expenditure across different income groups is not available until the early 1980s, we estimate the meat sub-demand system by estimating the full demand system in which total expenditure is equal to income. In this case, we can use the aforementioned approximated income distribution to approximate the AIDS aggregation bias term.

For the Japanese data, two types of aggregate AIDS models are estimated. The first is the usual AIDS based on per capita expenditure and without the adjustment term for aggregation bias, i.e.,

$$(I) \qquad \bar{w}_i = \alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{\bar{x}}{P}).$$

The second model is the AIDS model with the aggregation bias term based on the meat expenditure distribution as

$$(II) \qquad \bar{w}_i = \alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{\bar{x}}{P}) - \beta_i[\log(z^*) + \sum_{k=1}^{K} \rho_k^* \log(\lambda_k)],$$

where $\log(z) = - \sum_k \rho_k \log(\rho_k)$ and $\rho_k = \lambda_k \bar{E}_k / \sum_k \lambda_k \bar{E}_k$ and $\bar{E}_k$ is the mean of meat expenditure for the k^{th} income category.

For the U.S. data, two models are estimated. In order to analyze changes in consumer preferences, three additional terms were added to these two models. The first term is a time trend variable which was included for capturing gradual change in consumer preference. The second term is a dummy variable serving as an intercept shifter for capturing structural break in consumer preference occurring after 1975. The third term is a trend shifter effective after 1975 for modeling the possible switch in consumption trend for the commodities considered in this study. The two aggregate AIDS models for the U.S. are presented as follows,

$$(III) \qquad \bar{w}_i = \alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{\bar{x}}{P}) + d_{i1} T + d_{i2} D_{75} + d_{i3} TD_{75}.$$

$$(IV) \qquad \bar{w}_i = \alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{\bar{x}}{P}) - \beta_i[\log(z) + \sum_k \rho_k \log(\lambda_k)]$$

$$+ d_{i1} T + d_{i2} D_{75} + d_{i3} TD_{75}.$$

An important aspect of models (III) and (IV) is an attempt to isolate the impact of structural change from that of aggregation bias on the demand system. In empirical time series demand analysis, researchers often include a time trend to capture possible gradual changes in consumer preferences, or an intercept (or trend) shifter is used to test for possible structural breaks during the period of analysis. An important issue related to testing for structural change is that the aggregation bias term may tend to be closely correlated with the time trend. In this case, structural change test results based on a model without aggregation bias correction could be misleading because the effect of the aggregation bias term may be entangled with or proxied by the time trend variables. Therefore, in order to accurately test for structural change, one also needs to isolate the impact of structural change from the impact of aggregation bias. Only in the case where aggregation bias is explicitly accounted for can one claim that a statistically significant structural change effect is genuinely representing structural change.

Iterative three stage least squares is used to estimate the above AIDS models. The instruments for the Japanese model are a set of principal components constructed from Japanese macroeconomic variables. The instruments used for the U.S. models include the consumer price index for nonfood goods and services, the yield on three month treasury bills, the lagged price of corn received by producers, the average hourly wage rate of meat packing plant workers, the consumer price index for fuel and energy, the lagged average dressed weight of slaughtered cattle, lagged fat pounds trimmed from pork carcasses, and the lagged feed conversion rate for broilers.

The parameter estimates of the Japanese models are presented in Table 1. It can be seen that the parameter estimates are similar in magnitude between the two models. Percentage differences in estimated parameter values range from 5 to 11 percent.

For practical purposes, it is the elasticities (and not the AIDS parameter values themselves) that have direct policy implications. In order to compare the two Japanese models from a policy perspective, we calculate the elasticities based on the parameter estimates of the two AIDS models. The results are presented in Table 2. The calculated elasticities are generally also very similar across the three models.

Based on the above analyses, both Japanese AIDS models produce very similar results across all dimensions of comparison. This similarity is attributable to the relative lack of variation in the aggregation bias term in this application. All told, the similarity of parameter estimates between Japanese models suggests that aggregation bias in the aggregate Japanese AIDS model may not be severe and for practical purposes can be ignored in the modelling process.

The parameter estimates for the two U.S. models are presented in Table 3 . In general, most of the parameter estimates across the two models with and without the aggregation bias term are close to each other in magnitude. The percentage difference in parameter estimates between the two models range from 0-7.2 percent for all but one of the parameters. For the coefficients associated with beef, pork, and poultry prices in the meat sub-demand system, the percentage differences are less than 2 percent. The estimated elasticities for the two U.S. models are presented in Table 4. It is evident that the two sets of elasticities are also considerably close. The similarity of parameter estimates and estimated elasticities between the two models suggests that the simple mean model without accounting for the aggregation bias term might serve as an adequate representation of the true geometric mean model. Thus, for practical purposes, it would appear that the aggregation bias term can be ignored in a time series analysis of the U.S. complete demand system.

Since the impact of aggregation bias on the parameter estimates is quite small, the aggregation bias might be expected to have little affect on the results of testing for structural change. It is evident from Table 3 that the time trend variable (TREND) is statistically significant in every AIDS equations. The joint test of the significance of all variables related to testing structural change indicates that structural change has occurred in the U.S. meat demand at any conventional level of significance. This finding is consistent with the previous study by Wahl et al. It should be noted that our findings on structural change are net of the impact of aggregation bias, although testing results from the simple mean model were all consistent with the bias-adjusted model.

Summary and Conclusions

In this paper, we reemphasize the theoretical importance of explicitly accounting for aggregation bias in aggregate demand analysis based on aggregate time series data. We argue that simply ignoring the aggregation bias, i.e., directly using the simple average expenditure in the aggregate AIDS model, will generally lead to bias in the parameter estimates of the demand model. Given the fact that most time series data is of an aggregate nature and does not directly contain the detailed information one needs to explicitly model the aggregation bias, we derive an approximation to the structure of the aggregation bias term expressed in terms of both the proportions of consumer population and mean income (or expenditure) in different income categories and propose a readily implemented procedure for explicitly modelling this aggregation bias in an aggregate AIDS model. We apply the proposed procedure for modelling the aggregation bias to Annual Japanese Family Expenditure data and U.S. per capita consumption data. In both cases, we found that there is little difference in the parameter estimates and the implied elasticities between models which ignore the aggregation bias term and the models which explicitly take account of the aggregation bias. This suggests that the aggregation bias, to the extent that it is adequately approximated by proportions of consumer population and mean income (expenditure) in different income categories, is empirically negligible. This suggests that applied researchers can directly utilize the simple mean AIDS model without taking account of the aggregation bias, i.e., express the expenditure variable in aggregate AIDS on per capita basis, for both the Japanese and U.S. data.

In the case of U.S. data, after controlling for the impact of aggregation bias, we found that structural change in U.S. meat demand had occurred. Since the impact of aggregation bias is negligible, the outcome on testing structural change from the simple mean model is rationalized. This implies that Wahl et al.'s result that structural change had occurred in the U.S. meat demand, which was based on a model that did not take the aggregation bias into account, is supported.

We caution that the preceding observations are all based on an aggregate bias term that was approximated using a discrete representation of income or expenditure distributions across a number of income categories. It is an open question as to whether the aggregation bias term would have a more pronounced effect if an improved income or expenditure distribution representation were available. We are currently researching such a refinement of the analysis.

REFERENCES

Blundell, Richard, Panes Pashardes, and Guglielmo Weber. "What Do We Learn About Consumer Demand Patterns from Micro Data?" American Economic Review 83(1993):570-597.

Deaton, A. and J. Muellbauer. Economics and Consumer Behavior. Cambridge University Press, 1980a.

______________. "An Almost Ideal Demand System," Amer. Econ. Rev. 70(1980b):312:26.

Gorman, W.M. (1953). "Community Preference Fields," Econometrica, 21(1953):63-80.

___________. (1959) "Separable Utility and Aggregation," Econometrica 27(1959):469-81.

Green, H A.J. Aggregation in Economic Analysis, Princeton University Press, Princeton, New Jersey, 1964.

Japan Bureau of Statistics. Annual Report on the Family Income and Expenditure Survey. Office of the Prime Minister, Japan.

Muellbauer, J. "Aggregation, Income Distribution and Consumer Demand." Review of Economic Studies 42(1975):525-543.

____________. "Community Preferences and the Representative Consumer," Econometrica 44(1976):979-999.

Samuelson, P.A., Foundations of Economic Analysis, Cambridge, Mass., 1947.

Stoker, T.M., "Simple Tests of Distribution Effects on Macroeconomic Equation," J. of Political Economy 94(1986):763-795.

Theil, H. Linear Aggregation of Economic Relations, Amsterdam, 1954.

Theil, H. Principles of Econometrics. New York: John Wiley and Sons. 1971.

Wahl, Tom, Ron Mittelhammer, and Hongqi Shi. (1994) "Structural Change in U.S. Meat Demand Revisited." Staff Paper, Dept. of Ag. Economics, Washington State University, Pullman, WA.

Table 1. Estimated Parameters For Three Types of Japanese AIDS Models

Parameter[a]	Model I		Model II		% Change (Parameter)
α_1	-.804639	(4.36)	-.892698	(5.30)[b]	10.94
γ_{11}	-.757227	(4.26)	-.797071	(4.83)	5.26
γ_{12}	.447021	(3.68)	.474008	(4.10)	6.04
β_1	.290385	(6.53)	.304817	(7.66)	4.97
α_2	1.124925	(7.46)	1.185056	(8.24)	5.35
γ_{22}	-.251603	(2.74)	-.267577	(2.94)	6.35
β_2	-.172790	(4.79)	-.183177	(5.44)	6.01
R^2					
Beef Equation:	.7742		.8145		
Pork Equation:	.7065		.7445		

[a]The parameters α_1, γ_{11}, γ_{12}, and β_1 are for the beef equation. The parameters α_2, $\gamma_{21}(=\gamma_{12})$, γ_{22}, and β_2 are for the pork equation. The chicken equation is omitted in the estimation, but can be recovered using the adding up, homogeneity, and symmetry restrictions.

[b]The figures in the parentheses are absolute t-values.

Table 2. Marshallian Price Elasticities and Expenditure Elasticities for the Japanese AIDS Models[a]

Model Commodity[b]	e_{ibeef}	e_{ipork}	$e_{ichicken}$	$e_{iExpenditure}$
Model I				
Beef	-2.099124	0.116572	0.167143	1.815408
Pork	0.527301	-1.023708	-0.106337	0.602744
Chicken	0.775748	-0.149350	-1.063524	0.437127
Model II				
Beef	-2.072451	0.077778	0.142870	1.851810
Pork	0.520035	-0.998737	-0.099132	0.577834
Chicken	0.759335	-0.136282	-1.038957	0.415904

[a]The symbols e refer to elasticities. Elasticities are evaluated at the sample means.

[b]Model I refers to the aggregate AIDS without the adjustment term for aggregation bias. Model II refers to the aggregate AIDS model with the adjustment term for aggregation bias based on expenditure distribution.

Table 3. Parameter Estimates for U.S. AIDS Models With and Without Accounting for Aggregation Bias

	AIDS Model W/O Correct.		AIDS Model W/ Correct.		
Parameter Name	Parameter	|T-value|	Parameter	|T-value|	% Change (Parameter)
Beef Eq.:					
Intercept	13.8781	4.639	14.2920	4.720	-2.98
$\log(P_{beef})$	0.9762	5.634	0.9662	5.632	-1.02
$\log(P_{pork})$	0.4770	3.768	0.4707	3.736	-1.32
$\log(P_{poultry})$	0.1402	5.144	0.1404	5.172	0.14
$\log(P_{Non\text{-}meat})$	-0.9514	2.597	-0.9817	2.689	3.18
$\log(P_{other})$	-0.6420	2.189	-0.5956	2.067	-7.23
log(X/P)	-1.2680	3.683	-1.2978	3.793	-2.35
D_{75}	-1.0309	8.402	-1.0337	8.557	0.27
Trend	-0.0594	6.560	-0.0580	6.300	-2.36
D_{75}*Trend	0.0775	8.561	0.0778	8.724	0.39
Pork Eq.:					
Intercept	7.7648	3.053	8.1012	3.123	4.33
$\log(P_{pork})$	0.4689	2.699	0.4680	2.673	-0.19
$\log(P_{poultry})$	0.0485	1.436	0.0484	1.417	-0.21
$\log(P_{Non\text{-}meat})$	-0.6022	1.491	-0.6211	1.522	3.14
$\log(P_{other})$	-0.3922	1.510	-0.3661	1.409	-6.65
log(X/P)	-0.8001	2.657	-0.8278	2.725	3.46
D_{75}	0.0320	0.316	0.0315	0.313	-1.56
Trend	-0.0048	0.536	-0.0037	0.400	-22.92
D_{75}*Trend	-0.0022	0.283	-0.0021	0.269	-4.55
Poultry Eq.:					
Intercept	1.3325	2.009	1.3785	2.040	3.45
$\log(P_{poultry})$	0.3777	6.522	0.3777	6.549	0.00
$\log(P_{Non\text{-}meat})$	-0.3394	2.420	-0.3427	2.416	0.97
$\log(P_{other})$	-0.2270	2.921	-0.2238	2.850	-1.41
log(X/P)	-0.1382	1.741	-0.1418	1.774	2.60

Table 3. Continued

Parameter Name	AIDS Model W/O Correct.		AIDS Model W/ Correct.		% Change (Parameter)
	Parameter	|T-value|	Parameter	|T-value|	
D_{75}	0.0910	4.358	0.0915	4.406	0.55
Trend	0.0114	3.782	0.0116	3.751	1.75
D_{75}*Trend	-0.0076	4.688	-0.0077	4.710	1.32
Non-Meat Eq.:					
Intercept	60.7776	6.473	61.9649	6.535	1.95
$\log(P_{Non\text{-}meat})$	5.9675	4.369	5.8815	4.284	-1.44
$\log(P_{other})$	-4.0744	3.836	-3.9361	3.737	-3.39
log(X/P)	-4.5716	4.224	-4.6460	4.296	1.63
D_{75}	-0.1275	0.366	-0.1280	0.373	0.39
Trend	-0.0880	3.228	-0.0835	2.972	-5.11
D_{75}*Trend	0.0087	0.355	0.0091	0.376	4.60
R^2 :					
Beef	0.995		0.995		
Pork	0.983		0.983		
Chicken	0.979		0.979		
Non-Meat	0.985		0.985		

Note: All coefficients are multiplied by 100 for ease of presentation. $D_{75}=1$ for 1963 to 1975 and =0 otherwise. Trend=1,2,3....... .

Table 4. Marshallian Price and Expenditure Elasticities for the U.S. AIDS Models[a].

	AIDS Model Without Accounting for Aggregation Bias					
	Beef	Pork	Poultry	Non-meat	Other	Expend.
Beef	-0.6312	0.1801	0.0535	-0.2734	0.1336	0.5375
Pork	0.3783	-0.6365	0.0398	-0.3600	0.1849	0.3934
Poultry	0.2876	0.1006	-0.2441	-0.6340	-0.2334	0.7239
Non-meat	-0.0519	-0.0340	-0.1987	-0.5797	-0.0276	0.7131
Other	-0.0104	-0.0061	-0.0033	-0.0648	-1.0007	1.0853

	AIDS Model With Accounting for Aggregation Bias					
	Beef	Pork	Poultry	Non-meat	Other	Expend.
Beef	-0.6346	0.1780	0.0536	-0.2827	0.1591	0.5266
Pork	0.3741	-0.6369	0.0399	-0.3709	0.2214	0.3724
Poultry	0.2882	0.1005	-0.2442	-0.6394	-0.2218	0.7167
Non-meat	-0.0536	-0.0351	-0.0200	-0.5844	-0.0152	0.7084
Other	-0.0099	-0.0058	-0.0033	-0.0634	-1.0047	1.0870

[a]Elasticities are evaluated at the sample means.