

Strauss, John; Thomas, Duncan

**Working Paper**

## The Shape of the Calorie-Expenditure Curve

Center Discussion Paper, No. 595

**Provided in Cooperation with:**

Yale University, Economic Growth Center (EGC)

*Suggested Citation:* Strauss, John; Thomas, Duncan (1990) : The Shape of the Calorie-Expenditure Curve, Center Discussion Paper, No. 595, Yale University, Economic Growth Center, New Haven, CT

This Version is available at:

<https://hdl.handle.net/10419/160517>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station  
New Haven, Connecticut 06520

CENTER DISCUSSION PAPER NO. 595

THE SHAPE OF THE CALORIE-EXPENDITURE CURVE

John Strauss  
Rand Corporation

Duncan Thomas  
Yale University

March 1990

Notes: Support to both authors by Empresa Brasileira de Pesquisa Agropecuaria is gratefully acknowledged.

Partial support to Professor Strauss was also provided by the National Institutes of Health grant number R01-H021098-03 and 5S07-RR0-5710.

Center Discussion Papers are preliminary materials circulated to stimulate discussion and critical comments.

This discussion paper is a revision of an earlier draft written in December 1989.

We have benefitted from the comments of Harold Alderman, Janet Currie, Angus Deaton, Anil Deolalikar, James Heckman, Whitney Newey and T.N. Srinivasan. Maria-Helena Henriques and Mauricio de Vasconcellos have provided valuable help in understanding the data.

## Abstract

It is generally agreed that as income (or expenditure) rises, households switch to higher valued foods not necessarily with higher nutrient content. Some recent work has suggested that, even among the very poor, as income rises households purchase only additional taste and non-calorie nutrients: the calorie-income curve is very flat.

In this paper the effect of expenditure on calorie and protein demand is estimated with data from a large scale Brazilian expenditure survey. Non-parametric estimates, which impose very little structure on the relationship between nutrient intake and expenditure, demonstrate that the calorie-expenditure curve is positively sloped for households in the bottom three quartiles of the *per capita* expenditure (PCE) distribution; it is kinked when *per capita* calorie intake is between 2,500 and 3,000 calories per day and is flat at higher expenditure levels. The protein-expenditure curve is also positively sloped although it does not flatten out as much as the calorie curve, reflecting a switch into higher valued foods, richer in protein.

Using the non-parametric estimates as a guide, the nutrient expenditure functions are parameterized. The expenditure elasticities are estimated to be between 0.25 to 0.30 for calories and 0.40 for protein at the bottom decile of PCE (with standard errors less than 0.05). The hypothesis that nutrient elasticities are zero is unambiguously rejected.

## 1. Introduction

Studies in the early and mid-1970s showed that, for the malnourished, the binding constraint on nutrient intake is in calories and that deficiencies in other nutrients, especially proteins, were accompanied by deficiencies in calories (Sukhatme, 1970). Berg (1973) argued that while nutrient intake in a population will depend on many factors, such as the existence of nutrition programs, income is likely to be an important determining factor.

Many studies have attempted to measure the responsiveness of household calorie intake (or availability) to income since the seminal work of Reutlinger and Selowsky, (1976), which assumed an elasticity between 0.15 and 0.30, based on little information. There is, today, considerable disagreement in the literature over the size of this elasticity with estimates ranging between zero and one (see Table 1 and surveys by Alderman, 1989, and Behrman and Deolalikar, 1988).

It is generally agreed that as income (or expenditure) rises, households switch to higher valued foods not necessarily with higher nutrient content (Poleman, 1981; Behrman and Deolalikar, 1987). It had been thought, however, that calories respond positively to income at least at low income levels. It is likely the response declines with income and it has been conjectured that there may be a kink in the calorie-income curve. Some recent work, however, has questioned this conventional wisdom and suggests that even among the very poor, as income rises households purchase only additional taste and non-calorie nutrients: the calorie-income curve is very flat (Wolfe and Behrman, 1983; Behrman and Deolalikar, 1987; Behrman, Deolalikar and Wolfe, 1988). This is consistent with the controversial hypothesis that it is rare for individuals to be truly malnourished because of short and long run body adaptation to food shortages.<sup>1</sup> If the calorie-income curve is flat, then there are powerful

implications for policy. In particular, concern over low calorie intakes of the poor has been misdirected and the substantial resources allocated by governments and international agencies to alleviate this "problem" have been misspent.

In this paper, the effect of expenditure (rather than income) on nutrient demand (both calories and protein) is estimated with data from a Brazilian expenditure survey. We take advantage of the large sample size by doing exploratory data analysis and pre-testing on a random sub-sample of households; we then present estimates of expenditure elasticities using a different random sub-sample. Using non-parametric methods, which impose very little structure on the relationship between nutrient intake and expenditure, we find the calorie-expenditure curve is positively sloped for households in the bottom three quartiles of the *per capita* expenditure (PCE) distribution; it is kinked when *per capita* calorie intake is between 2,500 and 3,000 calories per day and is flat at higher expenditure levels. The protein-expenditure curve is also positively sloped although it does not flatten out as much as the calorie curve reflecting a switch into higher valued foods, such as meat and milk, which are rich in protein. Using parametric functional forms suggested by non-parametric estimates, we find that for urban households, the expenditure elasticity of the demand for calories is about 0.26 at the bottom decile of PCE and about 0.03 at the top decile (with standard errors of 0.02 and 0.01 respectively). The protein elasticities are higher: 0.47 at the bottom and 0.18 at the top decile of PCE (also estimated very precisely).

The next section describes the data paying particular attention to problems associated with estimating the relation between expenditures and nutrients. This is followed by non-parametric and parametric estimates of the relationship.

## 2. Data, measurement and estimation

The calorie-expenditure curve will be estimated with a very large household budget survey from Brazil. *Estudo Nacional da Despesa Familiar* (ENDEF) is a broad purpose survey conducted from August 1974 to August 1975 which covered over 53,000 households in a nationwide random probability sample.<sup>2</sup> Detailed information were collected on, *inter alia*, household consumption and expenditures together with demographic characteristics, labor supply, earnings and non-labor income of individuals.<sup>3</sup>

Each household in the ENDEF survey was visited on seven consecutive days and at each visit, food consumed and wasted<sup>4</sup> was weighed, distinguishing over three hundred food categories. Food weights were converted into nutrients using standard food composition tables and these weights were converted into a *per capita* nutrient intake measure based on a listing of all people present at each meal — including visitors, lodgers and employees.<sup>5</sup> Since these measurements cannot be made for meals taken out of the home, it will be assumed that the nutrient content of those meals is the same as meals taken at home.<sup>6</sup> Food expenditure is based on the value of food consumed (and wasted) including meals away from home; non-food expenditures are based on recall with the recall period varying from a fortnight to a year. In all cases the value of consumption from own production, gifts and transfers are included.

Table 1 presents estimates of the calorie-income (expenditure) elasticity from 20 recent studies. Most remarkable is the range of these estimates; some of this range can, however, be explained by methodological differences. (See Bouis and Haddad, 1989, for a discussion of some issues). All the estimated elasticities of calorie demand in the first panel are calculated indirectly from elasticities estimated in a food demand system. This is a convenient empirical strategy as it can be applied to *any* expenditure survey and does not require

knowledge of both quantities and expenditures. In all cases, however, indirect estimates tend to be high. Behrman and Deolalikar (1987) point out that this may be due to aggregation biases if there is large within group substitution from lower valued to higher valued foods, with similar calorie content. Clearly the extent of aggregation bias will depend both on the level of aggregation in the food demand system and on the scope for substitution. While this may explain the very high elasticities for India (Behrman and Deolalikar, 1987), it does not explain the high elasticity in Sierra Leone (Strauss, 1984).<sup>7</sup> In this paper, we exploit the fact that data are available for over 300 foods and convert those to nutrients in order to estimate the nutrient-expenditure function directly.

Even among the direct estimates, there is considerable heterogeneity. Elasticities based on income seem to be slightly lower than those based on expenditure. This is also true in the ENDEF data. Based on a regression of (log) *per capita* calorie intake on (log) *per capita* expenditure, the elasticity is 0.21; using, instead, (log) *per capita* income as the regressor, the elasticity is 0.11; (standard errors are less than 0.01 in both cases).

If current income measures long run income with random error, then a standard errors-in-variables argument predicts a downward bias in the estimated elasticity. It is, furthermore, reasonable to suppose that current nutrient intakes depend on labor supply in which case income might be endogenous. For this case, however, it is hard to sign the bias *a priori*. If unobserved taste for work is negatively correlated with taste for calories, the bias due to simultaneity may be downward. On the other hand, for low income households, it may be that increased calorie intake raises productivity and/or labor supply<sup>8</sup> in which case the estimated elasticity will be biased upwards and the bias would be greater for lower income households.

In the modern demand literature, it is often argued that current expenditure is a better measure of long-run well-being than current income because households are able to smooth consumption more than income. Nutrient intakes and total expenditure are, however, likely to be jointly determined. In addition, when the data sources for food expenditure and nutrient intake are the same — as in ENDEF — there are likely to be common measurement errors which will result in biased estimates of elasticities.<sup>9</sup> Pure measurement error in both nutrient intake and total expenditure should be uncorrelated with asset or wealth measures and, possibly, income in which case they may be used as instruments to purge estimates of simultaneity bias and bias due to measurement error. In the ENDEF data, when *per capita* expenditure is instrumented with household income, the calorie elasticity is 0.12 (standard error 0.01). If income is endogenous, then it is not an appropriate instrument; the estimated elasticity falls to 0.05 (standard error 0.02) when unearned income, a measure of wealth, is used as the instrument. These estimates are quite similar to those reported by Behrman, Deolalikar and Wolfe.

There may also be components of measurement error which are systematically correlated with income or wealth measures (Bouis and Haddad, 1988) — such as gifts of food into or out of the household and food eaten in the household by visitors or servants. If these are not accounted for, and they are hard to measure so are not always removed, they will lead to systematic overstatement of both nutrient intake and total expenditure for households who give away relatively more food or have relatively more visitors at meals (typically higher income households) and understatement for households who receive food or are absent from many meals (typically lower income households).<sup>10</sup>

Fortunately, considerable care was taken in the collection of the ENDEF data. Since nutrient intakes are based on actual consumption, gifts, wastage and



leftovers (whatever their disposition) are not included in the computation of calorie and protein consumption. Each person present at every meal is accounted for, including non-family and non-household members, so there is no problem of overestimating household *per capita* consumption on this account. Further, since data were collected over a seven day period, random variation from either independent measurement errors or daily dietary variation will be partly smoothed.

Nutrient requirements are lower for children than adults and, since poorer households typically have more children, failure to account for household composition is likely to result in a systematic upward bias in the estimated elasticity. There is, however, no correct method of controlling for composition. Intakes may be converted to an adult equivalent basis using, for example, international weights (FAO-WHO, 1973) and household expenditure could be similarly adjusted; it is, however, far from clear what weights should be attributed to each demographic group (Deaton and Muellbauer, 1986). We prefer, therefore, to include demographic controls in the regressions and permit the data to choose the appropriate weights. Family size and composition might properly be treated as endogenous, however, and, without any instruments to explain household composition which should not directly enter the demand function, estimates including demographic controls will be biased. Estimates will thus be presented including and excluding demographic controls as well as estimates based on a sample of households stratified on household composition.

A model explaining calorie intake should also include controls for activity levels of each individual; these are hard to measure and should clearly be treated as endogenous (Pitt, Rosenzweig and Hassan, 1989). We have no instruments with which to predict activity but, in an attempt to account for gross differences, examine rural and urban households separately.

An issue which has received surprisingly little attention in the nutrition literature is nonlinearity. For very poor households — many of whom spend around 60 to 70 per cent of their budget on food — it seems implausible that there should be no response of calories to expenditure. As the food share declines with income and people switch to higher valued foods and purchase more non-nutrient qualities (such as taste), the elasticity will surely decline with income. Elasticity estimates reported at the mean (as in Table 1) may be considerably lower than those for the poorest households, who are often of most interest. Those studies which do allow for some non-linearity, usually by introducing a quadratic term in expenditure or income, typically find elasticities do decline with expenditure.<sup>11</sup>

In descriptive studies, Poleman (1981) and Lipton (1983) have argued that nonlinearities may be quite sharp and suggest the calorie-income curve may be kinked. Poleman (p. 23) plots grouped data from a 1969/70 expenditure survey in Sri Lanka; calories rise with expenditure until an asymptote at about 2500 calories per person per day and then remain constant. He also shows that, among the poorest, the composition of foods consumed is essentially unchanged as income rises and it is only among the relatively better off that households switch into foods with a higher price per calorie. A related, but stronger hypothesis postulated by both Poleman and Lipton, is that the budget share of foods may actually increase with income for very poor households. There is some evidence to support this view (Thomas, 1986) and it is hard to imagine food shares rising without a concomitant increase in calorie intake.

### 3. Results

We shall exploit the size of ENDEF by stratifying in several dimensions. The focus in this paper will be on urban households; results for the rural sector are qualitatively similar although some differences will be pointed out. The sample of urban households is still very large and so we have split it into three random sub-samples with about 12,900 households in each.

The first sub-sample will be used for exploratory data analysis, including non-parametric estimation. These results will guide parametric estimation of the nutrient-expenditure curve — in particular the specification of the functional form and the covariates to be included in the model. The first sample will also be used to determine the appropriate choice of instruments in the two stage regressions. All the estimated elasticities are reported for the second sample; size of tests do not need to be adjusted (Lovell, 1983) since these estimates have not been used to guide any further data analysis.

#### (i) Exploratory data analysis : non-parametric estimates

Using the first sample, we begin with a contour plot (or three-dimensional histogram estimator) of the bivariate relationship between (log) *per capita* nutrient intake and (log) *per capita* expenditure (PCE); calories and protein are the measures of nutrients in Figures 1a and 1b respectively.<sup>12</sup> The relationships look quite nonlinear — especially for calories — although around median  $\ln PCE$  (which is 8.2) non-linearity seems unimportant and the correlation between nutrients and PCE is clearly positive.

Figure 1 presents a lot of detail. One way of summarizing the essence of the relationship, while imposing little structure, is to use a two-dimensional non-parametric estimator. Our first estimator is probably the best known: mean (log) nutrient intake is plotted against (log) PCE where means are defined for

each icosohile (5 percentile) of PCE. The shape of the calorie-expenditure curve is dramatic. In both the urban and rural sectors, there is a kink around the third quartile of PCE (at about 2,400 calories *per capita* per day in the urban sector and 3,000 calories in rural households<sup>13</sup>); the curve is positively sloped for poorer households and essentially flat for households in the top quartile of PCE. The protein-expenditure curve is positively sloped for all households; it tends to flatten out for higher expenditure households although the threshold is less dramatic.

Figure 2 has the force of simplicity. There are, however, a large number of non-parametric estimators in the literature which have theoretically more attractive properties and can be generalized to a multivariate context. (Hardle, 1988). A locally-weighted smoothed scatterplot (LOWESS) (Cleveland, 1979) is a nearest neighborhood-type estimator, in which each observation is replaced by its predicted value based on a weighted regression using the observations in a band around it.<sup>14</sup>

Figures 3a and 3b present LOWESS estimates of the nutrient-expenditure curves using a tri-cube weighting function, or kernel, for about 6000 urban households (randomly chosen from the first sub-sample). In addition to choosing the kernel, it is necessary to choose the bandwidth; one strategy is to use an automatic technique such as cross-validation (Efron, 1982). Since our aim is to uncover the shape of the nutrient-expenditure function, we have experimented with a range of bandwidths from 60 observations (1% of the sample) to 1200 observations (20%). In the top left hand panel, the bandwidth is 120 observations; clearly the curve is undersmoothed. When 300 observations (5% of the sample, top right hand panel) are included in each band, the shape of the curve is easier to identify and looks much like the icosohile means in Figure 2.

Enlarging the bandwidth smooths the curve without affecting its shape. The protein curve is similar although it does not flatten out as dramatically.

Expenditure elasticities of the demand for nutrients are apparently not zero — at least for poorer households. There is also evidence for the kind of threshold suggested by Poleman (1981) and Lipton (1981) — especially for calories. As expenditure rises, poor households purchase additional protein and calories in addition to other nutrients and non-nutrient qualities. When calories reach some threshold, however, households switch to higher protein foods while maintaining an approximately constant level of calorie intake. Failure to take account of this non-linearity will result in estimated elasticities which are biased downwards for the poorest. Our estimates from a regression linear in logs, discussed in the previous section, are almost certainly too low.

#### (ii) Parametric estimates

In addition to expenditure, nutrient intakes are likely to depend on prices and household characteristics. Although non-parametric estimates provide very useful guidance about the general shape of a curve in two dimensions, it is well known that in multivariate models they suffer from the "curse of dimensionality" (Huber, 1985): as the number of continuous covariates increases, the cells over which there is smoothing quickly become very wide. Asymptotically, additional discrete covariates have no impact on the rate of convergence of local estimators, (Bierens, 1987), suggesting that continuous covariates be made discrete. It is clear, however, that first order asymptotic results are not sufficient from a practical point of view. In almost all currently available samples, stratification on many discrete covariates will result in small subsamples. We take the view, therefore, that Figures 2 and 3 provide very valuable

input into the choice of a parametric functional form and therefore turn next to the estimation of parametric relationships.

A common functional form in the nutrition literature is a regression of (log) *per capita* nutrient intake,  $\ln N$ , on (log) *per capita* expenditure,  $\ln PCE$ , and, in some cases, a quadratic term in (log) *per capita* expenditure, including a set of other conditioning covariates,  $W$ , such as prices or regional dummies, seasonal dummies and household characteristics:

$$\ln N = \beta_0 + \beta_1 \ln PCE + \beta_2 \ln PCE^2 + W\gamma + \epsilon \quad [1]$$

Since the elasticity is a linear function of  $\ln PCE$ , this curve is unlikely to be able to pick up the kink in the nutrient functions so that elasticities at low expenditure levels may be under-estimated. It is also possible that a quadratic function will force elasticities at high expenditures to be negative. One might include higher order polynomials, such as cubics in  $\ln PCE$ , although it is well known that they too may suffer from tail-behavior dictating the shape of the curve in other parts of the distribution.

$$\ln N = \beta_0 + \beta_1 \ln PCE + \beta_2 \ln PCE^2 + \beta_3 \ln PCE^3 + W\gamma + \epsilon \quad [2]$$

Instead of resorting to semi-parametric methods such as series estimators which may also suffer from influential observations in the tails<sup>15</sup>, we propose a simple parametric form which can, in principle, replicate the shape of Figure 3: the log-inverse log model. To add flexibility, we also include a quadratic in inverse  $\ln PCE$ :

$$\ln N = \beta_0 + \beta_1 \ln PCE^{-1} + \beta_2 \ln PCE^{-2} + W\gamma + \epsilon \quad [3]$$

the elasticity being:

$$\eta = \frac{\partial \ln N}{\partial \ln PCE} = -\beta_1 \ln PCE^{-2} - 2\beta_2 \ln PCE^{-3}$$

We shall refer to [3] as the quadratic-inverse model. A two-parameter model may not be sufficiently flexible; instead of adding  $\ln\text{PCE}^{-3}$ , the logarithm of  $\ln\text{PCE}$  is included with the covariates:

$$\ln N = \beta_0 + \beta_1 \ln\text{PCE}^{-1} + \beta_2 \ln\text{PCE}^{-2} + \beta_3 \ln(\ln\text{PCE}) + W\gamma + \varepsilon \quad [4]$$

in which case the elasticity is a very flexible cubic function of inverse  $\ln\text{PCE}$ :

$$\eta = \frac{\partial \ln N}{\partial \ln\text{PCE}} = \beta_3 \ln\text{PCE}^{-1} - \beta_1 \ln\text{PCE}^{-2} - 2\beta_2 \ln\text{PCE}^{-3}$$

This model will be called the quadratic-inverse-log.

Using the second sub-sample of urban households, models [1] through [4] have been estimated for *per capita* calorie and protein intake. Each regression includes an additional 40 covariates (education of the head and spouse, age and sex of the head, existence of the spouse, twelve local market food prices, seven region dummies and twelve month dummies) which are suppressed from the presentation. Since the expenditure coefficients are hard to interpret, Table 2 presents elasticities evaluated at seven points in the PCE distribution and, beneath each estimate, its jackknifed standard error (Efron, 1982)<sup>16</sup>.

With large sample sizes, it is far from clear that conventional testing strategies are appropriate. Holding the probability of Type 1 errors constant, then as the sample size increases, the probability of Type 2 errors declines and so, it may be prudent to adopt a non-classical approach to choosing the size of the test. Schwarz (1978) proposed a Bayesian model selection criterion, related to the Akaike information criterion, (Akaike, 1978), which will asymptotically pick the model that is *a posteriori* most probable. For a  $\chi^2$  test statistic, the critical value is the logarithm of the sample size multiplied by the number of restrictions,  $r$ ; for our sample, which is over 12,800 observations, the critical value is  $9.5 \cdot r$ . Since the square of a  $t$  variate is asymptotically a  $\chi^2_1$ , the critical value of a  $t$  statistic is 3.08 according to the Schwarz criterion. We shall judge significance by the more stringent Schwarz criterion in the text,

although, for reference, standard errors and classical probability values are presented in the tables.

It is well known that the performance of high order polynomial models estimated by least squares can suffer from influential observations in the tails. The addition of high order polynomial expenditure terms in the regressions results in poor tracking of the non-parametric estimates in the tails of the expenditure distribution. Between the bottom and top deciles of PCE, however, estimated elasticities in models with four and five expenditure terms are remarkably similar to those with two or three expenditure variables. We restrict attention, therefore, to the two and three parameter models [1] through [4]. In all the regressions, expenditure has a significant impact on both calories and protein (with  $\chi^2$ s over 2000).

In the calorie regressions, each of the expenditure coefficients is individually significant. The two parameter models fail to track the non-parametric estimates as well as the three parameter models largely because of difficulty in replicating the kink. The quadratic polynomial model underestimates the elasticity at the bottom of the expenditure distribution; at the top, the elasticity is negative (-0.32 at the maximum reported PCE). The quadratic inverse model, [3], tracks the non-parametric curve better, especially in the lower deciles of expenditure but it is not flat in the upper deciles. The cubic polynomial model performs well and is flat at in the top deciles of PCE but even it is not sufficiently non-linear around the kink. The quadratic-inverse-log model, [4], also tracks the non-parametric estimate well, picks up the non-linearity better but it declines at the top of the PCE distribution displaying the same problem as in the quadratic polynomial model. Informally, among the parametric models we prefer the estimated elasticities in columns 2 and 4.



At the bottom decile of PCE, the calorie elasticity with respect to expenditure is over 0.40. It declines with expenditure to about half (just over 0.2) at the median and at the top decile of PCE it is quite small (about 0.1). In all cases, the elasticities are estimated very precisely and are significantly different from zero throughout the expenditure distribution.

In the protein regressions, each expenditure term is individually significant except in the cubic polynomial model: the cubic term is irrelevant and this model almost exactly tracks the quadratic model. The estimated function declines at the top of the expenditure distribution (with estimated elasticities of -0.19 to -0.26 at maximum observed PCE) for all models except the quadratic-inverse [3]. The estimated elasticities range from 0.53 at the bottom decile to 0.16 at the top, are very precise and are all significantly different from zero.

### (iii) Non-parametric instrumental variables estimates

It was argued in Section 2 that *per capita* expenditure and nutrient intakes are likely to share common measurement errors, in which case the estimates discussed above are biased and instrumental variables is the appropriate estimator. One might assume that the common measurement errors are uncorrelated with errors in household income, since income and expenditure data are collected independently. Income will then be a valid instrument, ignoring, for the moment, labor supply and efficiency wage considerations. Adopting a proposal of Newey and Powell, (1989), we present bivariate non-parametric instrumental variable estimates of the relation between nutrient intake and PCE in Figure 4. In the first step, the relation between  $\ln PCE$  and (log) total household income,  $\ln HHY$ , is estimated non-parametrically:

$$\ln PCE = f(\ln HHY) + u$$

The estimated residuals from that regression,  $\hat{u}$  are included in a second stage (multivariate) non-parametric regression:

$$\ln N = g(\ln PCE, \hat{u}) + \varepsilon$$

Again, the unknown functions  $f()$  and  $g()$  are estimated by LOWESS with the same sub-sample of urban households used in the non-parametric estimates discussed above.<sup>17</sup> The results are presented in Figure 4.<sup>18</sup>

The shape of the calorie-expenditure function, in the upper panel, is similar to the pictures in Figure 3. The curve is upward sloping for the poorest households, there is a kink around the third quartile of PCE ( $\ln PCE=8.8$ ) and the curve is flatter for households in the top quartile. The curve does not completely flatten out although this is mostly due to households with absolutely large residuals whose calorie intakes are higher than households with the same PCE but smaller residuals; there is a fair amount of noise at the top of the PCE distribution. For most households, namely those to the left of the kink, the plane of the function is downward sloping as residuals increase. This makes good sense if the residuals reflect measurement error in PCE which is not also in nutrient intake measures. The protein curve is in the lower panel; it is almost linear and, like calories, the plane slopes down in residual space.

#### (iv) Tests for exogeneity

There is a long tradition in economics which argues that household income should not be treated as exogenous. Firstly, expenditures are likely to depend on labor supply decisions; secondly, the decision to work or, at least, one's productivity may depend on nutrient intakes. Household unearned income, a measure of wealth, may be a better instrument. No unearned income is reported by 45% of all urban households (and almost two thirds of rural households). This puts a (rather high) lower bound on the size of the bandwidth in a local

estimator (including all nearest neighborhood and most kernel estimators). We turn, therefore, to parametric estimation to test exogeneity assumptions.

Still using the first sub-sample of urban households, Table 3 presents a series of Wu-Hausman tests for each of the four parametric models. Since we are primarily interested in the impact of expenditure on nutrient intake, and since Wu-Hausman tests tend to lack power as the dimension of a test rises, we test for differences between the expenditure coefficients in the regressions. It turns out that these  $\chi^2$ s test statistics are very similar to those including all covariates; of course, the degrees of freedom are not the same. In each of the 2SLS regressions, the identifying instruments are included as third order polynomials and the test statistics are based on jackknifed estimates of the variance-covariance matrix.

Classical probability values are reported beneath each statistic. According to the Schwarz criterion, the critical value for the two and three parameter models are 18.9 and 28.4 respectively. Maintaining that unearned income is exogenous, then  $\ln PCE$  is also exogenous in the protein regressions. Exogeneity of  $\ln PCE$  is rejected unambiguously in the calorie regressions.

Still maintaining unearned income is exogenous — and that measurement errors are uncorrelated with errors in the nutrient-expenditure function — is total household income a valid instrument?<sup>19</sup> The test statistics are reported in the second row of each panel of Table 3. It is impossible to distinguish between the estimates using the two sets of instruments and so, on efficiency grounds, household income is the appropriate instrument. Under classical testing, however, this inference would change in the two-parameter calorie model.

Assuming that total household income is uncorrelated with errors in the nutrient-expenditure function, then the hypothesis that  $\ln PCE$  is exogenous is unambiguously rejected in all models for both calories and protein with  $\chi^2$ s of

over 250 and 80 respectively. These results suggest the tests are not without power.

(v) Parametric instrumental variables estimates

Returning to the second sub-sample of urban households, Table 4 presents estimates of the expenditure elasticities of nutrient demand based on instrumental variable estimates with the same covariates used in the OLS regressions. With (log) household income as the identifying instrument, in the quadratic models, the expenditure elasticity at the bottom of the PCE distribution declines by about a third relative to the OLS estimates. At the bottom decile, the elasticity is 0.26; it is halved (0.11) at median PCE and is zero at the top of the PCE distribution. Although standard errors are about three times larger than in the OLS estimates, the elasticities are significant at least for all but the top decile of households. In the three parameter models, the elasticities at low levels of PCE are higher than in the two parameter models, and approximately the same as the OLS estimates. The three parameter IV estimates decline more rapidly than the OLS ones although at the top of the distribution, the IV elasticities increase sharply. The shape of the curve at low PCE is almost surely influenced by this positive slope at the top.

When unearned income is used as an instrument in the two parameter models, the elasticities at low levels of PCE are one third smaller than those instrumented with total income; these differences decline with PCE. The standard errors are much larger over the entire PCE distribution — and at low levels of PCE the elasticities are not significantly different from zero.

When a third expenditure term is added to the model, the elasticities change drastically and the standard errors become huge. At the bottom decile of PCE, the elasticity in the quadratic-inverse-log model is 2.3 but the jackknifed

standard error is 9.5. Throughout the PCE distribution, we cannot reject the hypothesis that there is no relation between income and calories.

Recall that 45% of households report no unearned income. For low expenditure households, unearned income is a poor instrument although over the entire PCE distribution the first stage regressions appear to perform fairly well.<sup>20</sup>

It appears there is a subtle interaction between the quality of the instruments and the functional form of the second stage regression. The three parameter models instrumented with unearned income are sufficiently flexible to allow tail observations to be very influential resulting in their poor performance. In contrast, the noise in the tails of the expenditure distribution is not as influential in the two parameter models.

The results for protein are much the same although, relative to least squares, the proportionate decline in elasticities is smaller than for calories. For the two parameter models, estimated elasticities are about 20% smaller at the bottom of the expenditure distribution when (log) total income is the instrument and about 40% smaller when instrumented with unearned income. In the latter case, the three parameter model estimates display the same behavior as in the calorie regressions.

Since the elasticity of demand for protein is about 0.1 higher than it is for calories, throughout the PCE distribution, all households switch into foods with higher protein content as expenditure rises. This results in a rise in the proportion of calories consumed from protein, the protein-calorie ratio, which has been suggested by nutritionists as an indicator of diet quality (see Pacey and Payne, 1985). Its expenditure elasticity, calculated directly, is about 0.20 at the bottom decile of PCE and 0.12 at the top decile.

Foods with higher protein content, such as meats and dairy products, tend to have higher unit prices and so the price of nutrients rises with expenditure. In fact, the implicit price of calories, household expenditure on food divided by calorie consumption, increases linearly with  $\ln PCE$ ; the expenditure elasticity is 0.40 at the bottom decile of PCE and 0.30 at the top.

The shapes of the nutrient-expenditure functions are quite similar in the rural sector. Instrumenting with (log) household income, the calorie elasticity is 0.32 at the bottom decile and 0.10 at the top; the threshold in the calorie curve is around 3000 calories *per capita per diem* which is higher than in the urban sector. With the same instruments, the protein-expenditure function is almost linear and the elasticity is around 0.23 throughout the distribution. As a result, the expenditure elasticity of the protein-calorie ratio rises with PCE from zero at low PCE to about 0.13 at the top decile of PCE. The calorie price elasticities are almost identical to those in the urban sector.

As expenditure rises, therefore, households purchase more expensive sources of calories, demand more non-calorie nutrients and, presumably, more taste. They also, however, buy more calories, at least until some threshold level of calorie intake.

#### (vi) Household composition effects

Standard demand theory would suggest household composition should enter the nutrient demand functions; it may be argued, however, that composition is endogenous. Since household size declines with PCE and larger households tend to have more children who consume fewer nutrients than adults, it is possible that the positive nutrient elasticities simply reflect changes in household composition. The robustness of the estimates to the inclusion of demographics is examined in Table 5.

The first column repeats the estimates for the quadratic inverse model instrumenting with total income (column 3 of Table 4). Implicitly, through the use of *per capita* expenditure, total household size (and an interaction with expenditure) is included in this model. More complicated demographics are included in the second column. It is common to include the number of people in each of a set of demographic groups in a linear fashion. We prefer to allow a more complex structure and include the ratio of the number of people in each group to household size (excluding one of the ratios to avoid linear dependence). In addition, we include the log of household size. Five age groups (0-4, 5-9, 10-14, 15-54 and  $\geq 55$ ) are distinguished by gender.

Changes in household composition do not explain the positive slope of the nutrient functions; in fact, the elasticities rise for poor households as the restriction tying the impact of expenditure and household size is relaxed. For calories, at the bottom decile of PCE, the elasticity rises by about 2 standard errors to 0.31 and by 1 standard error for protein to 0.42.

There may also be complicated interactions between the demographic variables and expenditure; exploiting the large sample size again, we stratify on household composition. In column three, households with two adults (one male, one female) and one child (aged less than ten) are included in the sample. The elasticities at the bottom of the expenditure distribution are higher: 0.35 for calories and 0.42 for protein at the bottom decile of PCE. The results are essentially the same if the sample is restricted to households with two adults and a child less than five years of age. The sample in column 4 includes households with two adults and two children. The elasticities are no different from those in the first column (which ignores demographics). Changing household composition does not explain the positive nutrient intake elasticities.

The estimates of the demographic effects are interesting in and of themselves. In the lower panel of Table 5, the coefficients from column 5 have been converted into "outlay equivalent ratios" (Deaton et al., 1989):

$$\pi_d = \frac{\partial N / \partial X}{\partial N / \partial m_d} \cdot \frac{m}{X}$$

where N is nutrient intake, X is total household expenditure, m the number of members in the household and  $m_d$  the number in demographic group d. The ratios indicate the amount of expenditure that a household would need to be given in order for *per capita* nutrient intake to remain unchanged when a person of type d joins the household; the ratio is divided by *per capita* expenditure to turn it into an elasticity.

Demographic composition would have no impact on nutrient intake if the  $\pi$  ratios were zero. Household composition clearly affects nutrient intake; each  $\pi$  ratio is significantly negative and the  $\chi^2$  for joint significance of the demographic ratios is 76.0 in the calorie regression, 43.3 in the protein regression. An additional child reduces *per capita* nutrient intake more than adults and, apart from infants and young children, women reduce *per capita* intake less than men.<sup>21</sup> The implied weights that might be associated with each demographic group (or equivalence scales) differ substantially in the calorie and protein regressions. Infants are equivalent to about 20% of a prime age male adult in calorie consumption and about 30% in protein consumption. Prime age women are equivalent to about 30% of a male in calorie consumption and 50% in protein consumption. The equivalence ratios for adult females are very small relative to international standards based on requirements that reflect only nutrient needs. This suggests that converting household information from survey data to an adult equivalent basis with these sorts of scales is not a good idea.



#### 4. Conclusions

Both the parametric and nonparametric evidence presented here reject the view that among poor households, nutrient intakes do not respond to changes in expenditure. In fact, our nonparametric estimates demonstrate that as expenditure rises, calorie consumption increases until some threshold level at which point households switch to foods with higher calorie prices. This is confirmed by elasticities computed from parametric regressions.

The estimated calorie and protein expenditure elasticities are 0.38 and 0.47 respectively at the bottom decile of *per capita* expenditure when it is treated as either measured without error or exogenous. This hypothesis is rejected with the ENDEF data, but while the elasticities decline when PCE is instrumented with household income they remain large. The elasticities are estimated to be between 0.25 to 0.30 for calories and 0.40 for protein at the bottom decile of PCE.

Clearly some of the heterogeneity in estimates of calorie elasticities in Table 1 can be explained by failure to instrument income or expenditure. In Section 2, we reported that the instrumented calorie elasticity is 0.12 when the calorie-expenditure function is log-linear. Previous studies which have failed to take account of non-linearities in the calorie-expenditure function may have severely underestimated the elasticity for poor households.

Methodological differences are probably not the whole story. Food intake measurements are likely to be quite noisy especially when households are visited only a small number of times. Large sample sizes, on the other hand, may compensate for difficulties in measurement. Differences in data quality and sample sizes almost surely contribute to an explanation of the heterogeneity of the estimates in Table 1. In any case, the view that nutrient elasticities are, in general, zero for the poor is premature.

## Notes

<sup>1</sup>Sukhatme (1982), Beaton (1983) and Pacey and Payne (1985) discuss the issue of body adaptation to deprivation, which is still unresolved both as to its magnitude and to the extent which it prevents functional disabilities.

<sup>2</sup>Excluding the sparsely populated rural North.

<sup>3</sup>Williamson-Gray (1982) has used aggregated data from ENDEF with cell means for 9 income classes and 22 regions to examine demand for nutrients and foods.

<sup>4</sup>Food waste includes gifts, leftovers, food given to animals and food thrown away.

<sup>5</sup>Households were defined using a dwelling definition.

<sup>6</sup>For poor workers this may understate nutrient intakes if an employer provides a disproportionately large share of the worker's nutrients through meals provided at work.

<sup>7</sup>Food consumption data have been converted to nutrient intakes and the calorie-expenditure relation estimated directly on the same data; the estimated elasticity is about 0.86 when expenditure is treated as exogenous and 0.99 when instrumented. The elasticity is essentially constant throughout the expenditure distribution.

<sup>8</sup>See Strauss (1986), Deolalikar (1988), Sahn and Alderman (1988) and Pitt, Rosenzweig and Hassan (1989) for empirical examples.

<sup>9</sup>The sign of this bias is *a priori* ambiguous; assuming the covariance in measurement errors is large relative to the variance in the measurement error in total expenditure, estimated elasticities will be upward biased.

<sup>10</sup>If the absolute magnitude of these errors increases as a function of the difference between total expenditure and mean expenditure, then a further upward bias is imparted to OLS estimates, which will not be removed by instruments correlated with income. A fixed effects estimator, provided panel data are available, will remove this source of bias (Bouis and Haddad, 1989). So too should netting out these flows — the strategy we adopt. Fixed effects estimates have the disadvantage that random measurement error in total expenditure, or income, will bias estimated coefficients downwards (Griliches and Hausman, 1986). Using income rather than expenditure is likely to exacerbate this problem. This may explain the low estimates reported by Bouis and Haddad.

<sup>11</sup>See, for example, Timmer and Alderman (1979), Pitt (1983) Chernichovsky and Meesook (1984), Garcia and Pinstrip-Anderson (1987), Ravallion (1988) and Sahn (1988), but see Strauss (1984) for an exception.

<sup>12</sup>The band width for each cell is 0.05 for the log of *per capita* calories and protein and it is 0.2 for the log of *per capita* expenditure.

<sup>13</sup>If calorie intake were measured per adult equivalent, rather than *per capita*, then the kink would be at a higher intake. Some of the difference between the rural and urban sector probably reflects higher activity levels in rural areas; it may also be that there is under-reporting of nutrient intake in urban areas because more meals are taken away from home.

<sup>14</sup> LOWESS estimates are calculated by creating a band around each observation ( $y_i$ ,  $x_i$ , say) and estimating a weighted linear regression of the dependent variable on the independent variables. Observations within the band are weighted by the tricube function,  $w_j = (1 - d_j^3)^3$ , where  $d_j$  is measure of the distance between observation  $j$  and the observation of interest,  $i$ . In the simple regression case,  $d_j = (x_j - x_i) / (x_i^* - x_i)$  where  $x_i^*$  is the furthest observation from observation  $i$  within the band. The weight is positive for each observation within the band; the weight is equal to one at the point itself, declines as points are further away and observations outside the band are given a weight of zero. The predicted value of  $y_i$  is then plot against  $x_i$  to form the LOWESS estimates. The estimator is easily generalized to the multivariate regression case (Cleveland, Devlin and Grosse, 1987) by defining a distance metric such as Euclidean distance. 6000 observations are used in the non parametric estimates to reduce computational time.

<sup>15</sup>We have attempted to estimate some series estimators. The model is linearly dependent if more than six expenditure terms are include in the OLS regressions and only five expenditure covariates can be included in the IV regressions. For such a large dataset, we would be restricted to a very low order series estimator.

<sup>16</sup>Jackknifed estimates of the variance-covariance matrix are consistent in the presence of heteroskedasticity and unknown forms of mis-specification. They are, in addition, more robust to influential observations than the infinitesimal jackknife or White (1980, 1982) estimator of variance.

<sup>17</sup>Euclidean distance is used for the estimation of  $g()$ .

<sup>18</sup>The bandwidths for both the first and second stage regressions is 300 observations; when the bands are increased (to 2400 observations) the shapes are much the same.

<sup>19</sup>By the Akaike information criterion, the model with total income is preferred to including earned and unearned income separately. Taking logarithms of total household income further improves the fit.

<sup>20</sup>In the  $\ln PCE$  regression, each of the three unearned income coefficients is individually significant; they are jointly significant ( $F=398$ ) and overall 55% of the variation in  $\ln PCE$  is explained.

<sup>21</sup>The hypothesis that an additional male has the same impact on nutrient intake as an additional female is rejected for all age groups except children less than 10.

## References

- Akaike, H., (1978), A Bayesian Analysis of the Minimum AIC Procedure, Annals of the Institute of Statistics and Mathematics, A30:9-14.
- Alderman, H., (1987), Cooperative Dairy Development in Karnataka, India: An Assessment, International Food Policy Research Institute, Research Report No. 64, Washington, D.C.
- Alderman, H., (1989), New research on poverty and malnutrition: What are the implications for research and policy?, mimeo, IFPRI and Cornell University.
- Bierens, H. J., (1987), Kernel estimators of regression functions, in T. F. Bewley (ed.), Advances in Econometrics, Cambridge University Press, New York.
- Beaton, G. H., (1983), Energy in Human Nutrition: Perspective and Problems, Nutrition Reviews, 41:325-340.
- Behrman, J. R., and A. B. Deolalikar (1987), Will Developing Country Nutrition Improve With Income? A Case Study for Rural South India, Journal of Political Economy, 95:492-507.
- Behrman, J. R., and A. B. Deolalikar (1988), Health and Nutrition, in H. Chenery and T. N. Srinivasan (eds.), Handbook of Development Economics, Vol. 1, Amsterdam, North Holland.
- Behrman, J. R., and B. L. Wolfe (1984), More Evidence on Nutrition Demand: Income Seems Overrated and Women's Schooling Underemphasized, Journal of Development Economics, 14:105-128.
- Behrman, J. R., A. B. Deolalikar, and B. L. Wolfe (1988), Nutrients: Impacts and Determinants, World Bank Economic Review, 2:299-320.
- Berg, A., (1973), The Nutrition Factor: Its Role in National Development, Brookings Institute, Washington D.C.
- Bouis, H. E., and L. J. Haddad (1988), Comparing Calorie-Income Elasticities Using Calories Derived from Reported Food Purchases and a Twenty-Four Hour Recall of Food Intakes: An Application Using Philippine Data, mimeo.
- Chernichovsky, D., and O. Meesook (1984), Patterns of Food Consumption and Nutrition in Indonesia: An Analysis of the National Socioeconomic Survey, 1978, World Bank Staff Working Paper No. 670, Washington, D.C.
- Cleveland, W. S., (1979), Robust Locally-Weighted Regression and Smoothing Scatterplots, Journal of the American Statistical Association, 74:829-836.
- Cleveland W.S., S.J. Devlin and E. Grosse, (1988), Regression By Local Fitting: Methods, Properties and Computational Algorithms, Journal of Econometrics, 37:87-114.

- Deaton, A., and J. Muellbauer (1986), On Measuring Child Costs: With Applications to Poor Countries, Journal of Political Economy, 94:720-744.
- Deaton A.S., J. Ruiz-Castillo and D. Thomas, (1989), The Influence of Household Composition on Household Expenditure Patterns: Theory and Spanish Evidence, Journal of Political Economy, 97:179-200.
- Deolalikar, A. B., (1988), Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India, Review of Economics and Statistics, 70:406-413.
- Edirisinghe, N., (1987), The Food Stamp Scheme in Sri Lanka: Costs, Benefits, and Options for Modification, International Food Policy Research Institute Research Report No. 58, Washington, D.C.
- Efron, B., (1982), The Jackknife, the Bootstrap and other Resampling Plans, SIAM CBMS-NSF Monograph, 38.
- Food and Agricultural Organization and World Health Organization, (1973), Energy and Protein Requirements, WHO Technical Report Series No.522, Geneva.
- Garcia, M., and P. Pinstrip-Andersen, (1987), The Pilot Food Price Subsidy Scheme in the Philippines: Its Impact on Income, Food Consumption and Nutritional Status, International Food Policy Research Institute Research Report No. 61, Washington, D.C.
- Griliches, Z., and J. A. Hausman (1986), Errors in Variables in Panel Data, Journal of Econometrics, 31:93-118.
- Hardle, W., (1988), Applied Nonparametric Regression, Faculty of Statistics, Rheinische-Friedrich-Wilhelms University, Bonn.
- Hausman, J. A., (1978), Specification Tests in Econometrics, Econometrica, 46:1251-1272.
- Huber, P. J., (1985), Projection Pursuit, Annals of Statistics, 13, 435-475
- Lipton, M., (1983), Poverty, Undernutrition and Hunger, World Bank Staff Working Paper No. 597, Washington, D.C.
- Lovell, M. C., (1983), Data mining, Review of Economics and Statistics, 55.1:1-11.
- Newey, W. and J. Powell, (1989), Instrumental variables estimation for nonparametric models, mimeo, Princeton University.
- Pacey, A., and P. Payne (eds.) (1985), Agricultural Development and Nutrition, Food and Agriculture Organization, Rome.
- Pinstrip-Andersen, P., and E Caicedo (1978), The Potential Impact of Changes in Income Distribution and Food Demand and Human Nutrition, American Journal of Agricultural Economics, 60:402-415.

- Pitt, M. M., (1983), Food Preferences and Nutrition in Rural Bangladesh, Review of Economic Statistics, 65:105-114.
- Pitt, M. M., M. R. Rosenzweig, and N. Hassan (1989), Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries, processed.
- Poleman, T. T., (1981), Quantifying the Nutrition Situation in Developing Countries, Food Research Institute Studies, 18:1-58.
- Ravallion, M., (1988), Income Effects of Caloric Undernutrition, processed.
- Reutlinger, S., and M. Selowsky (1976), Malnutrition and Poverty: Magnitude and Policy Options, Baltimore, Johns Hopkins University Press.
- Sahn, D., (1988), The Effect of Price and Income Changes on Food-Energy Intake in Sri Lanka, Economic Development and Cultural Change, 36:315-340.
- Sahn, D., and H. Alderman (1988), The Effects of Human Capital on Wages, and the Determinants of Labor Supply in a Developing Country, Journal of Development Economics, 29:157-184.
- Schwarz, G., (1978), Estimating the Dimension of a Model, Annals of Statistics, 6:461-464.
- Strauss, J., (1984), Joint Determination of Food Consumption and Production in Rural Sierra Leone: Estimates of a Household-Firm Model, Journal of Development Economics, 14:77-103.
- Strauss, J., (1986), Does Better Nutrition Raise Farm Productivity?, Journal of Political Economy, 94:297-320.
- Sukhatme, P. V., (1970), Incidence of Protein Deficiency in Relation to Different Diets in India, British Journal of Nutrition, 24:447-487.
- Sukhatme, P. V., (ed.), (1982), Newer Concepts in Nutrition and Their Implications for Policy, Maharashtra Association for the Cultivations of Science Research Institute, India.
- Thomas, D., (1986), Can the food share be used as a welfare measure?, unpublished Ph.D. dissertation, Princeton University.
- Timmer, C. P., and H. Alderman (1979), Estimating Consumption Parameters for Food Policy Analysis, American Journal of Agricultural Economics, 61:982-987.
- Trairatvorakul, P., (1984), The Effects on Income Distribution and Nutrition of Alternative Rice Price Policies in Thailand, International Food Policy Research Institute Research Report No. 46, Washington, D.C.

- von Braun, J., D. Puetz, and P. Webb (1989), Irrigation Technology and Commercialization of Rice in the Gambia: Effects on Income and Nutrition, International Food Policy Research Institute Research Report No. 75, Washington, D.C.
- Ward, J. O., and J. H. Sanders (1980), Nutritional Determinants and Migration in the Brazilian Northeast: A Case Study of Rural and Urban Ceara, Economic Development and Cultural Change, 29:141-163.
- White, H., (1980), A heteroskedasticity consistent covariance matrix and direct test for heteroskedasticity, Econometrica, 48:817-38.
- White, H., (1982), Instrumental Variables Regression With Independent Observations, Econometrica, 50:483-500.
- Williamson-Gray, C.,(1982),Food Consumption Parameters for Brazil and Their Application to Food Policy, International Food Policy Research Institute Research Report No.32, Washington D.C..
- Wolfe, B. L., and J. R. Behrman (1983), Is Income Overrated in Determining Adequate Nutrition, Economic Development and Cultural Change, 31:525-550.
- Wu, D-M., (1973), Alternative Test of Independence Between Stochastic Regressors and Disturbances, Econometrica, 41:733-750.

Table 1

## Estimated income and expenditure elasticities of calorie demand

Author(s)	Year	Expenditure or Income	Country studied	Elasticity at Mean
Calculated from Food Demand Equations:				
Pitt	1983	Expenditure <sup>a</sup>	Bangladesh	0.78-0.82
Behrman and Deolalikar	1987	Expenditure	India	0.77-1.18
Strauss	1984	Expenditure	Sierra Leone	0.82
Sahn	1988	Expenditure	Sri Lanka	0.62
Pinstrup-Anderson and Caicedo	1978	Income	Colombia	0.51
Calculated from Calorie Demand Equations:				
Behrman and Deolalikar	1987	Expenditure	India	0.17-0.37 <sup>b</sup>
Alderman	1987	Expenditure	India	0.41-0.44
Timmer and Alderman	1979	Expenditure	Indonesia: urban	0.26
			rural	0.51
Chernichovsky and Meesook	1984	Expenditure	Indonesia	0.54
Ravallion	1988	Expenditure	Indonesia	0.15
Garcia and Pinstруп- Anderson	1987	Expenditure	Philippines	0.12-0.34
Bouis and Haddad	1988	Expenditure	Philippines	0.05-0.45
Edirisinghe	1987	Expenditure	Sri Lanka	0.56
Ward and Sanders	1980	Income	Brazil	0.24-0.53
von Braun, Puetz and Webb	1989	Income	Gambia	0.37-0.48
Wolfe and Behrman	1983	Income	Nicaragua	0.01
Behrman and Wolfe	1984	Income	Nicaragua	0.05
Bouis and Haddad	1988	Income	Philippines	0.06-0.11
Trairatvorakul	1984	Income	Thailand	0.27-0.33

## Notes:

<sup>a</sup>Food, not total, expenditure is used.<sup>b</sup>Not significant at 5 per cent significance level.



Table 3

## Wu-Hausman tests for exogeneity

Functional form		ln(Calories)	ln(Protein)
Maintained Model	Alternate Model	$\chi^2$	$\chi^2$
Quadratic in lnPCE			
TSLS	OLS	22.33	6.76
Unearned income		[0.00]	[0.03]
TSLS	TSLS	7.76	5.20
Unearned income	ln(total income)	[0.02]	[0.07]
TSLS	OLS	272.97	87.88
ln(total income)		[0.00]	[0.00]
Cubic in lnPCE			
TSLS	OLS	19.07	2.37
Unearned income		[0.00]	[0.50]
TSLS	TSLS	6.21	1.86
Unearned income	ln(total income)	[0.10]	[0.60]
TSLS	OLS	269.33	85.98
ln(total income)		[0.00]	[0.02]
Inverse-quadratic in lnPCE			
TSLS	OLS	36.37	15.23
Unearned income		[0.00]	[0.00]
TSLS	TSLS	6.63	3.29
Unearned income	ln(total income)	[0.04]	[0.19]
TSLS	OLS	276.00	100.13
ln(total income)		[0.00]	[0.00]
Inverse-quadratic in lnPCE & lnlnPCE			
TSLS	OLS	21.81	2.06
Unearned income		[0.00]	[0.56]
TSLS	TSLS	6.16	0.91
Unearned income	ln(total income)	[0.10]	[0.82]
TSLS	OLS	270.29	86.11
ln(total income)		[0.00]	[0.00]

## Notes:

Tests based on first sub-sample of (12,818) urban households.  $\chi^2$  are tests for differences in expenditure coefficients using jackknifed variance-covariance matrices. 40 additional covariates included in each model. Classical p-values in parentheses. According to Schwarz criterion, critical values are 18.9 and 28.4 for the two parameter and three parameter models respectively.

Table 4

**Expenditure elasticities of nutrient demand  
Instrumented ln(per capita expenditure)**

File of PCE	Instruments=ln(household income)				Instruments = unearned income			
	Quadratic Polynom. (1)	Cubic Polynom. (2)	Quadratic Inverse Polynomial (3)	Quad Inv Polynom w/ lnlnPCE (4)	Quadratic Polynom. (1)	Cubic Polynom. (2)	Quadratic Inverse Polynomial (3)	Quad Inv Polynom w/ lnlnPCE (4)
<b>Calorie intake</b>								
05	0.241 [0.020]	0.418 [0.064]	0.320 [0.028]	0.440 [0.075]	0.140 [0.076]	1.405 [5.156]	0.223 [0.150]	3.587 [15.491]
10	0.219 [0.017]	0.331 [0.042]	0.261 [0.020]	0.314 [0.038]	0.129 [0.066]	1.072 [3.774]	0.184 [0.115]	2.292 [9.547]
25	0.180 [0.012]	0.203 [0.014]	0.182 [0.011]	0.167 [0.013]	0.109 [0.050]	0.586 [1.796]	0.131 [0.067]	0.837 [2.981]
50	0.133 [0.007]	0.090 [0.016]	0.113 [0.006]	0.074 [0.022]	0.085 [0.030]	0.162 [0.150]	0.085 [0.029]	-0.013 [0.695]
75	0.081 [0.007]	0.017 [0.022]	0.061 [0.008]	0.034 [0.017]	0.058 [0.012]	-0.106 [0.782]	0.050 [0.012]	-0.277 [1.590]
90	0.029 [0.012]	0.000 [0.014]	0.025 [0.010]	0.033 [0.010]	0.032 [0.021]	-0.157 [0.716]	0.025 [0.025]	-0.138 [0.631]
95	-0.004 [0.015]	0.018 [0.015]	0.008 [0.012]	0.044 [0.021]	0.015 [0.034]	-0.078 [0.185]	0.014 [0.033]	0.062 [0.492]
<b>Protein intake</b>								
05	0.375 [0.021]	0.505 [0.073]	0.452 [0.030]	0.521 [0.081]	0.251 [0.070]	1.528 [5.992]	0.309 [0.144]	3.732 [18.006]
10	0.349 [0.018]	0.431 [0.048]	0.390 [0.023]	0.420 [0.041]	0.239 [0.061]	1.191 [4.385]	0.279 [0.110]	2.423 [11.094]
25	0.303 [0.013]	0.320 [0.017]	0.304 [0.013]	0.296 [0.015]	0.218 [0.046]	0.700 [2.085]	0.235 [0.064]	0.954 [3.461]
50	0.248 [0.008]	0.216 [0.018]	0.227 [0.007]	0.205 [0.024]	0.194 [0.028]	0.272 [0.170]	0.194 [0.028]	0.095 [0.810]
75	0.186 [0.007]	0.139 [0.025]	0.165 [0.008]	0.150 [0.018]	0.166 [0.013]	0.000 [0.910]	0.160 [0.013]	-0.174 [1.849]
90	0.124 [0.011]	0.104 [0.016]	0.121 [0.011]	0.126 [0.011]	0.138 [0.021]	-0.052 [0.831]	0.133 [0.025]	-0.033 [0.731]
95	0.086 [0.015]	0.102 [0.017]	0.100 [0.012]	0.120 [0.022]	0.121 [0.032]	0.027 [0.211]	0.119 [0.033]	0.168 [0.574]

**Notes:**

Estimates based on second sub-sample of 12,898 households. Instruments include 40 exogenous covariates and third order polynomials in income. Jackknifed two stage standard errors in parentheses below elasticities.

Table 2

## Expenditure elasticities of nutrient demand

%iles of PCE	Functional form for lnPCE			
	Quadratic Polynomial (1)	Cubic Polynomial (2)	Quadratic Inverse Polynom (3)	Quadratic Inv w/ lnlnPCE (4)
<b>Calorie intake</b>				
05	0.402 [0.007]	0.437 [0.010]	0.451 [0.010]	0.422 [0.010]
10	0.368 [0.006]	0.387 [0.007]	0.383 [0.008]	0.383 [0.007]
25	0.308 [0.005]	0.308 [0.005]	0.289 [0.005]	0.312 [0.005]
50	0.237 [0.003]	0.223 [0.004]	0.207 [0.003]	0.230 [0.004]
75	0.156 [0.004]	0.141 [0.005]	0.142 [0.004]	0.146 [0.004]
90	0.076 [0.006]	0.074 [0.005]	0.096 [0.005]	0.071 [0.006]
95	0.025 [0.008]	0.040 [0.008]	0.074 [0.005]	0.030 [0.008]
<b>Protein intake</b>				
05	0.499 [0.009]	0.500 [0.013]	0.530 [0.013]	0.487 [0.012]
10	0.463 [0.007]	0.464 [0.009]	0.468 [0.010]	0.467 [0.008]
25	0.401 [0.006]	0.401 [0.006]	0.379 [0.006]	0.414 [0.006]
50	0.326 [0.004]	0.326 [0.005]	0.299 [0.004]	0.335 [0.005]
75	0.242 [0.004]	0.242 [0.006]	0.233 [0.005]	0.239 [0.005]
90	0.158 [0.007]	0.158 [0.007]	0.184 [0.006]	0.146 [0.007]
95	0.105 [0.008]	0.105 [0.010]	0.160 [0.007]	0.092 [0.009]

Notes: Expenditure treated as exogenous. Estimates based on second sub-sample of urban households (12,898 observations). 40 additional covariates included in each regression. Jackknifed standard errors in parentheses below elasticities. According to the Schwarz criterion, the critical value for the t statistic is 3.08

Table 5

Expenditure elasticities of nutrient demand  
Instrument log(per capita expenditure) with log(household income)  
robustness to inclusion of household composition  
in quadratic inverse model

File of PCE	Calorie intake				Protein intake			
	Full sample Demographics		Household composition		Full sample Demographics		Household composition	
	excluded (1)	included (2)	2 adults 1 child (3)	2 adults 2 children (4)	excluded (1)	included (2)	2 adults 1 child (3)	2 adults 2 children (4)
05	0.320 [0.028]	0.379 [0.026]	0.441 [0.116]	0.306 [0.068]	0.452 [0.030]	0.491 [0.029]	0.596 [0.159]	0.452 [0.093]
10	0.261 [0.020]	0.309 [0.019]	0.354 [0.087]	0.245 [0.051]	0.390 [0.023]	0.421 [0.021]	0.504 [0.120]	0.377 [0.069]
25	0.182 [0.011]	0.213 [0.010]	0.238 [0.050]	0.164 [0.030]	0.304 [0.013]	0.324 [0.011]	0.376 [0.070]	0.273 [0.039]
50	0.113 [0.006]	0.131 [0.005]	0.138 [0.025]	0.094 [0.019]	0.227 [0.007]	0.239 [0.006]	0.265 [0.033]	0.184 [0.023]
75	0.061 [0.008]	0.069 [0.007]	0.062 [0.025]	0.041 [0.023]	0.165 [0.008]	0.171 [0.008]	0.177 [0.027]	0.114 [0.028]
90	0.025 [0.010]	0.026 [0.010]	0.011 [0.034]	0.005 [0.029]	0.121 [0.011]	0.121 [0.011]	0.115 [0.040]	0.065 [0.036]
95	0.008 [0.012]	0.006 [0.011]	-0.012 [0.040]	-0.011 [0.032]	0.100 [0.012]	0.098 [0.012]	0.086 [0.047]	0.043 [0.040]

Outlay equivalent ratios

	Calorie intake		Protein intake	
	$\pi$ ratio	index	$\pi$ ratio	index
aged 0 - 4 years				
males	-3.867 [0.161]	18	-2.165 [0.089]	33
females	-3.797 [0.165]	18	-2.234 [0.094]	32
aged 5 - 9 years				
males	-2.267 [0.162]	31	-1.642 [0.102]	43
females	-2.804 [0.159]	25	-1.872 [0.098]	38
aged 10-14 years				
males	-1.098 [0.157]	63	-0.971 [0.101]	73
females	-1.875 [0.153]	37	-1.556 [0.096]	46
aged 15-54 years				
males	-0.695 [0.249]	100	-0.709 [0.161]	100
females	-2.050 [0.332]	34	-1.514 [0.212]	47
aged $\geq$ 55 years				
males	-1.369 [0.195]	51	-1.075 [0.132]	66
females	-2.687 [0.139]	26	-1.743 [0.085]	41

Notes:

Estimates based on second sub-sample. Columns 1 and 2 use all 12,898 observations; column 3 based on sample of 2 adult and 1 child aged <10 years (970 observations); column 4 based on 2 adult 2 children sample (868 observations). Demographic  $\pi$  ratios for full sample including demographics (column 2). Jackknifed standard errors in parentheses below elasticities and  $\pi$  ratios.

Figure 1a : Contour plot of  $\ln(\text{calories})$  against  $\ln\text{PCE}$  : Urban sector

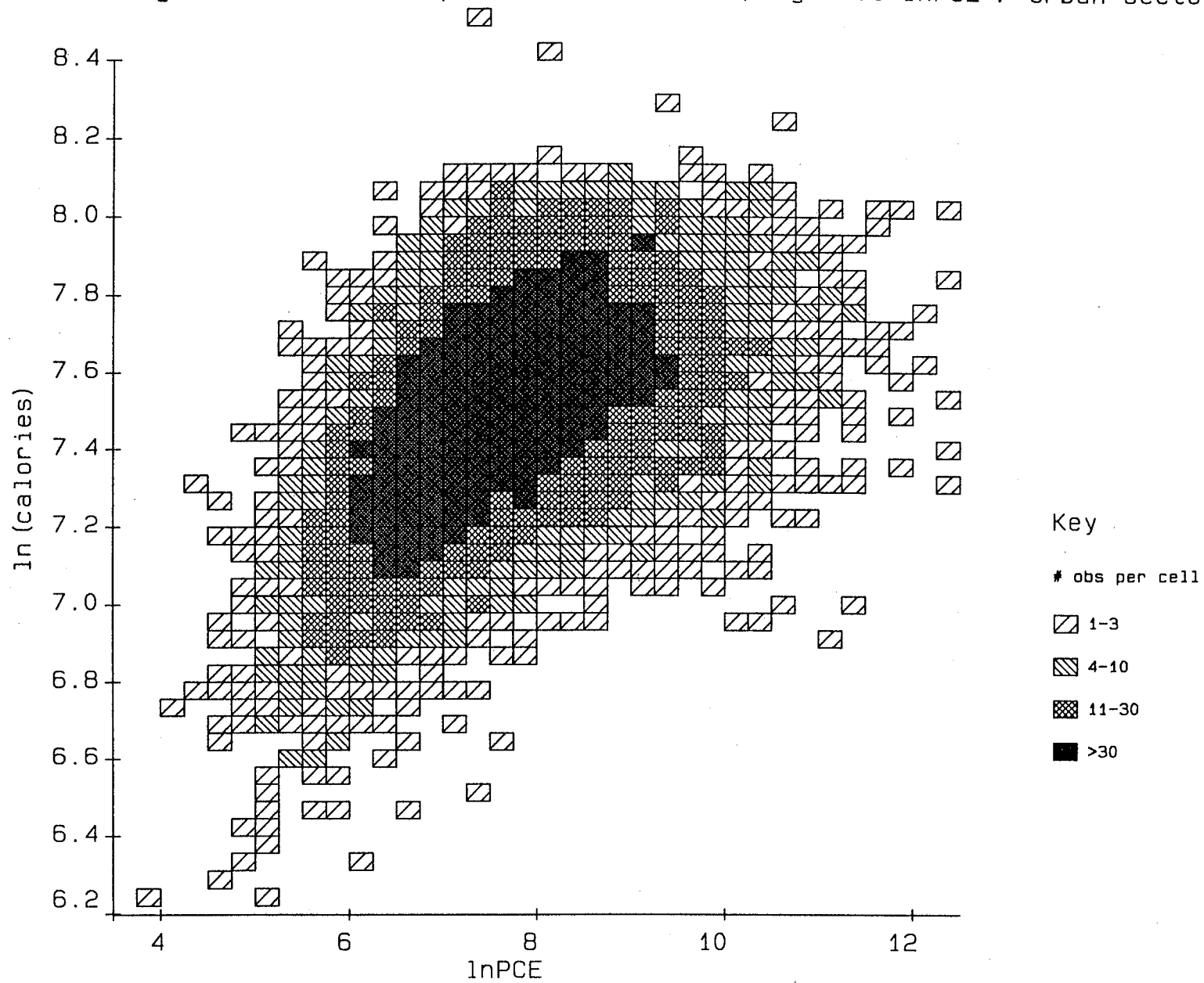
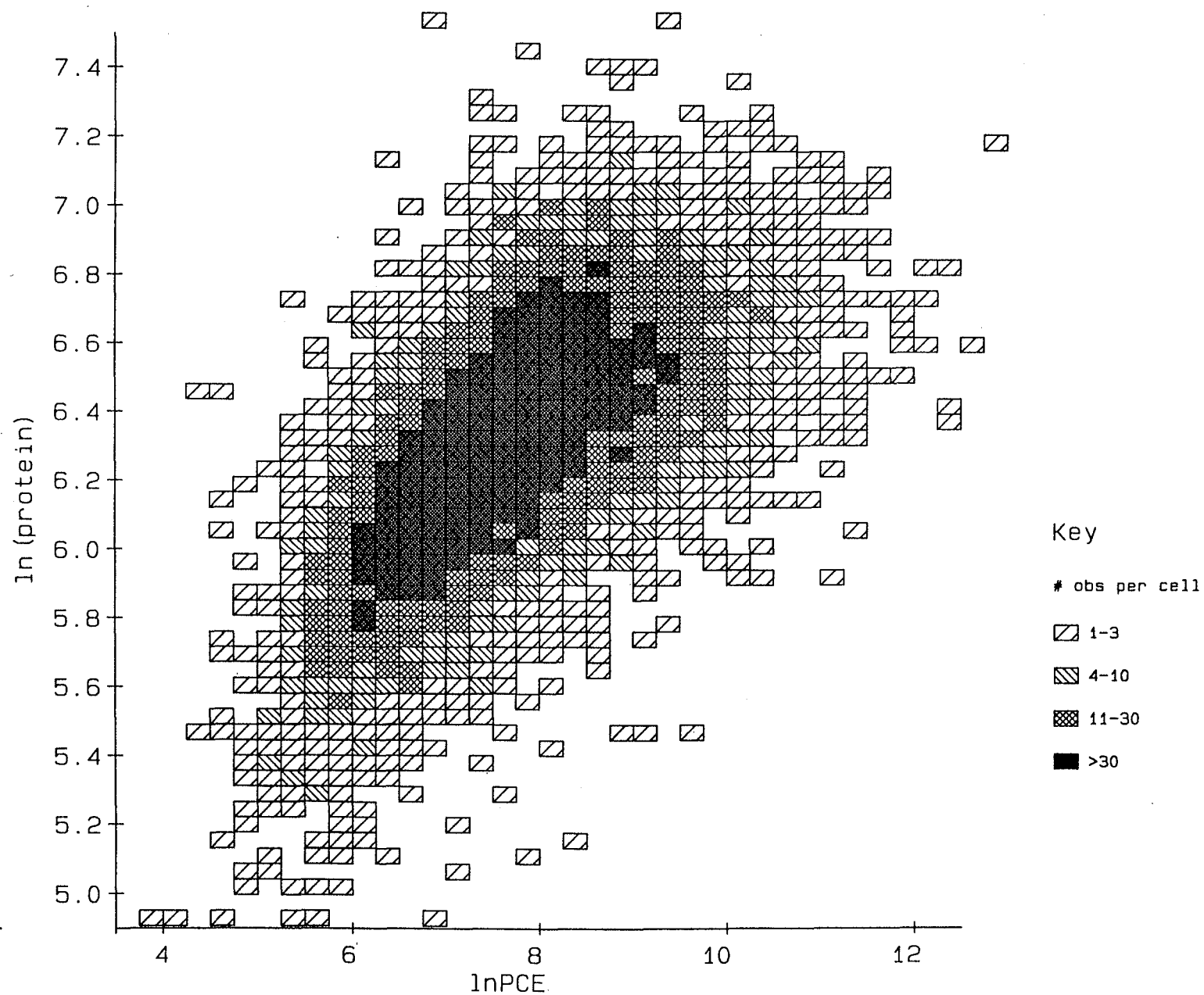


Figure 1b : Contour plot of  $\ln(\text{protein})$  against  $\ln\text{PCE}$  : Urban sector



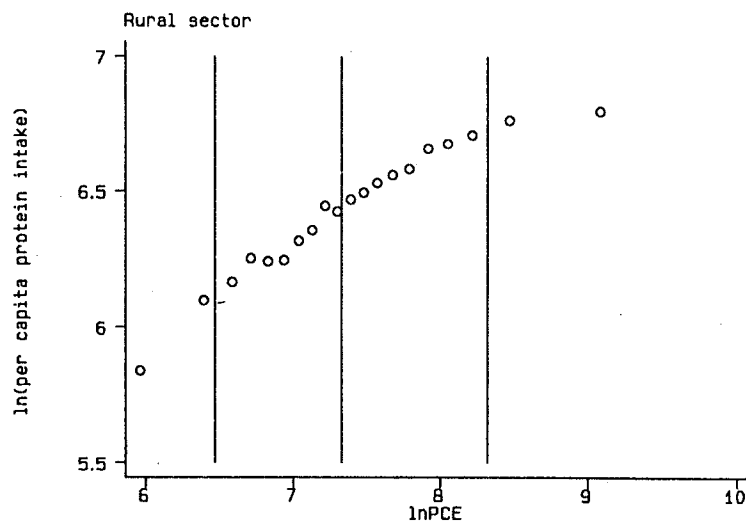
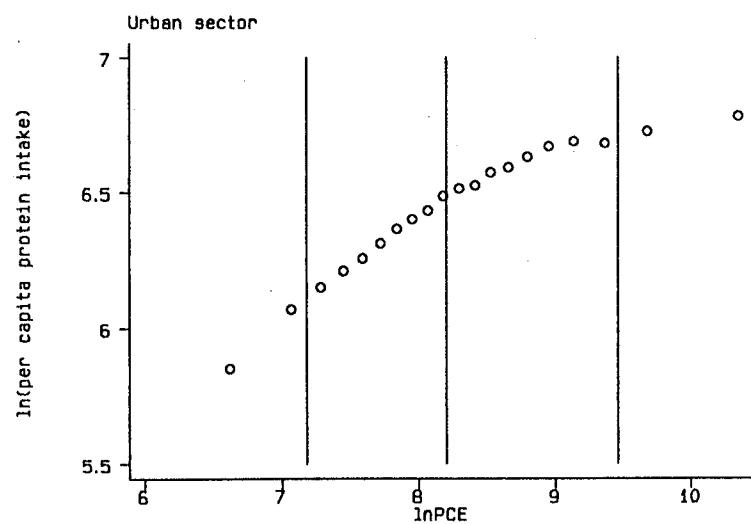
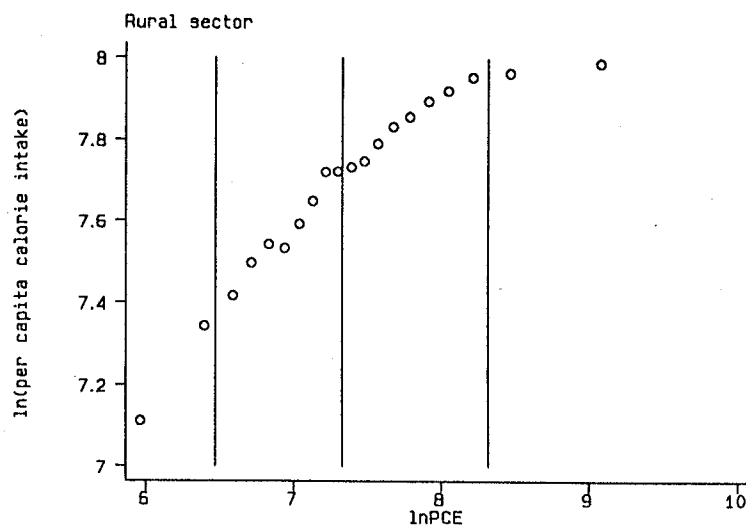
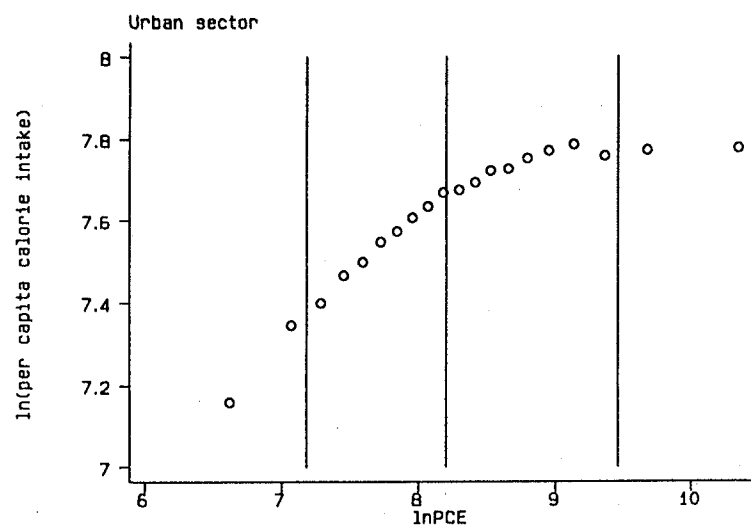


Figure 2 : Mean nutrient intake by icosohiles of PCE

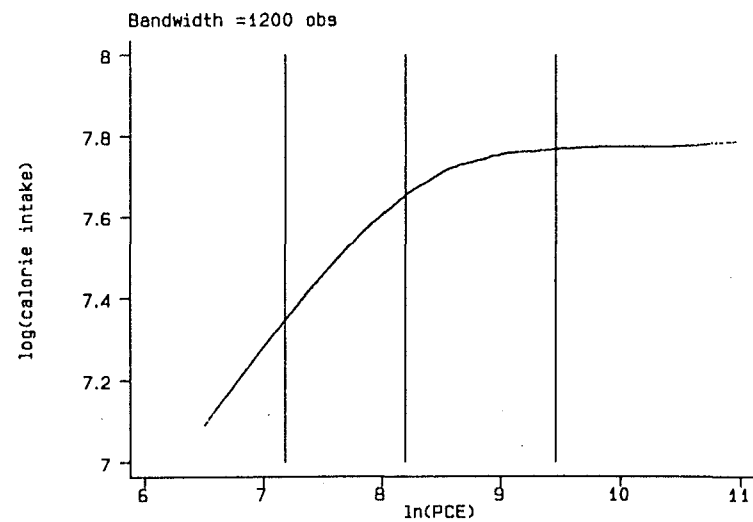
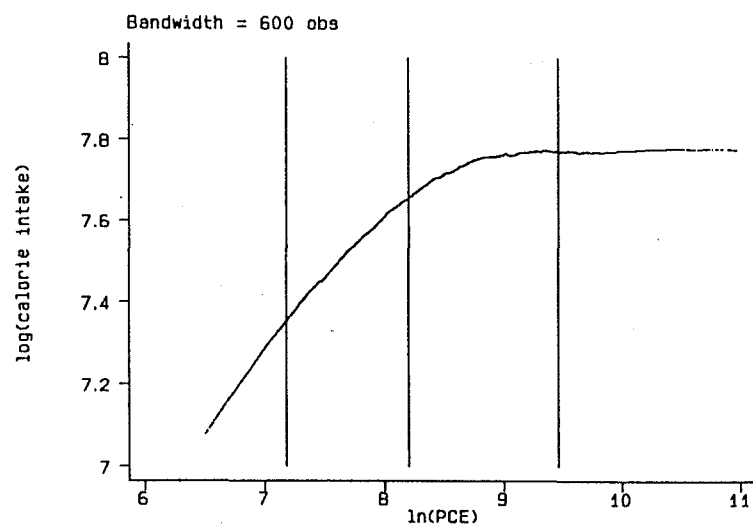
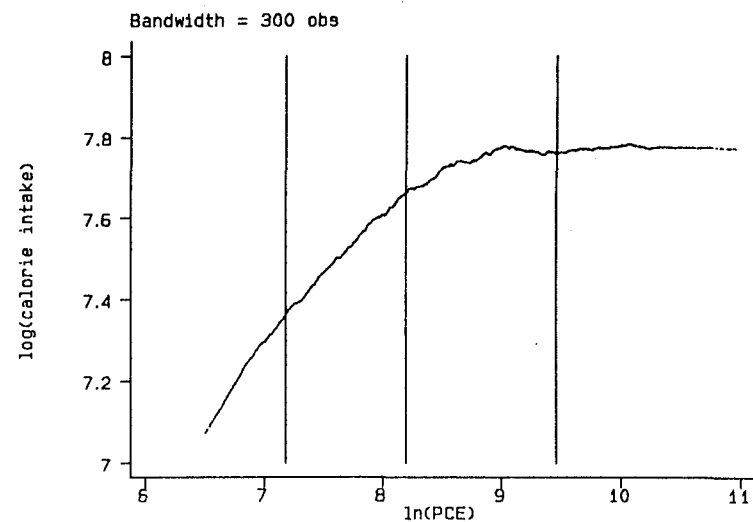
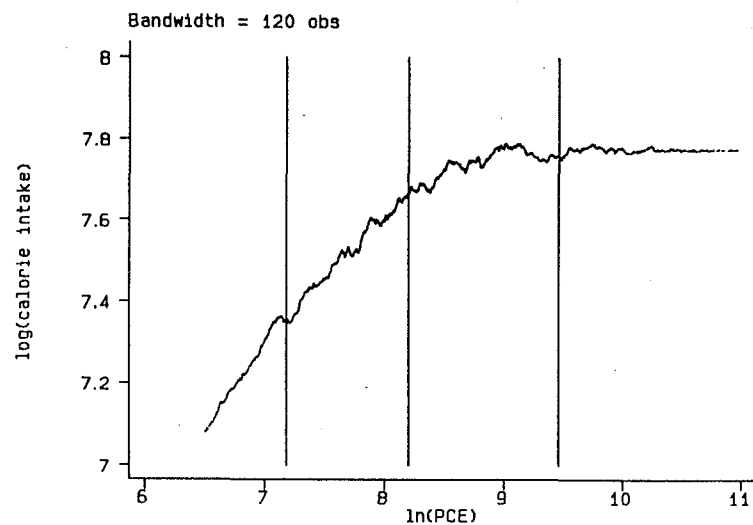


Figure 3a : LOWESS of  $\ln(\text{calorie intake})$  on  $\ln(\text{PCE})$



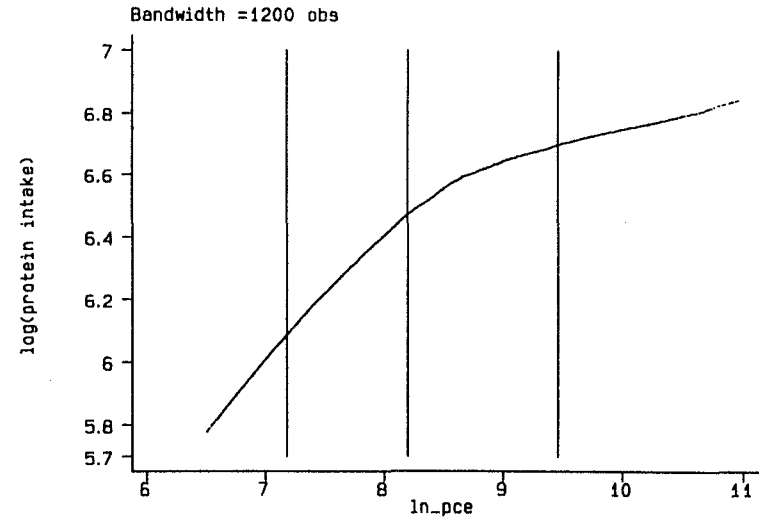
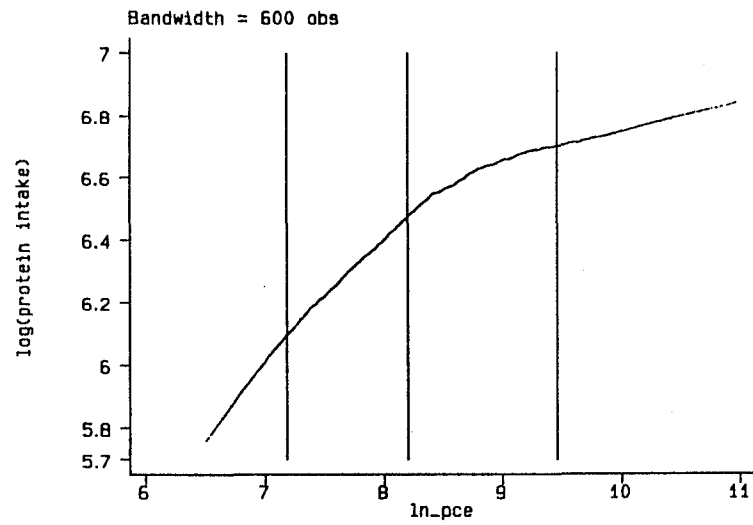
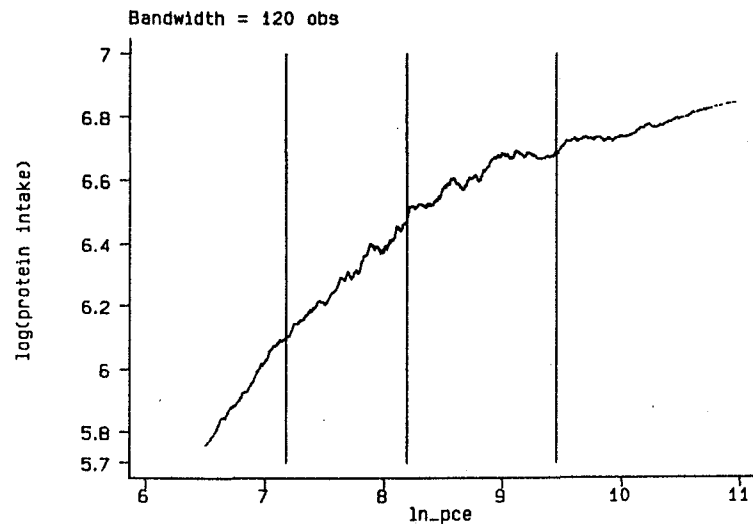
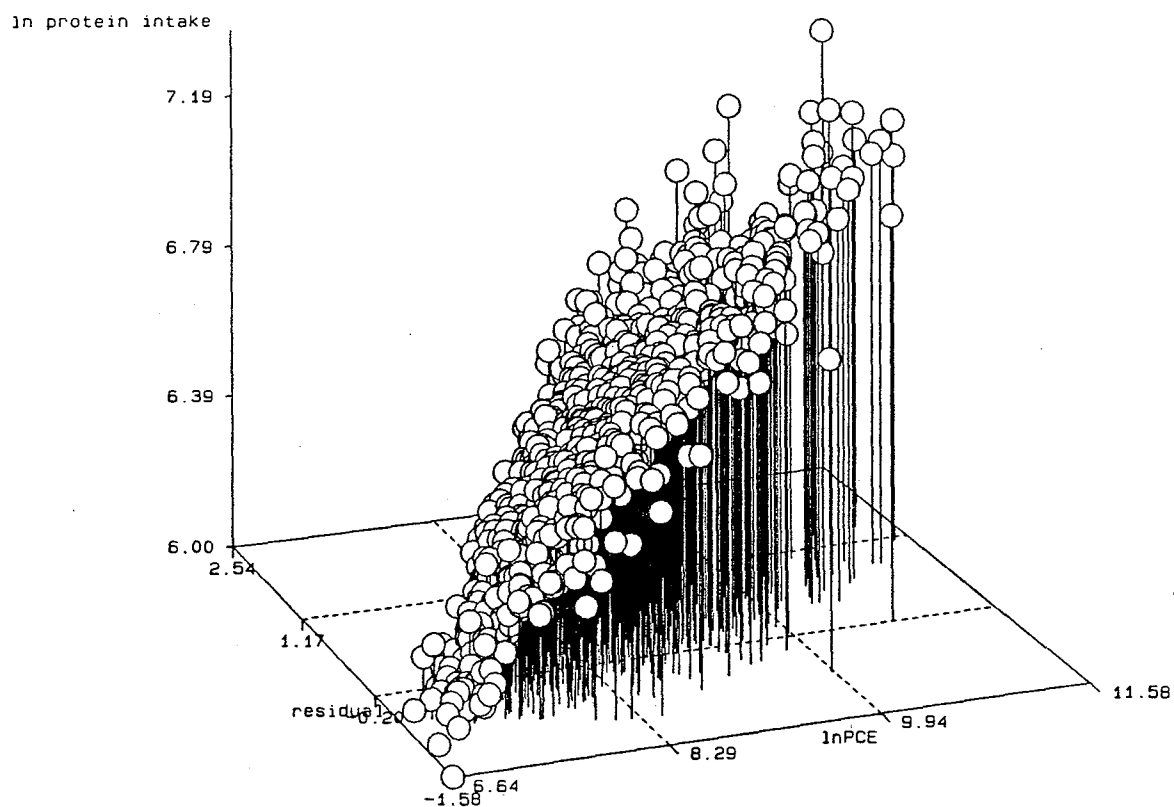
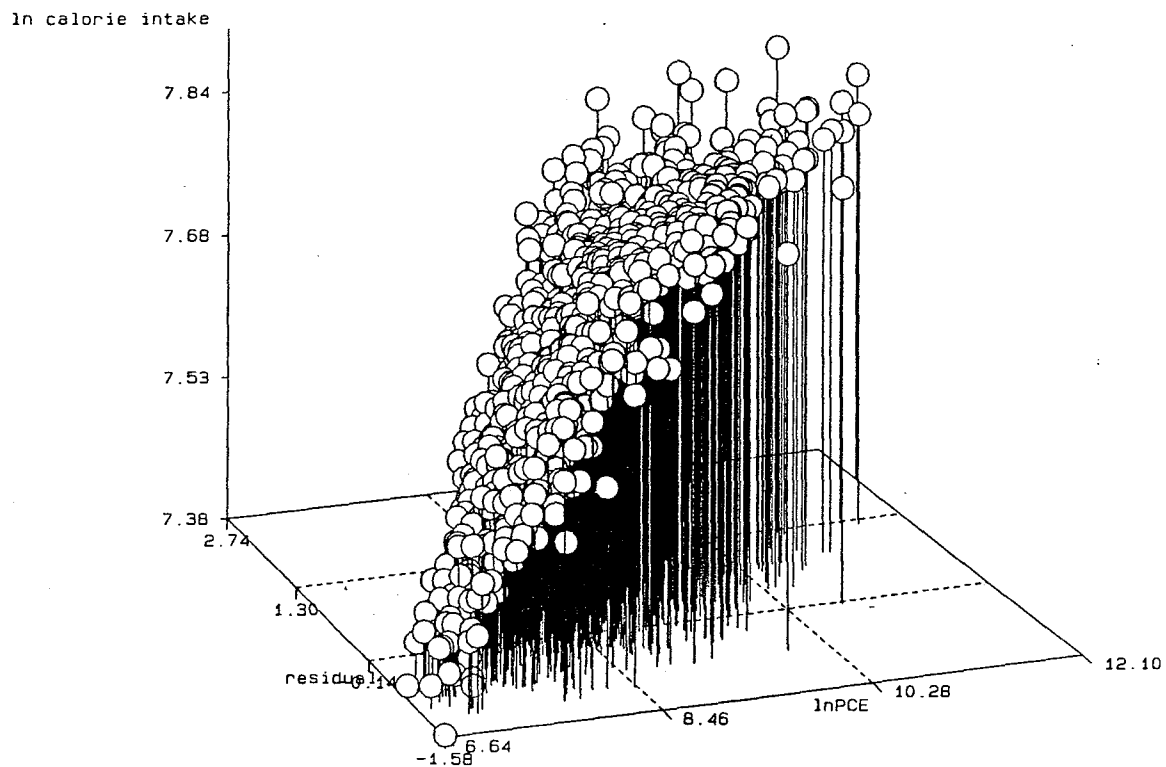


Figure 3b : LOWESS of  $\ln(\text{protein intake})$  on  $\ln(\text{PCE})$

Figure 4 : Non-parametric IV estimates



Bandwidth = 300 observations