MD

92-01

Working Paper

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

[redacted]

PIGLOG AGGREGATION

AND

ENGEL'S LAW

by

Robert G. Chambers and Rulon D. Pope*

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

Working Paper No. 92-01

January 1992

*Professor of Agricultural and Resource Economics, University of Maryland, and Professor of Economics, Brigham Young University, respectively.

This final file not done in here.

PIGLOG Aggregation and Engel's Law

For at least four decades, economists have been conc
useful, if not ideal, index numbers (e.g., Fisher, Gorman
This research was born of a desire to develop not only conc[illegible] but also empirically superior ways to measure aggregate response and welfare. Because the most prevalent aggregate measure is the simple sum or mean of micro data, aggregation theory often led to the nihilistic conclusion that existing measures or indices were inadequate. A major proponent of a more prospective approach is Gorman: He has systematically argued that economists should consider the calculation of "economic indices" rather than "mechanical indices" (Gorman, 1988). The difference between the two is that the former are developed to be consistent with the behavior of microeconomic agents whereas the latter are usually convenient and even reflective statistics constructed for a purpose other than measuring aggregate response or welfare. The mean is a good example of such a mechanical index. It is a useful summary measure of the location of the distribution of income. But for most aggregation problems, it is only a meaningful economic index if individual responses are linear with a common, constant marginal or slope effect with respect to the variable being aggregated (Gorman, 1953).

Deaton and Muellbauer (1980a), in pursuing economic indices, estimate systems of aggregable share equations which possess nonlinear Engel curves using the Almost Ideal Demand System (AIDS). The AIDS is a special case of a family of share equations which Muellbauer calls the Price Independent Generalized Log (PIGLOG) system. The associated aggregate income index is independent of consumer prices, relatively easy to compute, and has some other very attractive properties (Muellbauer, 1975). The aggregation criterion employed is that a commodity's aggregate budget share (total expenditure on the good divided by the sum of incomes) be representable as an aggregate function of an aggregate income index and consumer prices. The aggregate function is interpreted as an aggregate or representative budget share. Simple algebraic manipulation, however, allows reinterpreting Muellbauer's aggregate quantity as a weighted average of individual budget shares where the weights are each individual's share in total consumer income.

Because aggregation rules are inherently arbitrary, alternative aggregation rules should be considered: For example, one can plausibly specify a weighted average of individual budget shares as an aggregation rule, following Muellbauer, or just as plausibly one could also consider aggregate expenditure as a weighted average of individual expenditures. Weighted aggregation of expenditures or budget shares place different behavioral restrictions on individual behavior and thus are observationally different rules. Hence, one can discriminate empirically whether a weighted average of expenditures or budget shares is better supported by data.

This paper uses this observation to propose a simple nested procedure to test the appropriateness of Muellbauer's aggregation rule using microeconomic data. Thus, unlike virtually every other study on aggregation this paper does not impose an arbitrary aggregation rule and then force the data to accommodate it. Rather, Muellbauer's aggregation rule is generalized by adding a single parameter to individual demands and then tested statistically. For parsimony, however, we preserve the Muellbauer aggregate income index unchanged.

Our empirical application is the U.S. Engel curve for food. Because food expenditure is a relatively large proportion of income, it has had extensive study (e.g., Smallwood). However, relatively little attention has been paid to aggregation issues [see Deaton and Muellbauer, 1980b]. This is unfortunate because Engel's Law implies aggregate food demand must depend upon more than mean income. Aggregate food demand also should depend on the dispersion of income if Engel's law is valid. We concentrate on a single equation approach to the aggregation of Engel curves. The single-equation approach is especially useful because of the extensive amount of empirical work on food Engel curves done in a single-equation context. Moreover, the maintained assumptions of a single-equation approach are very weak because the maximization hypotheses implicit in a systems approach place cross-equation restrictions (often empirically rejected, e.g., Deaton and Muellbauer, 1980b) on each aggregable equation.

In what follows, we first introduce our notation, generalize Muellbauer's aggregation rule, and derive a micro-form consistent with the generalized aggregation rule (a generalized PIGLOG). We then specify a parametric representation of the form and suggest a nested testing procedure. The approach is then applied to survey data on food expenditures. The results support the generalization of the generalized PIGLOG equation introduced here. That is, the Muellbauer rule for aggregation of shares or expenditures is rejected by the data in favor of a more general specification. We close with an illustration of the effects of a mean preserving spread in the distribution of income on aggregate quantities to illustrate the importance of income distribution changes on food expenditures.

II. Notation and Proposed Test

The single product analogue of Muellbauer's (1975) aggregation rule is

$$(1) \qquad \bar{w}_j(y_{oj}, p) = \frac{\sum_{h=1}^{H} p_j X_{hj}}{\sum_{h=1}^{H} y_h} = \frac{\sum_{h=1}^{H} y_h w_{hj}}{\sum_{h=1}^{H} y_h}$$

for $j=1,\ldots,m$, where y_{oj} is the aggregate income index, p_j is the price of good j an element of the m vector p, X_{hj} is the demand by consumer h for good j, y_h is income or total expenditure of consumer h, and w_{hj} is the budget share of consumer h for good j, m is the number of goods and H is the number of consumers and $\bar{w}_j$ is a monotonic function (Equation (1) differs from Muellbauer's systems approach only in that (1) must hold for only commodity j rather than for all in commodities. Second, y_{oj} in Muellbauer is independent of j because a single index is used for all commodities). Muellbauer's aggregation rule creates an aggregate expenditure share, where the last expression in (1) indicates an interpretation of (1) as a weighted average of micro budget shares but not of expenditures. It is this latter interpretation of exact aggregation that leads to the interpretation of the function on the left side of (1) as an average or representative budget share. Indeed treating $y_h/\sum_h y_h$ as the relevant probability of y_h leads to the interpretation of $\bar{w}_j$ as an expected budget share. Aggregation requires the search for a reasonable y_{oj} and $\bar{w}_j$ which satisfies (1). As is apparent, the search inextricably depends on the micro responses X_{hj} and w_{hj}.

A simple y_{oj} is average income, $\bar{y}$. Gorman (1953) showed that when the right side of (1) is simply the numerator and $y_{oj} = \bar{y}$, then micro demands must be of the Gorman Polar Form. Engel curves are linear though they do not necessarily emanate from the origin (affine in income). In fact, even if $\bar{w}_j$

is retained when y_{oj} is still chosen as average income, micro demands must still assume the Gorman Polar Form. To see this, multiply (1) by $\sum_{h=1}^{H} y_h$ and set $y_{oj} = (1/H) \sum_{h=1}^{H} y_h = \overline{y}$

$$(1') \quad \sum_{h=1}^{H} y_h \overline{w}_j (\overline{y}; p) = p_j \sum_{h=1}^{H} X_{h_j}(y_h, p) .$$

This is a Pexider equation, where $\bar{w}_j$ and X_{h_j}: $R_+^{m+1} \rightarrow R$, and whose solution is (Aczel, Theorem 3.1):

$$p_j X_{hj}(y_h; p) = \gamma_{hj}(p) + \beta_j(p) y_h \qquad h = 1, \ldots, H$$

which is the Gorman Polar Form.[1] In order to highlight this implication, it is stated as the following result.

Result 1: Given that the measure of aggregate income is the mean, then demands aggregate as in (1) if and only if they assume the Gorman Polar Form.

Thus, any attempt to do aggregate work using mean income in (1) implies that a mean preserving spread in the distribution of income will not alter aggregate demand or the aggregate budget shares.

In contrast, the more general approach in (1) with $y_{oj} \neq \bar{y}$ allows for nonlinear Engel curves, an empirical "law" for many commodities such as food [e.g., Houthakker (1957), Blanciforti and Green (1983), Deaton and Muellbauer (1980b)]. Substantial effort has been devoted to the search for alternative income indexes. The PIGLOG index used in the AIDS is among the most popular alternatives where

$$\ln y_{0j} = \frac{\sum_h y_h \ln(y_h/c_{hj})}{\sum_h y_h} .$$

We adopt this index throughout this paper. However, at issue here is the appropriate aggregation rule.

A reasonable but alternative aggregation rule to (1) is

$$\bar{E}_j(y_{oj}, p) = \frac{\sum_{h=1}^{H} y_h p_j X_{hj}}{\sum_{h=1}^{H} y_h} \tag{2}$$

for some j. In this case, the interpretation of (1) as a weighted average of budget shares is extended to expenditures. Equation (1) does not imply (2) and it is an empirical matter as to whether (1) or (2) is the more appropriate description of the data. It may well be that (1) is descriptive for some goods while (2) is appropriate for other goods.[2]

Here the focus is on a parsimonious extension of (1) which allows (2) as a special case. A simple generalization replaces y_h in the numerator of (1) or (2) with an arbitrary transformation of y_h, $g_h(y_h)$. For example, when $g(y_h)$ in (2) is 1, then Muellbauer's rule (1) is obtained for that good. When $g_h(y_h) = y_h$, then the expenditure weighted average in (2) is obtained. In other cases near these extremes, one obtains "nearly" weighted averages of expenditures or budget shares. Therefore, the focus here is on the existence of a $y_{oj}(y_1,\ldots,y_H)$ and $\bar{z}_j$ such that

$$\bar{z}_j(y_{oj}(y_1,\ldots,y_H), p) = \frac{\sum_{h=1}^{H} g_h(y_h)\, z_{hj}(y_h, p)}{\sum_{h=1}^{H} y_h}, \tag{3}$$

where z_{hj} is any arbitrary micro response function of interest. Since the approach varies substantively from Muellbauer, the Appendix derives the micro functions consistent with (3) and the PIGLOG index for income where $\ln y_h$ denotes the natural logarithm of y_h.

Result 2: Given that the measure of aggregate income is the PIGLOG index,

$$y_{oj} = \frac{\sum_{h=1}^{H} y_h \ln(y_h/c_{h_j})}{\sum_{h=1}^{H} y_h},$$

then the aggregate responses defined in (3) exist if and only if micro responses assume the form

(4) $$g_h(y_h)\, z_{hj}(y_h, p) = \alpha_j(p)\, y_h + \beta_j(p)\, \sigma_h(y_h) + \gamma_{hj}(p) \qquad h=1,\ldots,H,$$

where $\sum_{h=1}^{H} \gamma_{hj}(p) = 0$.

It is important to note that this is derived for a single-equation problem and that no cross equation restrictions such as Slutsky symmetry or Engel aggregation conditions are required as in Muellbauer. The form in (4) virtually dominates empirical analyses of demand as the AIDS system is ubiquitous (e.g., Fulponi, Green and Alston). Interpreting z_{hj} as expenditures and parameterizing g_h as y_h^{θ} yields:

(5) $$y_h^{\theta} p_j X_{hj} = \alpha_j(p)\, y_h + \beta_j(p)\, y_h \ln(y_h/c_{hj}) + \gamma_{hj}(p) \qquad h=1,\ldots,H$$

where $\sum_{h=1}^{H} \gamma_{hj}(p) = 0$. We will refer to the left side of (5) as the *augmented expenditure. Two polar cases of interest in (5) are $\theta = 0$ where (1) is obtained and $\theta = 1$ where (2) holds.*[3]

When $\theta = 0$, (5) becomes the PIGLOG expenditure equation consistent with (1). As in Muellbauer, the aggregate budget share becomes

$$\overline{w}_j = \alpha_j(p) + \beta_j(p) \left[\frac{\sum_h y_h \ln y_h / c_{h_j}}{\sum_h y_h} \right]$$

where the bracketed term may be interpreted as the log of aggregate income. When $\theta = 1$, a PIGLOG expenditure equation is defined by (2). In this case

$$\overline{E}_j = \alpha_j(p) + \beta_j(p)\left[\frac{\sum_h y_h \ln y_h/c_{h_j}}{\sum_h y_h}\right].$$

Though homogeneity will not be of concern in our empirical study (our data are cross-sectional), the ability to impose homogeneity on (5) should be demonstrated. Because demands are homogeneous of degree zero in p and y_h, $p_j X_{hj}$ is homogeneous of degree one. Hence, the left side of (5) must be homogeneous of degree $\theta + 1$ in y_h and p. A simple way to impose like homogeneity on the right side of (5) is to set $\alpha_j(p) = -\beta_j(p) \ln\eta_j(p)$, where $\eta_j(p)$ is a positive function. This yields

$$-\beta_j(p)\ y_h \ln \eta_j(p) + \beta_j(p)\ y_h \ln(y_h/c_{hj})$$
$$+ \gamma_{hj}(p) = \beta_j(p)\ y_h\ [\ln(y_h/c_{hj}) - \ln \eta_j(p)]$$
$$+ \gamma_{hj}(p) = \beta_j(p)\ y_h \ln(y_h/c_{hj}\ \eta_j(p))$$
$$+ \gamma_{hj}(p),\ h = 1,\ \ldots,\ H.$$

Thus, for arbitrary $t > 0$, homogeneity requires

$$\beta_j(tP) t y_h \ln[t y_h/(c_{hj}\ \eta_j(tP)) + \gamma_{hj}(tP) =$$
$$t^{\theta+1}\ \{\beta_j(p)\ y_h \ln[y_h/(c_{hj}\ \eta_j(p))] + \gamma_{hj}(p)\}.$$

Hence, $\beta_j(tP) = t^{\theta}\beta_j(P)$, $\eta_j(tP) = t\ \eta_j(P)$ and $\gamma_{hj}(tP) = t^{\theta+1}\ \gamma_{hj}(p)$ are sufficient to ensure that demand functions corresponding to (4) are positively homogeneous of degree zero in p and y_h.

The macro form corresponding to (5) is (summing and dividing by $\sum_{h=1}^{H} y_h$)

$$\overline{z}_j(y_{oj}, p) = \alpha_j(p) + \beta_j(p) \ln y_{oj}$$

where $\ln y_{oj}$ is the aggregate income index equal to $\sum_{h=1}^{H} y_h \ln(y_h/c_{hj}) / \sum_{h=1}^{H} y_h$. Therefore (5) not only generalizes Muellbauer but retains use of the PIGLOG income index. This index is among the easiest to compute other than merely average income.

III. Empirical Test

The data are from a random sample of 763 households from the 1985 Consumer Expenditure Survey, U.S. Bureau of Labor Statistics. A rich set of consumer characteristics are available. However, for broad expenditure categories like food, commonly used socio-economic characteristics seem to be empirically unimportant in determining expenditure patterns.[4] However, family size is a salient characteristic and it enters (5) through the individual specific terms γ_{hj} and c_{hj} (see e.g., Lewbel, 1988). Food expenditure and income are mean scaled.

Because this is a cross-sectional analysis, prices are taken as constant across individuals. A preferred approach would have considered cross-sectional price differences (Cox and Wohlgenant) but the data were unavailable. Hence α_j, β_j and γ_{hj} do not vary with respect to prices. Thus, one is left to model the effects of family size fs_h on food expenditure. The specification adopted is $\gamma_{hj}(p) = \gamma_{oj}(p) + \gamma_{1j}(p)fs_h$ and $c_{hj} = fs_h^{c_j}$, h = 1,..., H. Hence, in our cross-section, the specification employed here is:

$$(6) \quad y_h^{\theta_j} pX_{hj} = \gamma_{oj} + \gamma_{1_j}fs_h + \alpha_j y_h + \beta_j y_h \ln(y_h/fs_h^{c_j}) + \varepsilon_{hj} \quad h=1,\ldots,H,$$

where $E(\varepsilon_{jh}\ \varepsilon_{jh}') = \delta_{hh'}$, where $\delta_{hh'} = \sigma^2$ when h = h' and zero otherwise and $E(\varepsilon_{jh}\ y_h) = 0$.[5] The form in (6) allows family size to affect expenditures in two distinct ways. First, $\gamma_{oj} + \gamma_{1j}\ fs_h$ allows variations in fs_h to affect average expenditures (an intercept effect) regardless of the income level. Second $c_{hj} = fs_h^{c_j}$ allows the marginal propensity to consume to vary with family size. Because $\sum_h \gamma_{hj} = 0$, $\gamma_{0j} = -\gamma_{1j}fs$ where $\overline{fs}$ is average family size (2.6 persons). Thus (6) is estimated in the following form:

$$(6') \quad y_h^{\theta_j} p_j x_{h_j} = \gamma_{1j}(fs_h - \overline{fs}) + \alpha_j y_h + B_j y_h \ln(y_h/fs_h^{c_j}) + \varepsilon_{hj}.$$

Table 1 presents the empirical results of nonlinear least squares estimation using a gradient procedure. Log likelihood values are also reported assuming ε_{hj} is normally distributed. Estimates for the unrestricted model are plausible with food being a normal good. Moreover, in accordance with Engel's Law, the unrestricted budget share declines with income. The unrestricted income elasticity evaluated at mean values of expenditure, income, and family size is .49.

As is apparent from the asymptotic t values in the unrestricted model, the null hypotheses of $\theta = 0$ or $\theta = 1$ are both rejected at traditional confidence levels. Though there is strong evidence for rejecting both the share and expenditure aggregations in (1) and (2), the log likelihood ratio is larger for the expenditure weighted average aggregation rule than for the share weighted average representation. Though the estimated parameters for Muellbauer's model are generally reasonable, the same cannot be said of the weighted average expenditure rule in (2). Note particularly the negative sign for γ: family sizes below the mean have higher average expenditures, ceteris paribus. This and the large log-likelihood ratio (or t in the unrestricted model) casts severe doubt on the appropriateness of the weighted expenditure aggregation rule for food expenditure.

Given these results, the aggregation rule empirically supported is the "nearly" weighted average

$$\bar{z}_j(y_{oj};p) = \frac{\sum_h y_h^{.153} p X_{hj}}{\sum_h y_h}$$

where $\ln y_{oj} = \sum_h y_h \ln(y_h/fs^{.711})/\sum_h y_h$ and $\bar{z}_j(y_o,p) = .786 - .439 \ln y_{oj}$.[6]

IV. The Impact of Changing Distribution

The estimates in Table 1 parameterize a coherent aggregation scheme for augmented expenditures. It is interesting to see how changes in the income distribution affect this new definition of aggregate food expenditure. A simple and useful way to do this is to examine the impact of a mean preserving spread in the income distribution using the estimated model.

The estimated augmented expenditure represented by Table 1 is concave in income because $\hat{B}_j = -.439 < 0$. Further, within the relevant data range, the parameter estimates in Table 1 imply that each of the estimated demands is concave in income. This is required by Engel's law. Jensen's inequality thus implies that a mean preserving spread of the income distribution should reduce aggregate demand.

To examine a similar effect on aggregate income, it is useful to note the connection of the PIGLOG income index to Shannon's entropy. Entropy as defined here is:

(7) $$\text{Ent} = -\sum_h \lambda_h \log \lambda_h,$$

where $\lambda_h = y_h/\sum_h y_h$. Because $0 \le \lambda_h \le 1$ and $\sum_h \lambda_h = 1$, λ_h can be interpreted as the probability that income (normalized) will assume the value $y_h/\sum_h y_h$. Entropy is a standard measure of inequality or the variance of income (Shorrocks).

Using entropy (Theil, 1967), the PIGLOG index can be written

(8) $$\ln y_{oj} = -\text{Ent} + \ln(\sum_h y_h) - \frac{\sum_h y_h \ln fs_h^c}{\sum_h y_h}.$$

When the estimate of the weighting parameter $\hat{\theta}$ is small, the interpretation of $\bar{z}_j$ is closer to an aggregate share than to an aggregate