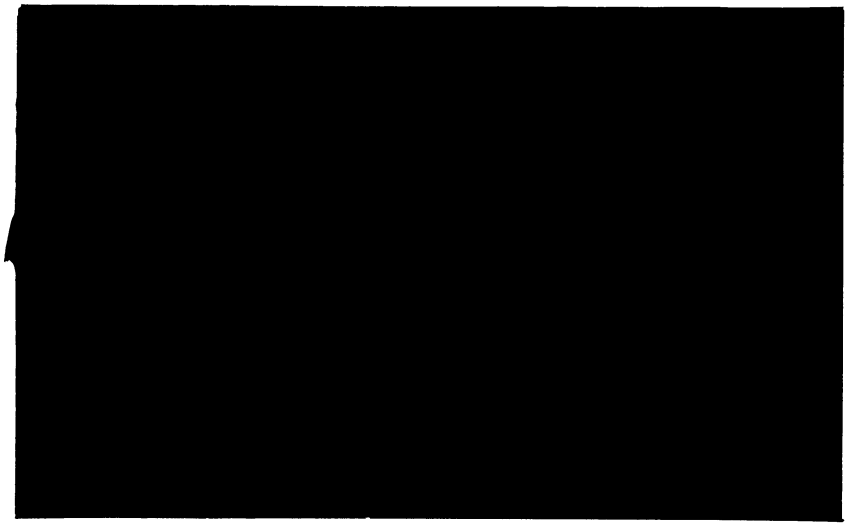
MD

91-05

Working Paper

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS
SYMONS HALL
UNIVERSITY OF MARYLAND
COLLEGE PARK 20742



ENGEL'S LAW AND LINEAR IN MOMENTS AGGREGATION

by

Robert G. Chambers and Rulon D. Pope*

Department of Agricultural and Resource Economics
University of Maryland
College Park, Maryland 20742

Working Paper No. 91-05

February 1991

*Respectively Professors, University of Maryland and Brigham Young University.

Engel's Law and Linear in Moments Aggregation

Robert G. Chambers and Rulon D. Pope*

*Respectively Professors, University of Maryland and Brigham Young University.

Engel's Law and Linear in Moments Aggregation

The estimation of demand relationships is central to the practice of agricultural economics. Though in theory such analyses should often be straightforward, seldom is the empirical researcher that fortunate. The first main problem is that data are often limited. For example, time-series data typically are aggregated well beyond the behavioral response postulated in theory and seldom have available the variety of socioeconomic data needed to permit consistent disaggregation. But aggregate time series data generally do contain significant price variation which most cross-sections lack. Thus, empirical agricultural demand analysis is typically carried out either at the very highest levels of aggregation (Blanciforti and Green (1983a), (1983b); George and King; Huang) or in the absence of price variation. Only rarely have researchers had available a consistent set of panel data and even more rarely have researchers been able to integrate disparate cross-section and time-series data sets (for an attempt at the latter see Jorgenson, Lau, and Stoker).

The nexus between the typical cross-section or microeconomic approach to demand estimation and the more aggregate time-series approach is the aggregation problem. Generally one seeks ways of consistently aggregating individual (cross-section) relationships so that aggregate (time-series) variables can be used in econometric analysis of aggregate models based on microeconomic principles. For example, Gorman early asked under what condition does aggregate demand (defined as the sum of individual demands) depend only upon per capita income and price? This is equivalent to asking when can one safely ignore the income distribution in analyzing aggregate demand? The answer is well known: "only the Gorman Polar Form exactly aggregates in the sense described above." But the Gorman Polar Form is linear in income and thus contradicts Engel's law. Thus, for a commodity like food, where Engel's law has been repeatedly verified, meaningful consistent aggregation requires that one look beyond Gorman's results.

Muellbauer discovered a coherent system of micro budget shares, aggregate income, and an aggregation rule under the assumption that the representative income index depends only on the distribution of income across individuals. This form was labelled the price independent generalized linear (PIGL) equation. It yields nonlinear Engel curves and thus generalizes the Gorman Polar Form while permitting consistency with Engel's law. The most prominent subclass of this form is the "Almost Ideal Demand System" (AIDS) developed by Deaton and Muellbauer. The AIDS model has enjoyed increasing acceptance in agricultural economics as a model of both individual and aggregate behavior (Blanciforti and Green (1983a), (1983b); Eales and Unnevehr; Fulponi; Green and Alston). However, the AIDS model as it is typically applied to aggregate data is only consistent with aggregation under some exceedingly restrictive assumptions, namely that the expenditure and the demographic distributions remain stable over the entire sample (formally, the requirement is that the number of households in the population divided by Theil's entropy measure (p. 639) of income dispersion is constant). For some of the longer data series to which this model has been applied (e.g., the entire postwar era as in Blanciforti, Green, and King) this presumption seems *a priori* implausible. This restriction is necessary because the PIGL representative expenditure is not typically available for aggregate data. Hence, the appropriate PIGL representative expenditure must be replaced by a surrogate which is typically taken to be aggregate expenditure (expenditure or income per capita) under the presumption that this surrogate is proportional (where the degree of proportionality is constant) to the appropriate index. Thus the primary theoretical and empirical advantage of the AIDS model, that it is sensitive to changes in the distribution of income, is typically assumed away in most aggregate AIDS models. For while the AIDS model can yield non-linear Engel curves in the aggregate and hence aggregate behavior seemingly in accordance with Engel's law, this same model also violates the essence of Engel's law. Consider the simple thought experiment where a fixed amount of income was transferred from the very richest individual to the very poorest. (This changes the income entropy.) Engel's law would imply that aggregate food consumption would rise (this after all may be the basic premise of the Food Stamp Program). But the most naive version of the aggregate AIDS model, which took the entropy as constant, would suggest no change.

This paper constructs a simple food consumption model which is consistent with Engel's law at both the individual and the aggregate level. Doing so required an aggregate model that depends on the income distribution in an empirically tractable manner. Our starting point is the literature on aggregation with multiple indices. Drawing on the multiple index aggregation literature (Gorman (1981); Lewbel; Lau) this paper proposes a linear in the moments (LIM) of income aggregation procedure. The model rests on the recognition that aggregate demand for any commodity ultimately depends on the distribution of income across individuals and prices. Hence, an obvious alternative is to model aggregate demand in terms of prices and something that fully characterizes the distribution of income, i.e., its moments.

Our empirical analysis examines cross-sectional demand as the first logical stage in building a coherent structure. Indeed, as Chambers and Pope

(1990) demonstrate, the first step in the proper specification of an aggregate demand equation must begin with a study of the nature of micro responses from an aggregation perspective. The empirical goal thus is to analyze micro demands in order to determine empirically the number of moments of income which are needed to describe aggregate demand consistently.

Using the 1985 Consumer Expenditures Survey, and the LIM model, our empirical evidence indicates that at least two moments of after tax income are needed to portray aggregate food demand. Therefore, consistent aggregate representation of aggregate food demand response requires measuring at least the mean and variance of income. Because these indices are commonly available (although aggregation studies usually use only the mean the variance is easily approximated from available decile data or income), a simple aggregation structure is suggested which is more tractable than Muellbauer's aggregate PIGL model.

II. Aggregation Results and Interpretation

The linear in moments (LIM) model was anticipated by Gorman's seminal work. For a given commodity, his results led to micro and macro Engel curves respectively:

(1) $X_j = \alpha_j(p) + B(p)\ y_j \qquad j = 1, \ldots, N$

(2) $\bar{X} = \bar{\alpha}(p) + B(p)\bar{y}$,

where X_j is the j^{th} individual food demand, p is an m vector of common prices, y_j is the j^{th} individual's income, N is the number of consumers, and bars denote averages (e.g., $\bar{\alpha}(p) = \sum_j \alpha_j(p)/N$). In particular, the macro form is linear in the first empirical moment of the income distribution (mean income).

As it turns out, Gorman's (1953) model is a special case of the more general aggregation problem which is to find continuous functions $\phi^1(y)$, ..., $\phi^K(y)$ $(K < N)$, and $\bar{X}(\phi^1, \ldots, \phi^K; p)$ where $y = [y_1, \ldots, y_N]$ such that[1]

$$\bar{X}(\phi^1, \ldots, \phi^K; p) = \sum_{j=1}^{N} X_j.$$

The general solution to this problem must be expressible as

$$(3) \quad X_j = \alpha_j(p) + \sum_{k}^{K} \beta_k(p)\,\tilde{\phi}_k(y_j) \qquad j=1,\ldots,N;$$

$$(4) \quad \bar{X} = \alpha(p) + \sum_{i=1}^{N} \sum_{k=1}^{K} \beta_k(p)\,\tilde{\phi}_k(y_i) \quad .$$

Our approach is similar but is only meant to apply to a single food demand equation (and thus is more robust). We require that there exist a continuous function relating aggregate food demand to the income distribution. Because if finite moments exist they completely characterize the income distribution in most circumstances (Rao, p. 106), we take this to mean that aggregation requires the existence of a continuous function $\hat{X}(\mu_1, \ldots, \mu_K; p)$ $(K < N)$ where $\mu_k = \sum_{j=1}^{N} \frac{y_j^k}{N}$ such that

$$(5) \quad \hat{X}(\mu_1, \ldots, \mu_K; p) = \sum_{j=1}^{N} \frac{X_j(y_j, p)}{N} \quad .$$

In other words, we require the first K empirical moments of the income distribution to characterize aggregate demand.

The solution to equation (5) (see Appendix A for a deviation) is expressible as

$$(6) \quad \begin{aligned} &(a) \quad \frac{X_j(y_j, p)}{N} = \alpha_j(p) + \sum_{k=1}^{K} \beta_k(p)\,y_j^k \\ &(b) \quad \hat{X}(\mu_1, \ldots, \mu_K; p) = \alpha(p) + \sum_{k=1}^{K} \beta_k(p)\,\mu_k \end{aligned}$$

with $\alpha(p) = \sum_{j=1}^{N} \alpha_j(p)$. Equation (6) represents the LIM model.

Taking expectations conditional upon p in (1), one finds, however, that

$$E(X_j|p) = \alpha_j(p) + B(p)E(y_j|p) \quad j=1,\ldots,N.$$

Thus, one sees immediately that individual demand is linear in the zeroth and first moments of income as a result of using the corresponding empirical moments in the aggregate relationship. Taking conditional expectations in the more general model gives

$$E(X_j|p) = \alpha_j(p) + \sum_{k=1}^{K} \beta_k(p)E(y_j^k)$$

which shows that the individual regression equations are linear in the moments of y. Equation (6a) also has the empirical advantage that it is K-order flexible in the sense of Diewert. Moreover, it can always be interpreted as the K^{th} order expansion around zero so it will be able to approximate any smooth function to an arbitrarily close degree. Hence, the empirical model has extremely attractive approximate properties.

However, it is often more convenient and interpretable to have a macro representation of (6) involving central moments. Yet, Lau has argued that all aggregate indices must be additively separable saying, "... non-additive symmetric functions . . . are not admissible in the aggregate demand function."[2] Thus variance and any other higher central moment is symmetric but not additively nor even strongly separable and would appear to be ruled out. However, Chambers and Pope (1990) have shown that a central moment representation similar to (6) is always possible. For two moments, (6b) becomes

(6b′) $\bar{X} = \bar{\alpha}(p) + B_2 V + \bar{y}[B_1(p) + B_2(p)\bar{y}]$

where V is the variance. Therefore, the necessary condition that either (6b') or (6b) be an appropriate macro aggregate is that micro demands be of the form (6a). For three moments (6) becomes

$$(6b") \quad \overline{X} = \overline{\alpha}(p) + V[3B_3(p)\overline{y}^2 + B_2(p)] + S\,B_3(p) + \overline{y}[B_1(p) + (B_2(p))\overline{y} + B_3\overline{y^2}]$$

where S is the third central empirical moment. Using binomial expansions one can convert demand functions written in terms of moments about zero into their central moment counterpart for K of any finite dimension.

III. Empirical Results

The empirical analyses here use the 1985 Continuing Consumer Expenditure Surveys to implement the LIM model.[3] Because transfers such as food stamps are not central to our purpose here (and involve substantial controversy, Devaney and Fraker; Buse and Chavas), and because negative incomes squared are indistinguishable from like positive incomes squared, observations with incomes less than 0 are omitted. Further, in order to reduce measurement error and/or to estimate more of a consumption or use function, all observations with less than 0 expended are omitted.[4] The latter truncation avoids the use of Tobit or other such estimators but seems prudent given data limitations (Smallwood, p.8). Therefore, the relevant sample size consists of 9561 households. Attention will focus only on the sample and not a weighting procedure to obtain U.S. estimates (and the accompanying controversy on heteroskedasticity, e.g., Buse and Chavas).

No observations on prices are recorded and no attempt was made to connect with other possible data sources to obtain some notion of price variation (Cox and Wohlgenant). The data, however, do permit the inclusion of many possible socio economic variables. Those included in econometric studies vary by the specific commodity studied. These range from the inclusion of only family size (Lewbel) to a large number of socio-economic variables (e.g., Guilkey, Haines, and Popkin). Across a wide number of studies, family size appears to be the most central and significant of these variables. Our own investigation substantiates this conclusion. Since the focus here is more directly on the impact of income on choice, we will concentrate on the parsimonious model which includes only family size as the lone socio-economic variable. Appendix B reports estimates of a model with a more extended list of variables and similar income effects.

A second issue is the proper definition of income. Ideally one would want a measure of permanent income. Given the available data, the options are pre- or post-tax annual income. After tax income seems conceptually preferable and there are a large number of missing values for pre-tax income. Though we have explored pre-tax income with results similar to those reported here, only after-tax results are included here.

Consistent aggregation in the LIM model and other models of this variety requires that family size be incorporated in the individual-specific $\alpha_j(p)$ terms. This variable is included in linear form (in Appendix B the socio-economics variables included are primarily dummy variables). Further, only $K \leq 3$ will be considered. This will enable the aggregate to capture the first three and most readily interpretable central moments of the income distribution. Therefore, the empirical regression reported here is:

(7) $X_j = \alpha_0 + \alpha_1 fs_j + B_1 y_j + B_2 y_j^2 + B_3 y_j^3 + \varepsilon_j \quad j = 1, \ldots, N$

where X_j is food expenditures normalized by its mean, y_j is after tax personal income normalized by its mean, and ε_j is the disturbance (j=1,...,N).

The distribution of ε_j (j=1,...,N) (as is conventional) is taken to be heteroskedastic of the form variance(ε_j) = $ky_j^\theta fs_j^\lambda$ (j=1,...,N). The error sum of squares in (7) was minimized at approximately $\theta = 0$ and $\lambda = 1$. Rather than calculate the more complicated maximum likelihood estimator and the information matrix, weighted least squares results are presented under the exact assumption that $\varepsilon \sim (0, \sigma^2 I_N fs)$, where ε and fs are N vectors, σ^2 is the homoskedastic portion of the variance, and I_N is the identity matrix of dimension N. Thus, all of the variables in (7) are divided by the square root of family size and the resulting equation is estimated by ordinary least squares. Appendix B also contains conventional Tobit estimates with the extended set of socio-economic variables. Weighted least squares and tobit results are similar.

Although the error structure here appears valid, it fails White's test for heteroskedasticity. In a sample of nearly 10,000 observations, this is unsurprising. However, random samples of approximately 1000 easily "pass" the test at the .01 significance level. Thus, a conservative approach is followed here by reporting White's heteroskedastic consistent estimators. This yields consistent estimates of both the parameters and their standard errors even in the presence of heteroskedasticity.

The primary focus of our empirical effort is to determine how sensitive the aggregate should be to the distribution of income, i.e., how many empirical moments are needed to characterize the aggregate data accurately. To this end, the model was estimated both with and without the y^3 term. For the more general model (Model 1), all estimated coefficients are significantly different from zero except $\hat{B}_3$. The probability value for $\hat{B}_3$ is .14 and thus, at conventional significance levels, the hypothesis that B_3 is zero would not be rejected. Model 1 will be neglected in further discussions.

The coefficient estimates for Model 2 in Table 1 have reasonable signs and imply reasonable curvature properties. Food expenditures are increasing in family size and concave in income. Thus, the data support Engel's law. These results indicate that aggregate food demand should be modelled using the first two empirical moments of income.

Corresponding to (6b'), our estimate of aggregate food consumption is defined by:

$$\bar{X} = .721 - .013\ V + \bar{y}\ [.307 - .013\ \bar{y}]$$

Therefore, a <u>ceteris paribus</u> increase in the variance of income reduces aggregate food consumption. This result follows from the concavity of micro demands in income (Engel's law) and contrasts markedly with the results that would be obtained from the typical aggregate AIDS representation. By Jensen's inequality any mean preserving spread in the distribution of income will of necessity lower average or total food consumption if our model results reflect empirical reality.

To illustrate, the variance of normalized income for our sample is .779 while the mean is one. A 10% rise in the variance of income (holding the mean constant) will then reduce total expenditure by $N\ \partial\bar{X}/\partial V = 9561(.10\ V)\ .013 = 12.43$. In contrast a policy which leads to a 10% more egalitarian distribution would increase food expenditure by the same amount.

Table 2 presents some simple numerical calculations to illustrate the relationship between mean income, variance and total consumption. Although the effect of dispersion is empirically significant, the largest relative