Nutrient Demand Elasticities with Noisy Measures of Household Resources

John Gibson[†]
University of Waikato

Abstract

Many studies suggest that changes in household economic resources (incomes and expenditure) have little effect on nutrient intakes and child malnutrition in developing countries. This paper examines the impact that errors-in-variables have on inferences about the importance of household incomes to the calorie and protein demands of households. Results are based on a new household survey from Papua New Guinea, with repeated observations on households during the year. These repeated observations allow regression estimates to be corrected for the differing reliabilities of the explanatory variables.

JEL: I32, O15
Keywords: Errors-in-variables, Income, Nutrition,

Contributed Paper:
43rd Annual Conference
Australian Agricultural and Resource Economics Society
Christchurch
January 20-22, 1999

[†] Department of Economics, University of Waikato, Private Bag 3105, Hamilton, New Zealand.
Fax: (64-7) 838-4331. E-mail: jkgibson@waikato.ac.nz.

I. Introduction

The appropriate strategy for solving the nutritional problems of poor countries depends crucially on the value of the income elasticity of nutrient demand. If the elasticity is high – as was once believed – economic growth may be all that is needed to eliminate hunger and malnutrition. But if the elasticity is low, economic growth will bring few nutritional gains and policy makers may need to heed proposals for directly targeting the nutrition of the poor. A low elasticity also sits uneasily with Sen's entitlements approach to famines; if calorie intakes of the poor do not respond to incomes, one would not expect entitlement failures to result in mass starvation (Ravallion, 1990).

Despite the importance of nutrient demand elasticities, there is little consensus on their approximate values. Most attention has been paid to the calorie demand elasticity, estimates of which range from 1.18 to 0.01 (Strauss and Thomas, 1995). The wide range is partly explained by whether the estimates come directly from calorie demand equations, or are indirect conversions from food demand equations. The estimates also are affected by whether the calories measured are those actually consumed (intake) or just those available to the household, and whether household resources are measured by income or by expenditure (Bouis and Haddad, 1992).

Much recent research suggests that the income elasticity of nutrient demand is low, even for poor people who have inadequate diets. Although the poor may increase their food expenditures almost proportionately with income, this extra spending goes on food attributes other than nutrients – taste, appearance, degree of processing, variety, and status (Behrman, Deolalikar and Wolfe, 1988). Thus, according to these researchers, "increases in income will *not* result in substantial improvements in nutrient intakes" (Behrman and Deolalikar, 1987, p.505, italics in the original). Moreover, it is claimed that other factors, such as women's schooling, are more important than incomes in determining nutrient demands (Behrman and Wolfe, 1984). This recent literature on calorie demand elasticities claims that the high elasticities that were previously believed to exist are likely to be the result of either bad data or bad estimation methods (Bouis, 1994).

This paper reports estimates of the income elasticity of nutrient demand for Papua New Guinea (PNG). One feature of the results is that they are based on a survey with repeated observations on (a subset of) households during the year, which allows regression estimates to be corrected for the effect of errors-in-variables. Another feature is that estimates are reported for both calorie and protein demands, because it is protein shortages which are the more binding constraint in PNG.

One potential objection to the results presented here, that it is nutrition constraining income rather than the reverse, can be dismissed immediately. As Subramanian and Deaton (1996) have recently pointed out with Indian data, the cost of food energy for daily work activity is much too low for the nutritional efficiency wage story to be plausible. The same result holds in Papua New Guinea: On average, the extra 600 calories needed for a day's physical work can be bought in the form of the traditional diet (banana, root crops, sago) for less than three percent of the minimum daily wage. Even in urban areas, where traditional staples are expensive because of transport costs, 600 calories from the traditional diet costs only seven percent of the minimum wage, and buying this energy in the form of imported cereals costs just four percent of the daily wage. Moreover, and in contrast to most countries, the ready access to land under customary tenure allows even the poorest in PNG to grow their own food, so there is little apparent destitution and even less reason for believing that malnutrition explains poverty, rather than results from it.

II. Income and Nutrients: Themes From the Literature

A wide ranging set of studies of income and expenditure elasticities of calorie demand have recently been summarised by Strauss and Thomas (1995). Four themes emerge from the patterns of the elasticity estimates. First, the largest estimates of nutritional responses to income changes come from indirect calculations of the calorie elasticity, based on (calorie) weighted averages of expenditure elasticities for broad food groups. One problem with this method is that it assumes a constant conversion factor between expenditure on a food group and the quantity of calories obtained. If there is "quality-shading", where richer households buy more expensive (e.g., "tastier or "more convenient") calories within the broad food groups, the elasticity of calorie *quantities* with respect to income will be overstated (Behrman and Deolalikar, 1987). This error can cause a substantial upward bias; Behrman and Deolalikar estimate indirect elasticities that are between 3.2 and 4.5 times higher than the directly estimated calorie elasticity (0.77 vs 0.17 using levels estimators and 1.18 vs 0.37 using first differenced estimators).

The second theme that emerges is that when household resources are measured by income rather than by expenditure, the calorie elasticity is lower. This is consistent with the argument that in developing countries, current expenditures are less subject to short-term fluctuations than are current incomes (Grootaert, 1995), making expenditures the better proxy for permanent income. However, it is not clear that being better than income is a sufficient condition for using current expenditure as a proxy for permanent income. The high within-year variability of expenditures reported by Scott (1992) suggests that even current expenditure is not a very reliable proxy for permanent income – we may get quite different estimates of permanent income depending on the particular period of the year that expenditures are observed.

The third theme that emerges concerns the way that calories are measured. The data commonly available to economists refer to calorie *availability*, which is derived from the same household budgets as the total expenditure data used as the regressor. This creates the possibility of common errors between nutrient consumption and expenditures. Unlike standard (uncorrelated) measurement error bias, which causes truncation of regression coefficients towards zero, correlated error bias can increase the size of measured elasticities (Bouis and Haddad, 1992). A further problem with calorie availability data is that it may not adequately control for *wastage* and *leakages*. Calorie consumption is overstated for households that give away or waste relatively more food or have relatively more visitors at meals, while it is understated for households that are absent from many meals or receive food gifts. Because the first group are likely to be rich and the second group poor, uncorrected leakages will cause the calorie elasticity to be overstated.

The fourth theme that emerges is that elasticities measured at a single evaluation point, such as the mean per capita income or expenditure level, may understate the elasticity that applies to much poorer households. This would occur if the calorie elasticity is subject to non-linearities, which is a plausible claim (Ravallion, 1990). Because it is the poorly nourished who are the subject of public concern, elasticities reported at mean expenditures may be inappropriate summary statistics of the opportunities and constraints facing the group who are the main concern of public policy.

III. Nutrient Demand Estimates and Measurement Error

To see the effect that measurement error has on the estimated income elasticity of demand for nutrients, it is helpful to start with the case of a bivariate regression of nutrients, y^* on permanent income, x^* (suppressing the constant for ease of notation):

$$y^* = \beta x^* + \varepsilon. \tag{1}$$

The observed data, y and x only imperfectly measure their theoretical counterparts, y^* and x^*:

$$y = y^* + v \qquad \text{with } v \sim N(0, \sigma_v^2)$$

$$x = x^* + u \qquad \text{with } u \sim N(0, \sigma_u^2)$$

and the errors v and u are assumed to be uncorrelated with each other and with y^* and x^*. It is well-known that under these assumptions, the errors in measuring y^* cause no problems of bias:

$$y = \beta x^* + \varepsilon + v \;= \beta x^* + \varepsilon^* \tag{2}$$

while the errors in measuring x^* cause the OLS estimate of β to be biased towards zero (Greene, 1997, p.437). The degree of this 'attenuation bias' depends on the proportion of the variation in x that is due to variation in the true value, x^*:

$$\text{plim } \hat{\beta} = \frac{\text{cov}(x^* + u,\ y^* + v)}{\text{var}(x^* + u)} = \frac{\text{cov}(x^*, y^*)}{\text{var}(x^*) + \text{var}(u)} = \beta \cdot \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_u^2} \tag{3}$$

where this proportion, $\sigma_{x^*}^2 \big/ (\sigma_{x^*}^2 + \sigma_u^2)$ is known as the 'reliability ratio' of x and will henceforth be denoted λ. It is clear that if an estimate of λ is available, the true value of β can be recovered from the attenuated value estimated by OLS, and this point is returned to below.

A more realistic nutrient demand relationship uses the multiple regression model:

$$\mathbf{y}^* = \mathbf{X}^* \beta + \varepsilon \tag{4}$$

where, potentially, all of the variables in the matrix of explanatory variables, $\mathbf{X}^*$ are measured with error:

$$\mathbf{X} = \mathbf{X}^* + \mathbf{U}\,.$$

If each column of the measurement error matrix $\mathbf{U}$ is independent of every other column, and is independent of any measurement errors in $\mathbf{y}$, and of the true values $\mathbf{y}^*$ and $\mathbf{X}^*$, then the degree of bias for the vector of OLS coefficients is given by:

$$\text{plim}\,\hat{\boldsymbol{\beta}} = \left[\mathbf{Q}^* + \boldsymbol{\Sigma}_{uu}\right]^{-1}\mathbf{Q}^*\boldsymbol{\beta} \qquad = \qquad \boldsymbol{\beta} - \left[\mathbf{Q}^* + \boldsymbol{\Sigma}_{uu}\right]^{-1}\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta} \tag{5}$$

where $\mathbf{Q}^*$ is the moment matrix of the true $\mathbf{X}^*$ matrix, and $\boldsymbol{\Sigma}_{uu}$ is the covariance matrix of the measurement errors in the x_i variables. It can be shown from equation (5) that even if only a single explanatory variable is measured with error, all of the coefficients will be biased (Greene, 1997, p.440). If estimates of the reliability ratios, λ_i are available for each of the i explanatory variables, then the errors-in-variables estimator can be used to recover the true coefficient values:

$$\hat{\boldsymbol{\beta}}_{EIV} = \left[\mathbf{X}'\mathbf{X} - \mathbf{S}\right]^{-1}\mathbf{X}'\mathbf{y} \tag{6}$$

where **S** is a diagonal matrix with elements $N(1-\lambda_i)s_i^2$, where N is the number of observations, λ_i is the reliability ratio, and s_i^2 is the variance of the variable (StataCorp, 1997). However, this estimator will not recover the true coefficient values if the errors in **U** are correlated with any errors in measuring **y***, because in this case of correlated measurement error the degree of bias for the vector of OLS coefficients is:

$$\text{plim}\,\hat{\boldsymbol{\beta}} = \left[\mathbf{Q}^* + \Sigma_{uu}\right]^{-1}\mathbf{Q}^*\boldsymbol{\beta} + \left[\mathbf{Q}^* + \Sigma_{uu}\right]^{-1}\boldsymbol{\Gamma} \tag{7}$$

which depends on the vector of covariances, **Γ** between the measurement error in **X** and **y**.

IV. Data and Model Specification

The data used for this study come from a random sample of 1144 Papua New Guinea households, residing in 120 rural and urban communities ("clusters"), who were interviewed between January and December 1996 (Gibson and Rozelle, 1998). The clusters were selected from the enumeration areas of the 1990 Census, using a stratified sample (15 strata). Household weights were derived from (i) the unequal sampling rates between strata, (ii) the variation between the 1990 Census estimates of the size of each cluster and the actual size found during the survey, and (iii) the deviation of the actual number of households surveyed in each cluster from the target number. Except where noted, all results presented below take account of the clustered, stratified and weighted nature of the sample.

The survey used a closed interval recall method, with households interviewed twice so that the start of the recall period was signaled by the first interview. These two interviews were usually two weeks apart, which is the length of the pay period in Papua New Guinea. Expenditure data were collected on all food (36 categories) and other frequent expenses (20 categories) during the recall period. The expenditure estimates include the imputed value of own-production, net gifts received, and stock changes, so they should be a comprehensive monetary measure of consumption during the recall period. An annual recall covered 31 categories of infrequent expenses. An inventory of durable assets was used to estimate the value of the flow of services from these assets, including rental services from owner-occupied dwellings.

The survey also collected data on the quantities of purchased and self-produced foods, the quantities of food gifts given and received, and the quantities of food stock changes. These food quantity data were collected for the same 36 categories of food that expenditures were collected for. The Pacific Islands Food Composition Database was used to compute the calorie quantities from the food quantity data. One item where food quantities were not available was cooked meals eaten out of the home; calories from this source were derived as the average "price" each household paid for all other calories plus a 50 percent premium to reflect processing margins (Subramanian and Deaton, 1996).

Twenty clusters were chosen, randomly, as a "longitudinal sub-sample". Expenditures by households in these clusters were observed for two periods of the year, roughly seven months apart. All other variables collected by the survey were also gathered again during these revisits. The correlation between these two sets of observations on the same household allow the reliability ratio to be calculated for each variable collected by the survey and subject to a caveat noted below, this may allow the errors-in-variables estimator (equation 6) to be implemented. These reliability ratios range from 0.9, for variables like schooling, to 0.6 for variables like expenditure and income sources (Table 1). This pattern is in accord with the result that response error tends to increase with the complexity of the concept being measured (Fuller, 1987, p. 8).

If the households in the longitudinal sub-sample are representative of the whole sample, the correlations in Table 1 can be used as estimated reliability ratios to correct the regression estimates for the effects of errors-in-variables. There is no statistically significant difference between the average values of household characteristics for the longitudinal sub-sample and the average for the households in the other 100 clusters (the "cross-sectional sub-sample"), thus the longitudinal sub-sample appears to be representative of the whole sample (Table 2). In particular, households in the two sub-samples have the same average level of expenditure per adult equivalent and the same pattern of economic activity. Households in the two sub-samples have similar access to transport infrastructure, which affects price (and hence, expenditure) fluctuations during the year by allowing movement of goods from surplus to shortage areas. Finally, a similar proportion of clusters in the two sub-samples are in dry areas, which tend to have more variable food production over the course of the year.

Model Specification

To allow comparison with earlier results in the literature, two of the models estimated are based on specifications that have been used elsewhere. The first of these specifications uses a quadratic in (log) per capita expenditures, (log) household size, years of schooling for adult women in the household, and several household demographic ratios. This specification follows one used by Behrman and Wolfe (1984), in a much-quoted study which showed women's schooling having a stronger effect than income on nutrient demand. The second specification is based on Subramanian and Deaton (1996) and includes various economic activity variables in addition to household demographic terms, and in one variant allows separate intercepts for each cluster to proxy for omitted price and other locational effects.

IV. Results

Figure 1 shows the scatterplot of (log) per capita calorie and protein availability versus the (log) of total household expenditure per capita expenditure (PCE), which is used as the proxy for permanent income. The fitted values from the bivariate regression of calories (or protein) on expenditure is also reported, along with the coefficient estimates. It is evident that protein availability responds to changes in expenditure more strongly,(with an elasticity of 0.60 compared with 0.38) and with less variation about the mean response, than does calorie availability. However, there is little in Figure 1 to support the notion that changes in permanent income have no effect on the availability of calories.

Figure 1 also shows the fitted values of a nonparametric regression which can be compared with the linear regression estimates to see if non-linearities are an important feature of the PNG data. For each point (x_i, y_i) on the scatterplot, the smoothed point (x_i, $\hat{y}_i$) was formed from a locally weighted regression of a first order polynomial. The weights were formed for the 300 nearest neighbouring points of (x_i, y_i), using the "tricube" function:

$$w_k(x_i) = \left[1 - \left(\frac{|x_k - x_i|}{\max \aleph(x_i)\ |x_j - x_i|}\right)^3\right]^3$$

A new set of weights, δ_i were then defined for each (x_i, y_i) based on the size of the residual y_i- $\hat{y}_i$. Larger residuals have smaller weights, to guard against outliers distorting the smoothed plots. New fitted values were computed as before, but using the new weights – full details of this LOWESS method are found in Cleveland (1979).

Although some non-linearity is apparent in the calorie-expenditure relationship, it proved to be statistically insignificant when modelled as a spline function. Specifically, when dummies for each population quartile were interacted with (log) PCE, none of these splines were significantly different from the linear component (p<0.61). It was also usually the case that quadratic terms in (log) PCE were statistically insignificant. Hence, the rest of the paper concentrates on parametric estimates of the relationship between nutrients and household resources.

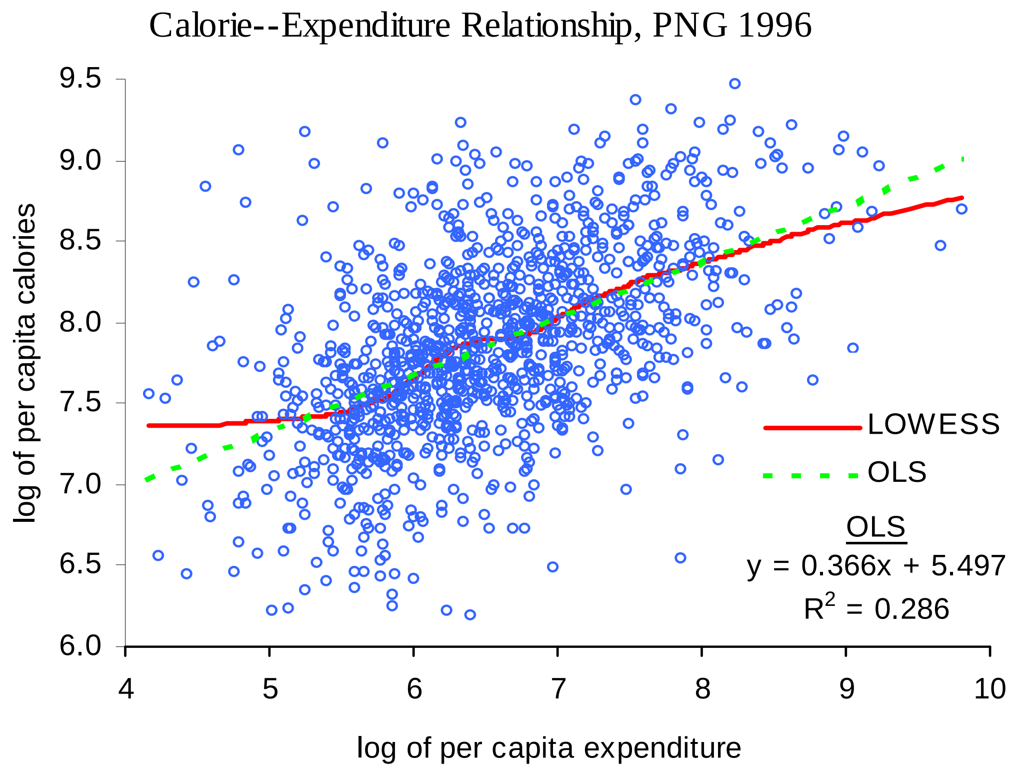
Table 3 reports the OLS results, which show similar patterns for both calories and protein across the various specifications. The expenditure elasticities fall once household size is added to the model, because size is negatively correlated with both PCE and with nutrient demand, and larger households also have more children and so have lower nutrient needs per capita. In the column (2) estimates, the quadratic on PCE is not statistically significant in the calorie regression although it is in the protein regression. The implied elasticities at the mean of (log) PCE are 0.35 for calories and 0.60 for protein, while the protein elasticity is 0.69 at one standard deviation below the mean of (log) PCE. In contrast to Behrman and Wolfe's results for Nicaragua, women's schooling exerts a negative influence on nutrient demands in PNG (especially calories) once household expenditures are controlled for. The most likely explanation is that households with a high level of women's schooling also have a high overall schooling level, so economically active adults are more likely to be engaged in sedentary occupations, lowering the demand for calories. The results in columns (3) and (4) show further increases in the income elasticities of calorie and protein availability as more variables are added to the model. Overall, the results in Table 3 show that the bivarate elasticities reported in Figure 1 are quite conservative estimates, and therefore entirely defensible measures of the nutrient demand elasticities.

The results in Table 4 can be considered as upper bound estimates of the nutrient elasticities. These results have had attentuation bias due to errors-in-variables removed, based on the estimator described in equation (6). The resulting calorie demand elasticity is approximately 0.5, while the protein demand elasticity is approximately 0.9. These differences from the OLS values indicate the potential biases caused by using proxies for permanent income that have large measurement errors, caused by within year variability.

The results in Table 5 can be considered as lower bound estimates of the nutrient elasticities, and are applicable if measurement errors in the explanatory variables are correlated with errors in the dependent variable. This is likely in a nutrient availability study because estimates of food expenditure are common to the nutrient estimates and to the total expenditure estimates. Although estimators that correct for correlated measurement error can not be implemented here, because of lack of information on the covariance between the errors u and v, if nonfood expenditure is used as an instrument for household total expenditure, the resulting elasticities will be biased downwards whether or not correlated measurement error is present. These downwardly biased estimates of the calorie elasticity are approximately 0.20, and for the protein elasticity approximately 0.50.

V. **Conclusions**

Elasticities of nutrient availability with respect to household economic resources have been estimated for Papua New Guinea. Under a variety of different assumptions about the error in measuring permanent income with current expenditures, the income elasticity of calorie availability varied from approximately 0.2 to 0.5, while the elasticity of protein availability varied from 0.5 to 0.9. None of these results are consistent with the view that changes in permanent income have little or no effect on nutrition.


Calorie--Expenditure Relationship, PNG 1996
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
log of per capita calories
4
5
6
7
8
9
10
log of per capita expenditure
LOWESS
OLS
OLS
y = 0.366x + 5.497
R2 = 0.286

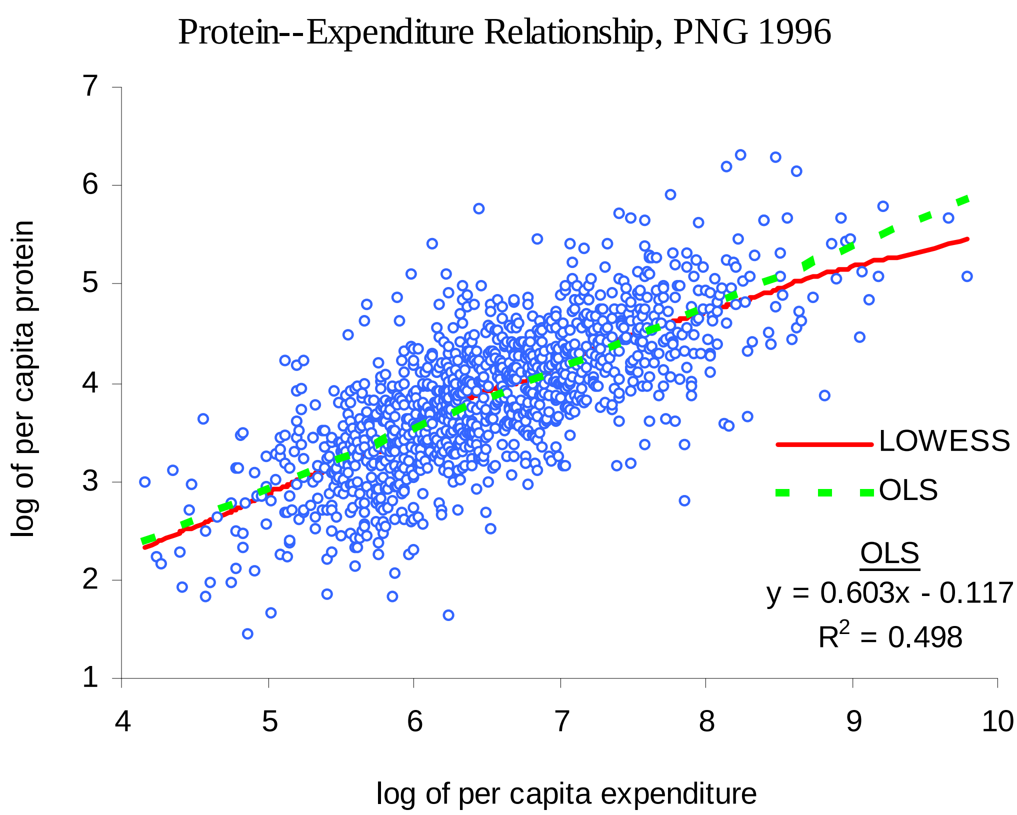

Protein--Expenditure Relationship, PNG 1996
7
6
5
4
3
2
1
log of per capita protein
4
5
6
7
8
9
10
log of per capita expenditure
LOWESS
OLS
OLS
y = 0.603x - 0.117
R2 = 0.498

Table 1: Correlations Between Two Estimates of Each Explanatory Variable[a]

Explanatory variable	Weighted correlation[b]	Unweighted correlation
ln per capita expenditure (PCE)	0.6586	0.6243
[ln PCE]^2	0.6491	0.6155
ln household size	0.7816	0.7776
rf15+	0.5957	0.6161
rf714	0.7391	0.7200
rf06	0.6544	0.6846
rm714	0.7500	0.7480
rm06	0.8212	0.8356
Women's school years	0.8758	0.8950
Head's school years	0.9186	0.9249
Age of head	0.8996	0.8700
Food crop income	0.5956	0.5766
Wage & business income	0.8278	0.8221

Note: Variables beginning with *r* are demographic ratios, so that e.g., rf714 is the ratio of females aged 7-14 to total household members. The income variables classify households according to the main income source of the household head.

[a] Estimates made approximately seven months apart, for a random sub-sample of 162 households in 20 clusters.

[b] Weighted by the inverse of the probability that each observation is included in the sample of households.

Table 2: Comparison of the Cross-sectional and Longitudinal Sub-samples

	Cross-sectional sub-sample[a] N=982	Longitudinal sub-sample[a] N=162	t-test for equal means[b]
Total expenditure per adult equivalent[c]	901.3 (83.2)	924.9 (130.7)	0.15 [0.89]
Household size	5.8 (0.2)	6.1 (0.3)	0.80 [0.43]
Years of school of household head	4.0 (0.3)	4.8 (0.8)	0.94 [0.36]
Age of household head	40.1 (0.6)	40.8 (1.7)	0.41 [0.69]
Female head (%)	7.9 (1.3)	8.4 (2.5)	0.16 [0.88]
Head's main income is wage job (%)	20.1 (2.4)	26.2 (8.9)	0.66 [0.52]
Head's main income is tree crop agriculture (%)	38.0 (4.4)	42.7 (9.1)	0.46 [0.65]
Minutes walk to nearest road, airstrip or port[d]	60.8 (15.4)	43.5 (21.9)	0.65 [0.52]
Dry climate (<2500mm rainfall/year) (%)[d]	39.5 (6.9)	42.6 (14.0)	0.20 [0.85]

Notes:
[a] Standard error of the mean in ( ) corrected for clustering, sampling weights and stratification.
[b] p-level for two-sided hypothesis test in [ ].
[c] Kina per year, in 1996 national average prices, where the value of the poverty line is used as the spatial price deflator and K1.3=US$1 in 1996.
[d] Data collected at cluster level, and weighted by the population in each cluster.

Table 3: OLS Estimates of Nutrient Availability Regressions, Papua New Guinea, 1996

	(1)		(2)		(3)		Within cluster (4)	
	β	|*t*|	β	|*t*|	β	|*t*|	β	|*t*|
				Calories				
ln PCE	.312	(6.8)	.763	(2.1)	.397	(10.4)	.533	(16.0)
[ln PCE]^2	…	…	-.031	(1.1)	…	…	…	…
ln household size	-.320	(8.2)	-.271	(7.1)	-.245	(6.2)	-.170	(4.2)
rf15+	…	…	.194	(1.6)	.090	(.8)	.123	(1.1)
rf714	…	…	-.151	(1.1)	-.108	(.8)	-.037	(.3)
rf06	…	…	.078	(.5)	.029	(.2)	-.069	(.5)
rm714	…	…	-.029	(.2)	-.032	(.2)	.089	(.7)
rm06	…	…	-.025	(.2)	-.103	(.8)	-.079	(.7)
Women's school years		…	-.026	(4.3)	…	…	…	…
Head's school years	…	…	…	…	-.015	(3.6)	-.011	(3.0)
Age of head	…	…	…	…	-.003	(2.0)	-.003	(1.8)
Food crop income	…	…	…	…	.081	(1.5)	.032	(.5)
Wage & business income		…	…	…	-.083	(1.7)	-.044	(.9)
Urban location	…	…	…	…	-.175	(2.1)	-1.541	(24.8)
Constant	6.362	(18.3)	4.732	(3.9)	5.896	(19.1)	6.027	(24.1)
Zero slopes *F*-test	$F_{(2,104)}$=198.2		$F_{(9,97)}$=72.1		$F_{(9,97)}$=68.0		$F_{(10,96)}$=81.3	
R^2	.373		.399		.423		.625	
				Protein				
ln PCE	.571	(12.5)	1.325	(4.1)	.604	(15.9)	.736	(22.2)
[ln PCE]^2	…	…	-.056	(2.2)	…	…	…	…
ln household size	-.189	(4.9)	-.199	(4.9)	-.173	(4.6)	-.100	(2.5)
rf15+	…	…	-.004	(.0)	-.037	(.3)	-.021	(.2)
rf714	…	…	-.020	(.2)	-.004	(.0)	-.002	(.0)
rf06	…	…	.394	(1.9)	.324	(1.8)	.212	(1.4)
rm714	…	…	.043	(.3)	.060	(.5)	.152	(1.5)
rm06	…	…	-.066	(.4)	-.159	(1.2)	-.075	(.6)
Women's school years		…	-.006	(1.0)	…	…	…	…
Head's school years	…	…	…	…	-.003	(.6)	-.007	(1.8)
Age of head	…	…	…	…	-.003	(1.7)	-.003	(2.2)
Food crop income	…	…	…	…	.180	(3.2)	.030	(.4)
Wage & business income		…	…	…	-.022	(.4)	-.139	(2.0)
Urban location	…	…	…	…	-.011	(.1)	-.276	(3.4)
Constant	.395	(1.2)	-2.095	(2.0)	.214	(.8)	-.127	(.6)
Zero slopes *F*-test	$F_{(2,104)}$=143.5		$F_{(9,97)}$=58.4		$F_{(9,97)}$=43.7		$F_{(10,96)}$=33.6	
R^2	.518		.530		.539		.701	

Note:
The reported absolute *t*-values are corrected for the clustered, stratified, and weighted nature of the sample.

Variables beginning with *r* are demographic ratios, so that e.g., rf714 is the ratio of females aged 7-14 to total household members. The omitted group is male adults. There are three economic activity groups, with households whose head's main income is from tree crops omitted. The within cluster regression contains 119 dummy variables for clusters.

Table 4: Upper Bound (Correcting for Errors-in-variables) Estimates of Nutrient Availability Regressions, Papua New Guinea, 1996

	(1)		(2)		(3)		(4)	
	β	|*t*|	β	|*t*|	β	|*t*|	β	|*t*|
Calories								
ln PCE	.552	(8.6)	.469	(6.3)	.585	(7.1)	.630	(7.9)
ln household size	…	…	-.334	(5.4)	-.209	(3.0)	-.240	(3.5)
Women's school years		…	…	…	-.051	(6.2)	…	…
Head's school years	…	…	…	…	…	…	-.043	(8.2)
Zero slopes *F*-test	$F_{(1,105)}$=74.7		$F_{(2,104)}$=194.3		$F_{(3,103)}$=141.3		$F_{(3,103)}$=153.1	
R^2	.284		.373		.391		.397	
Protein								
ln PCE	.923	(13.5)	.905	(11.8)	1.007	(12.0)	1.049	(12.3)
ln household size	…	…	-.072	(1.2)	.037	(0.5)	.012	(0.2)
Women's school years		…	…	…	-.045	(5.9)	…	…
Head's school years	…	…	…	…	…	…	-.039	(7.7)
Zero slopes *F*-test	$F_{(1,105)}$=183.3		$F_{(2,104)}$=136.1		$F_{(3,103)}$=94.0		$F_{(3,103)}$=106.1	
R^2	.503		.512		.521		.521	

Note: The reported absolute *t*-values are corrected for the clustered, stratified and weighted nature of the data. Models are estimated in deviation-from-the-mean form, so no constant terms are included. Some of the more fully parameterised models could not be estimated because the low reliability ratios and highly collinear variables (especially ln PCE and [ln PCE]^2) prevented the inversion of the matrix **A** for the estimator: $\mathbf{b}=\mathbf{A}^{-1}\mathbf{X}'\mathbf{W}y$.

Table 5: Lower Bound (IV) Estimates of Nutrient Availability Regressions, Papua New Guinea, 1996

	(1)		(2)		(3)		Within cluster (4)	
	β	|*t*|	β	|*t*|	β	|*t*|	β	|*t*|
				Calories				
ln PCE[a]	.167	(3.9)	.028	(.1)	.209	(5.3)	.328	(7.4)
[ln PCE]^2	…	…	.012	(.4)	…	…	…	…
ln household size	-.377	(8.9)	-.325	(7.6)	-.308	(7.4)	-.265	(5.2)
rf15+	…	…	.130	(1.0)	.060	(.5)	.113	(1.0)
rf714	…	…	-.258	(1.8)	-.234	(1.7)	-.101	(.8)
rf06	…	…	-.090	(.6)	-.155	(1.0)	-.189	(1.3)
rm714	…	…	-.098	(.6)	-.119	(.8)	-.009	(.1)
rm06	…	…	-.145	(1.0)	-.239	(1.8)	-.186	(1.5)
Women's school years		…	-.012	(1.9)	…	…	…	…
Head's school years	…	…	…	…	-.007	(1.3)	-.005	(1.3)
Age of head	…	…	…	…	-.003	(1.8)	-.003	(2.0)
Food crop income	…	…	…	…	.076	(1.5)	.025	(.4)
Wage & business income		…	…	…	-.017	(.3)	.011	(.2)
Urban location	…	…	…	…	-.077	(.9)	-1.213	(15.1)
Constant	7.393	(22.3)	7.763	(5.8)	7.199	(22.6)	7.252	(24.0)
Zero slopes *F*-test	$F_{(2,104)}$=117.1		$F_{(9,97)}$=29.4		$F_{(12,94)}$=36.6		$F_{(10,96)}$=34.9	
R^2	.331		.346		.372		.591	
				Protein				
ln PCE[a]	.441	(10.2)	.874	(2.1)	.426	(9.6)	.497	(11.1)
[ln PCE]^2	…	…	-.033	(1.1)	…	…	…	…
ln household size	-.240	(5.8)	-.251	(5.9)	-.233	(5.9)	-.210	(4.3)
rf15+	…	…	-.072	(.5)	-.065	(.5)	-.033	(.3)
rf714	…	…	-.112	(.8)	-.123	(.9)	-.076	(.6)
rf06	…	…	.245	(1.2)	.150	(.8)	.081	(.5)
rm714	…	…	-.025	(.2)	-.022	(.2)	.038	(.3)
rm06	…	…	-.175	(1.1)	-.288	(1.9)	-.199	(1.4)
Women's school years		…	.007	(1.1)	…	…	…	…
Head's school years	…	…	…	…	.005	(1.0)	-.000	(.0)
Age of head	…	…	…	…	-.003	(1.5)	-.004	(2.4)
Food crop income	…	…	…	…	.175	(3.3)	.021	(.3)
Wage & business income		…	…	…	.040	(0.7)	-.075	(1.1)
Urban location	…	…	…	…	.082	(1.1)	.104	(1.0)
Constant	1.313	(4.2)	-.044	(.0)	1.452	(4.7)	1.292	(4.8)
Zero slopes *F*-test	$F_{(2,104)}$=111.3		$F_{(9,97)}$=28.2		$F_{(12,94)}$=27.1		$F_{(10,96)}$=17.4	
R^2	.496		.504		.509		.671	

Note:
The reported absolute *t*-values are corrected for the clustered, stratified, and weighted nature of the sample. Variables beginning with *r* are demographic ratios, so that e.g., rf714 is the ratio of females aged 7-14 to total household members. The omitted group is male adults. There are three economic activity groups, with households whose head's main income is from tree crops omitted. The within cluster regression contains 119 dummy variables for clusters.

[a] Treated as endogenous and instrumented for by the (log) of non-food expenditure per capita.

References

Anand, S. and Harris, C. (1994), "Choosing a welfare indicator", *American Economic Review,* 84, 226-231.

Behrman, Jere and Barbara Wolfe, 1984, More evidence on nutrition demand: income seems overrated and women's schooling underemphasized, *Journal of Development Economics* 14(1): 105-128.

Behrman, Jere and Anil Deolalikar, 1987, Will developing country nutrition improve with income? A case study for rural South India, *Journal of Political Economy* 95(3): 492-507.

Bouis, Howarth, and Lawrence Haddad, 1992, Are estimates of calorie income elasticities too high?: A recalibration of the plausible range, *Journal of Development Economics* 39(4): 333-364.

Deaton, Angus. (1997) *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy* Johns Hopkins, Baltimore.

Grootaert, C. (1983), "The conceptual basis of measures of household welfare and their implied survey data requirements", *The Review of Income and Wealth*, 29, 1-21.

Ravallion, Martin, 1990, *Economic Development and Cultural Change*

StataCorp (1997), *Stata Statistical Software: Release 5.0*, College Station: Stata Corporation.

Strauss, John and Duncan Thomas, 1995, Human resources: empirical modelling of household and family decisions, in: J. Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics Vol. III*, (North-Holland, Amsterdam) 1883-2023.

Subramanian, Shankar and Angus Deaton, 1996, The demand for food and calories, *Journal of Political Economy* 104(1): 133-162.

Wolfe, Barbara and Jere Behrman, 1983, Is income overrated in determining adequate nutrition? *Economic Development and Cultural Change* 31(3): 525-549.