The Demand for Food and Calories
Author(s): Shankar Subramanian and Angus Deaton
Reviewed work(s):
Source: *Journal of Political Economy*, Vol. 104, No. 1 (Feb., 1996), pp. 133-162
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/2138962
Accessed: 13/01/2012 19:39

# The Demand for Food and Calories

Shankar Subramanian

*Cornell University and Indira Gandhi Institute of Development Research*

Angus Deaton

*Princeton University*

We investigate nutrition and expenditure in rural Maharashtra in India. We estimate that the elasticity of calorie consumption with respect to total expenditure is 0.3–0.5, a range that is in accord with conventional wisdom. The elasticity declines only slowly with levels of living and is far from the value of zero suggested by a recent revisionist literature. In these Indian data, the calories necessary for a day's activity cost less than 5 percent of the daily wage, which makes it implausible that income is constrained by nutrition rather than the other way around.

## I. Introduction

This paper is concerned with the relationship in poor countries between economic status, as represented by income or by total expenditure, and nutritional status, as represented by calories consumed. In the development literature, there are two different strands of inquiry.

We thank the Director of Economics and Statistics, Maharashtra, and the Chairman and Governing Council of the National Sample Survey Organization for allowing us to use their household-level data from the thirty-eighth round survey. We have received helpful comments from Kaushik Basu and Kirit Parikh and from seminar participants at the Indira Gandhi Institute of Development Research, the Gujarat Institute of Area Planning, the Delhi School of Economics, Princeton, and Northwestern. We are grateful to Harold Alderman, Howarth Bouis, and two referees for detailed and thoughtful comments on the previous version and to Anil Deolalikar and Tom Walker for access to the data from the International Crops Research Institute for the Semi-arid Tropics.

[*Journal of Political Economy*, 1996, vol. 104, no. 1]

The first, on efficiency wages, which begins with Leibenstein (1957) and especially Mirrlees (1975) and Stiglitz (1976), postulates that productivity depends nonlinearly on nutrition. In extreme cases, this relationship predicts the existence of unemployment since those who do not get enough to eat are insufficiently productive to make it profitable for employers to hire them at wages below what is paid to those in work. That this mechanism is part of the explanation for destitution in India has been recently argued by Dasgupta (1993, pt. 4). More broadly, Fogel (1994) sees a nutritionally determined inability to work as a brake on economic growth through most of human history. The second line of inquiry takes nutrition to be conditioned by income and by the demand for food, and the main object of research has been the Engel curve for calories. This is the main topic of this paper, which uses household-level data from the Indian National Sample Survey, although we also produce evidence on whether the link from nutrition to productivity is a plausible cause of poverty in our data.

In the recent literature on calorie Engel curves, there is debate on the extent to which nutrition responds to income. For many years, conventional wisdom has held that hunger and malnutrition would be eliminated by economic growth. The demand for calories will rise with income, if not one for one as consumers substitute quality for quantity, then with an elasticity substantially greater than zero. In contrast, some recent studies—Behrman and Deolalikar (1987), Bouis and Haddad (1992), and Bouis (1994)—have argued that the elasticity is close to zero, so that "increases in income will *not* result in substantial improvements in nutrient intakes" (Behrman and Deolalikar 1987, p. 505). If this position is accepted, there are important implications for the way economists think about development. In accord with some popular beliefs, economic policies that are good for growth do not imply the elimination of hunger. Indeed, even policies that increase the incomes of the poorest may not improve their nutrition. A zero elasticity also creates a chasm between the way economists think about living standards, where real income is the measure of welfare, and people know what is good for them, and the way living standards are often characterized by nutritionists and development practitioners, who see development largely in terms of guaranteeing that people have enough to eat. Trade policy, pricing and tax policy, and project evaluation would all be done differently in a world in which the promotion of real incomes did not meet the objectives of development. To take but one example, economists think of substitution possibilities as welfare enhancing; if it is possible to substitute across a wide range of foods, consumers are well protected against changes in relative prices. To the nutritionist concerned only with adequate diet, welfare is *decreased* by voluntary substitution away from approved to disapproved food.

In this paper, and like Behrman and Deolalikar, we look again at the evidence from India, but using data on rural households from the Maharashtran state sample of the thirty-eighth round of the National Sample Survey carried out in 1983. Since one of Behrman and Deolalikar's three villages is in Maharashtra, the comparison with their results is close enough to be instructive. Although the National Sample Survey data are not a panel, so that we are not able to control for individual heterogeneity by estimating differenced regressions, our sample has compensating advantages. We have 5,630 households, so that it is possible to estimate the calorie elasticity more precisely than is possible with the sample of 253 households collected by the International Crops Research Institute for the Semi-arid Tropics (ICRISAT). Precision is important here because Behrman and Deolalikar's preferred point estimate of the elasticity, although insignificantly different from zero, is 0.37, which is close to the conventional wisdom (cf. the influential study by Reutlinger and Selowsky [1976]). Our sample is also large enough for us to follow the example of Strauss and Thomas (1990) and use nonparametric estimation to explore the shape of the relationship between calories and total outlay. Strauss and Thomas found that for Brazil, the elasticity of calories with respect to total expenditure is higher for poorer consumers, and as argued by Ravallion (1990), the result is plausible in general.

We also look at the cost of calories and at how patterns of demand and nutrition change with total expenditure. Some goods contain many calories and are cheap, some contain few and are expensive, and it is the substitution of the latter for the former, to the extent that it occurs, that drives a wedge between the elasticities of food expenditure and of calories. The National Sample Survey collects data on 149 individual foods, and we use this detail to document how consumers move from cheap calories to expensive calories, what foods are involved, and whether the change takes place between broad groups of goods or within them. Counting the cost of calories is also relevant for evaluating explanations of unemployment based on efficiency wages. Indeed, the Maharashtran data make it quite implausible that malnutrition is the explanation for poverty, as opposed to its consequence. Food energy is cheap in rural Maharashtra, and the additional 600 or so kilocalories that a farmer in the tropics might require for daily physical activities (Dasgupta 1993, p. 422) can be purchased in the form of the customary diet (coarse cereals) for about 4 percent of the daily wage. If nutrition is a trap, it is one from which there is a ready escape.

Our estimates of the total expenditure elasticity of calories are within the range found by Behrman and Deolalikar (as would be almost any sensible estimate). For the poorest households, the elasticity is around 0.55 and falls fairly gently through the cross section to perhaps 0.40 for the better-off households. Those at the top of the distribution pay almost twice as much for their calories, with much of the switch accounted for by substitution out of cheap, coarse grains. Except for very poor households, where there is evidence of quality upgrading even within coarse grains, the price per calorie rises much less within broad groups of foods than between them. As a result, calorie elasticities would not be much overestimated if total calories were calculated by applying conversion factors to broad groups, provided that the groups are appropriately defined. The total expenditure elasticity of food is around 0.75, a figure that is more or less equally divided between the elasticity of calories and the elasticity of the price of calories, with the latter largely set by the familiar switch from cereals to other foods as living standards rise. Our results are therefore very much in the traditional camp and provide no support for the notion that nutrition will not increase with higher standards of living, that the calorie elasticity is zero. We recognize the (very real) possibility that our estimates of the elasticity are biased upward by common measurement errors in calories and total expenditure, but we show that these effects cannot explain our results if the true elasticity is close to zero.

The remainder of the paper is organized as follows. Section II discusses the data, and Section III presents some preliminary tabulations that show the main features of expenditure and nutrition patterns. Section IV adopts a somewhat more sophisticated approach and provides parametric and nonparametric estimates of calorie elasticities, with controls for demographic structure and other factors. We also provide further analysis of quality shifts in consumer purchases and the increases in prices paid per calorie. Section V places our results within the broader literature, discusses a number of caveats, and assesses the extent to which the results can be generalized.

## II. Data Description

The data used here are taken from the thirty-eighth round (1983) of the National Sample Survey (NSS) for rural households in Maharashtra state in (south) western India. There are two NSS surveys, one collected by the National Sample Survey Organization itself and the other, which is the one used here, collected by the state-level directorate of statistics and known as the state sample. The rural part of this sample consists of 5,630 households, 10 in each of 563 villages, and was collected in four subrounds over a single calendar year. Respondents are asked to recall how much they had consumed of each of more than 300 items over the last 30 days and to report expenditures in rupees as well as physical quantities when appropriate. There are 149 food items in the list. Data were also recorded on the *number* of meals given to employees, the number given to guests at ceremonies, and the number given to guests at other occasions, as well as the number of meals received as payment in kind, the number of meals purchased, and the number of meals eaten at home. A "meal" is a person-meal, so that if the headman gives a feast to 300 guests at his daughter's wedding, 300 meals would be recorded. Consumption from own production and consumption from receipts in kind were valued at prices then prevailing locally. Other information collected included data on landownership; the age, sex, marital status, and educational level of family members; the occupation of the household head; the household's religion and caste; and indicators of wealth, such as house area, type of construction, state of repair, and access to piped drinking water and electricity. In recent years, the NSS has not attempted to collect income data as part of its consumption surveys since it is widely believed that attempts to do so would compromise the response rate and the quality of the data. As a result, we follow much of the previous literature in studying the relationship between nutrition and total expenditure or, for short, total outlay.

We computed calorie intakes from the basic data using the calorie content tables in Gopalan, Sastri, and Balasubramanian (1971). This is mostly a straightforward if tedious exercise. The full detail of reported food consumption is used, with weights converted to calories using the calorie content factors; the result is, of course, calorie "availability" rather than intake, a distinction that could be important if the income elasticity of "wastage" is high enough to compromise our results. For a number of items, such as meals taken at restaurants, we used the average "price" of calories from the rest of the data plus a 50 percent premium to reflect processing margins. We also made an attempt to remove those calories not actually consumed by members of the household and to add in those not reported in food purchases. A substantial amount of food is supplied by better-off Indian households to guests and to servants, and many poor and not so poor individuals receive meals as part of their employment. The result is that for rich households, total calorie availability will overstate calorie consumption by household members, with a corresponding understatement for poorer households, so that there is a real possibility of overstating the elasticity of calorie consumption to income or total expenditure.

The correction is made as follows: total calories were regressed on the number of meals given to guests at ceremonies, $m_1$, the number of meals given to guests on other occasions, $m_2$, the number of meals given to employees, $m_3$, and the number of meals taken at home, $m_4$. This gives

$$\text{cals} = \underset{(1.9)}{1{,}550 m_1} + \underset{(2.9)}{1{,}379 m_2} + \underset{(7.7)}{607.8 m_3} + \underset{(.1)}{726.9 m_4}, \quad R^2 = .897 \qquad (1)$$

(standard errors are reported in parentheses), so that meals given to guests contain about twice as many calories as those normally eaten at home or given to servants and employees. Note that 600 calories is about a third of mean daily intake (see table 1 below). We use these estimates to adjust available calories by subtracting from the total $m_1$ through $m_3$ multiplied by their respective coefficients, and then adding the total meals received multiplied by the weighted average of the second and third coefficients in (1), with weights proportional to the total number of meals given to guests and employees, respectively. The adjustment generates a *negative* calorie intake for six households, which report that they gave many people meals but spent very little on food. These households are eliminated from the analysis below. Note that we have no way of allocating these adjustments to the various food groups: we do not know what the composition is of these various meals. In consequence, when we look at total calories, we typically work with the adjusted figure. However, calorie consumption and prices per calorie for food groups are not adjusted.

Total expenditure (outlay) is taken net of purchases of durables; exclusion of these lumpy purchases is a standard procedure to help make consumption a less noisy measure of well-being. In contrast to *calories*, we make no correction to *expenditures* for meals provided to guests or employees. To the extent that employees are household servants, the expenditures presumably contribute to the living standards of members of the household and are appropriately included in total outlay. By contrast, expenditures on meals for farmworkers should be excluded, and our failure to do so, since we have excluded the calorific values of the meals, will tend to bias *downward* our estimate of the elasticity of calories to total expenditure.

In this paper we confine ourselves to the rural sample, partly because nutritional issues are often thought to be more important for rural people and partly for comparison with earlier results, particularly those of Behrman and Deolalikar. However, it is also true that urban households obtain much more of their food in the form of precooked meals. It is difficult to make an accurate assessment of the calories from that source and impossible to allocate the expenditures and calories to particular foods.

## III. Tabulations: Expenditures and Calories

Table 1 shows how consumers allocate their budgets, from which foods they get their calories, and how much each calorie costs if purchased via each of the various foods. We look both at the disposition over broad categories of expenditures, in panel A, and at the allocation within the important cereals category, in panel B. Columns 1–3 show the expenditure patterns, expressed as shares of the budget. They are calculated from the budget shares of each of the 5,630 households in the survey, averaged over the whole sample in column 1 and then over the bottom and top deciles of per capita household expenditure in columns 2 and 3, respectively. The average of per capita household expenditure is 115 rupees per person per month; the corresponding means for the top and bottom deciles are 48 rupees and 282 rupees, respectively. Columns 4–6 show the distribution of calories over the various goods. As shown in the bottom row, per capita daily calories are 2,120 on average and 1,385 and 3,382 in the two extreme deciles. These are the unadjusted figures; when the adjustments for servants and guests are made, the average falls somewhat. More important, the calorie availability for poor people rises, by 46 calories, and that for the highest decile falls, by 215 calories. Columns 7–9 show how many rupees of expenditure on each good were required to generate 1,000 calories. On average, rural Maharashtran households in 1983 spent 1.14 rupees per 1,000 calories, with the poorest decile paying substantially less, 88 paise, and the best-off substantially more, 1 rupee and 50 paise. For reference, the rural daily wage in 1983 was about 15 rupees.

Cereals, particularly coarse cereals, provide cheap calories, and so they bulk much larger in the calorie share (71 percent) than in the expenditure share (41 percent). At the other extreme, meat calories cost 12 rupees per thousand and account for less than 1 percent of total calories and 5 percent of the budget. Processed foods and beverages produce even fewer calories per unit of expenditure. It is these differences, together with the fact that the budget turns away from cereals, especially coarse cereals, that account for the increase in the price per calorie between the least and the most well off. Rice and jowar together account for more than half of calories on average and closer to two-thirds for the poorest group. For neither of these groups does the price per calorie change very much from the bottom to the top decile. Nor does the substitution toward wheat that accompanies rising incomes do anything to increase the overall price of calories; wheat provides calories more cheaply than rice, and its quality does not rise with income. What accounts for the increase in the price of calories with increasing living standards is not within-group substitu-

TABLE 1

EXPENDITURE PATTERNS, CALORIE CONSUMPTION, AND PRICES PER CALORIE, RURAL MAHARASHTRA, 1983

| | EXPENDITURE SHARES (%) | | | CALORIE SHARES (%) | | | PRICE PER CALORIE (Rupees per 1,000 Calories) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean (1) | Bottom 10% (2) | Top 10% (3) | Mean (4) | Bottom 10% (5) | Top 10% (6) | Mean (7) | Bottom 10% (8) | Top 10% (9) |
| | | | | A. Food Groups | | | | | |
| Cereals | 40.7 | 46.0 | 31.0 | 70.8 | 77.3 | 57.3 | .64 | .51 | .79 |
| Pulses | 8.9 | 10.2 | 7.8 | 6.6 | 6.2 | 7.2 | 1.51 | 1.44 | 1.60 |
| Dairy | 8.1 | 4.9 | 11.8 | 2.8 | 1.3 | 4.9 | 3.69 | 3.59 | 3.92 |
| Oils and fats | 9.0 | 9.2 | 9.2 | 5.9 | 4.8 | 7.6 | 1.74 | 1.67 | 1.81 |
| Meat | 5.1 | 3.4 | 6.4 | .7 | .4 | 1.0 | 11.7 | 11.0 | 12.2 |
| Fruits and vegetables | 10.5 | 8.5 | 12.0 | 3.5 | 2.3 | 5.4 | 3.90 | 3.83 | 3.85 |
| Sugar | 6.5 | 7.4 | 5.9 | 7.2 | 7.0 | 8.0 | 1.01 | .94 | 1.09 |
| Other food | 11.3 | 10.4 | 16.1 | 2.5 | 0.8 | 8.6 | 17.4 | 16.8 | 15.9 |
| | | | | B. Cereals | | | | | |
| Rice | 11.6 | 9.0 | 10.9 | 15.2 | 10.1 | 16.5 | .95 | .89 | 1.02 |
| Wheat | 5.6 | 3.8 | 7.9 | 8.5 | 4.7 | 14.4 | .79 | .73 | .82 |
| Jowar | 18.2 | 27.4 | 9.3 | 37.8 | 52.9 | 21.6 | .50 | .43 | .55 |
| Bajra | 3.0 | 2.7 | 1.3 | 6.6 | 4.9 | 3.2 | .48 | .48 | .50 |
| Other coarse cereal | 1.2 | 2.8 | .3 | 2.2 | 4.5 | .6 | .66 | .58 | .99 |
| Cereal substitutes | 1.1 | .5 | 1.3 | .6 | .2 | .8 | 2.23 | 2.22 | 2.22 |
| Total food (or total calories) | 67.4 | 73.4 | 54.1 | 2,120<br>2,098 | 1,385<br>1,429 | 3,382<br>3,167 | 1.14 | .88 | 1.50 |

NOTE.—Mean refers to mean over the whole sample, bottom 10% to mean over households in the bottom decile of per capita household expenditure, and top 10% to mean over households in the top decile of per capita household expenditure. The figures in the last row of panel B are unadjusted and adjusted total calories, respectively, where the adjustment corrects for meals given to others or not received from others; see the text for a full description. Shares of calories and of expenditures are calculated on an individual household basis and are averaged over all appropriate households. Calorie prices are averages over consuming households.

tion, but the substitution between groups, away from cereals and toward dairy products, edible oils, processed foods, and beverages.

It is useful to have some simple formulas that serve to characterize these effects and also to provide a link with Behrman and Deolalikar's work. We use capital letter subscripts to denote groups of goods (typical group $G$) and lowercase letters to denote goods within groups (typical good $i$ in group $G$). Any one of the composite of goods in the table is to be regarded as a group, and the individual goods, which we do not observe, are taken to be quite homogeneous, so that a kilogram of good $i$ has a constant calorie availability $k_i$ no matter who buys it. Total calories $c$ are given by

$$c = \sum_G \left( \sum_{i \in G} q_{Gi} k_{Gi} \right), \tag{2}$$

where $q_{Gi}$ is the quantity of the good. The elasticity of calories with respect to total outlay $x$, $\epsilon_{cx}$, is then easily shown to be

$$\epsilon_{cx} = \sum_G \eta_G \sigma_G \left( \sum_{i \in G} \eta_{iG} \omega_{iG}^c \right), \tag{3}$$

where $\sigma_G$ is the share of calories obtained from group $G$, $\eta_G$ is the total expenditure elasticity of expenditure on group $G$, $\omega_{iG}^c$ is the share of group $G$'s calories that come from good $i$, and $\eta_{iG}$ is the elasticity of expenditure on good $i$ with respect to total expenditure on the group. Note also that the prices per calorie are the ratios of group expenditure $x_G$ to group calories $c_G$, so that the elasticity of group $G$'s price per calorie with respect to total expenditure is given by

$$\xi_G = \frac{\partial \ln(x_G/c_G)}{\partial \ln x} = \eta_G \left( 1 - \sum_{i \in G} \eta_{iG} \omega_{iG}^c \right). \tag{4}$$

If we substitute this equation in (3), we can express the calorie elasticity as

$$\epsilon_{cx} = \sum_G \sigma_G (\eta_G - \xi_G). \tag{5}$$

Equation (5) plays the same role as equation (7) of Behrman and Deolalikar. To the extent that consumers, as they grow richer, substitute expensive nonnutrient characteristics for nutrients *within* the group, $\xi_G$ will be large, and the overall outlay elasticity of calories will be reduced. As Behrman and Deolalikar emphasize, counting calories by applying conversion factors to broad groups effectively assumes that each $\xi_G$ is zero, so that if the assumption is false, there could be substantial overestimation of the elasticity. Note, however, two important points. First, there is no theoretical requirement that $\xi_G$ be positive, and indeed, in table 1, we see that households in the top decile of per capita expenditure (PCE) pay less for calories from "other food" than households in the bottom decile of PCE. If such a phenomenon were widespread, the broad group procedure could actually bias elasticities downward. Second, even if there were no substitution within groups, so that all the $\xi_G$'s were zero, there is still scope for substitution between groups. As table 1 shows, the price for calories over all goods can rise with total outlay, implying a substantial difference between the elasticities with respect to total expenditure of food and calories, respectively. If $\epsilon_{xx}$ is the elasticity of demand for food and $w_G$ is the share of the budget devoted to group $G$, $\epsilon_{xx} = \Sigma\, w_G \eta_G$, so that from (5)

$$\epsilon_{xx} = \epsilon_{cx} + \sum_G \eta_G(w_G - \sigma_G + \xi_G \sigma_G \eta_G^{-1}). \tag{6}$$

If it is generally the case that there is a positive relationship across goods between, on the one hand, total expenditure elasticities and, on the other, the difference between the budget share and the calorie share, then the food elasticity will be greater than the calorie elasticity, even if the within-group elasticities are zero.

A preliminary assessment of the size of these various factors can be made from table 1 or, more simply, by running crude double logarithmic regressions between the logarithms of group expenditures and of group calories on the logarithm of per capita household expenditure. We shall do better than this in the next section, but the results are not misleading, and they allow us to describe the broad features of the data, to quantify the difference between the food and calorie elasticities, and to decompose the rise in the price of calories into its within- and between-group substitution components. We use the full breakdown of goods in table 1, with cereals subdivided, so that there are 13 groups in all. The regressions include the logarithm of total household size as well as the logarithm of per capita total household expenditure, and the calorie shares in equation (5) are those at the mean, taken from column 4 of table 1. The elasticity of total (unadjusted) calorie availability with respect to per capita total expenditure is 0.451, a figure that can be obtained directly or from substitution of the detailed elasticities into the right-hand side of equation (5). If equation (5) is recalculated, setting each $\xi_G$ to zero, we obtain 0.490. The total outlay elasticity of food expenditure is 0.772, so that although the elasticity of the price of calories (0.320) drives a very substantial wedge between the food and calorie elasticities, almost all of the effect comes through the substitution between broad groups, with very little due to substitution within groups. The estimates of $\xi_G$ are mostly small. Only for jowar (0.15), other coarse cereals (0.27), meat (0.15), and fruit (0.43) are the estimates greater than 10 percent; for the admittedly heterogeneous groups of fruits and vegetables (−0.05) and other foods (−0.17), the estimates are negative.

We shall refine these estimates in the next section. However, the broad substantive result remains true: as households become richer, they substitute between broad groups of goods and away from those cereals, particularly coarse cereals, that provide cheap calories. As a result, the average price per calorie rises. There is relatively little substitution within groups, at least if these groups are sensibly defined in a way that is sensitive to the characteristics of the foods involved. From the methodological perspective, it follows that it is not likely to be seriously misleading to calculate calorie availability by applying calorie conversion factors to broad groups of goods.

## IV. Estimation of the Calorie Expenditure Relationship

In this section, we look more closely at the relationship between calories and total expenditure. One of our main concerns is possible nonlinearity. It is quite plausible that poor people, whose income is insufficient to buy sufficient food, should have an elasticity of calories to total expenditure that is much higher than that for those who have enough to eat. Indeed, Strauss and Thomas (1990) found such a relationship for Brazil, with an elasticity of 0.26 for the lowest decile that fell to 0.03 for the highest decile. To allow for such a possibility, we use nonparametric procedures to estimate the regression functions. We begin with the expectation of calories conditional on household per capita expenditure. As before, we work with logarithms and write

$$m(x) = E(y|x) \tag{7}$$

for the regression function, where $y$ is the logarithm of per capita (adjusted) calorie availability, and $x$ is the logarithm of per capita total household expenditure, excluding durables.

We estimate $m(x)$ using a smooth local regression technique whose superiority over kernel and other methods has been demonstrated by Fan (1992, 1993). The procedure works as follows. At any given point $x$, we run a weighted linear regression of the logarithm of calories per head on the logarithm of per capita expenditure. The weights are chosen to be largest for sample points close to $x$ and to diminish with distance from $x$; they are also set so that, as the sample size increases, the weight given to the immediate neighborhood of $x$ is increased so that, in the limit, only $x$ is represented. In our case, for the local regression at $x$, observation $i$ gets the (quartic kernel) weight

$$w_i(x) = \frac{15}{16}\left[1 - \left(\frac{x - x_i}{h}\right)^2\right]^2 \tag{8}$$

if $-h \leq x - x_i \leq h$ and zero otherwise. The quantity $h$ is a bandwidth that is set (in this case by inspecting alternative plots) so as to trade off bias and variance, and that tends to zero with the sample size. The procedure is similar to LOWESS (Cleveland 1979) but is somewhat simpler to implement since it does not require the identification of nearest neighbors. Our main concern here is to plot the regression function so that, instead of calculating local regressions for each point in the sample, we choose an evenly spaced grid of 100 points in the distribution of log per capita expenditure and calculate local regressions for each. The estimate of $m(x)$ is the predicted value from the local regression at $x$, and the local estimated slope coefficient, $\hat{\beta}(x)$ say, is used as an estimate of the slope $m'(x)$.

Standard errors for the regression function and its slope are obtained by bootstrapping following the directions in Efron and Tibshirani (1993, chap. 7.3) with modifications to allow for the clustered structure of the NSS data. The rural sample of 5,630 households comes from 563 sample villages with 10 households drawn from each. Since there is likely to be some intracluster homogeneity of calorie consumption conditional on per capita expenditure, the effective sample size is less than 5,630. Our bootstrapping preserves this feature of the sampling design by two-stage resampling. At the first stage, we resample with replacement from the 563 clusters, and at the second stage, we resample 10 households from each of the selected clusters, again with replacement. We generate 50 such two-stage bootstrap samples, reestimate the local regressions, and use the standard errors over the replications to estimate the standard errors of the original regression.

Figure 1 shows a contour map of the estimated joint density of our two main variables, the logarithm of (adjusted) per capita calories and the logarithm of household expenditure per capita. This figure corresponds to a scatter diagram, but in order to better see the shape of the underlying mass, we have estimated the joint density using a bivariate quartic kernel according to the methods in Silverman (1986, chap. 4.2) and graphed the contour lines. The shape of this figure anticipates a good many of the results to come. Although the contours are not the perfect ellipses that would be generated by a joint normal distribution, they are close to being so, and the regression function

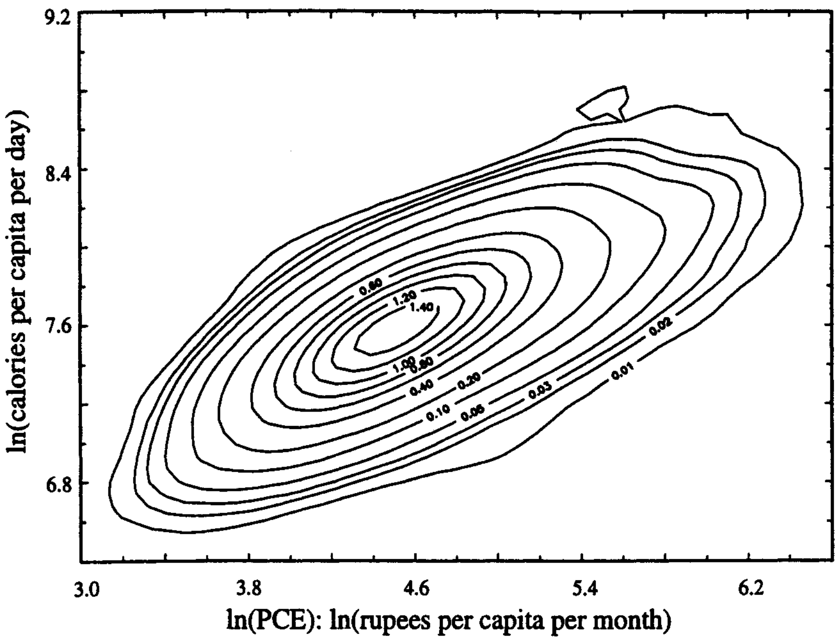


FIG. 1.—Estimated joint density of the logarithms of per capita calories and expenditures.

of log per capita calories on log per capita expenditure is close to being linear. The logarithmic transformation is almost enough to remove both the skewness in the distributions and the nonlinearity in the relationship.

Figure 2 shows the local regression estimate of the regression function corresponding to the joint density in figure 1; the bandwidth is 0.5. The confidence intervals around the line show two standard errors on either side, for both the clustered and unclustered bootstrap. As is to be expected, the confidence bands that allow for clustering are wider, but the design effect is not very large and the regression line is tightly estimated, especially in the middle of the range of per capita expenditure. Although the slope of the line is steeper at lower levels of PCE, the estimated "curve" is very close to being a straight line. Nor is this an effect of oversmoothing; the matching graphs with lower bandwidths are similar in general shape, although there is more variance in detail. There is certainly nothing here that matches Strauss and Thomas's findings for Brazil, perhaps because these rural Indian households are much poorer, and even those at the top of the distribution are not at the point at which calories are unimportant. Figure 3 shows the slopes of the curve in figure 2, again with confi-

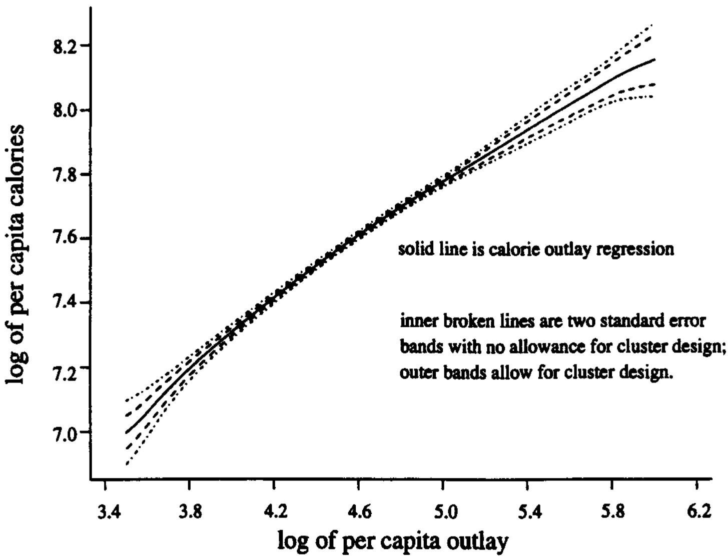


FIG. 2.—Regression function for log calories and log per capita expenditure, Maharashtra, India, 1983.

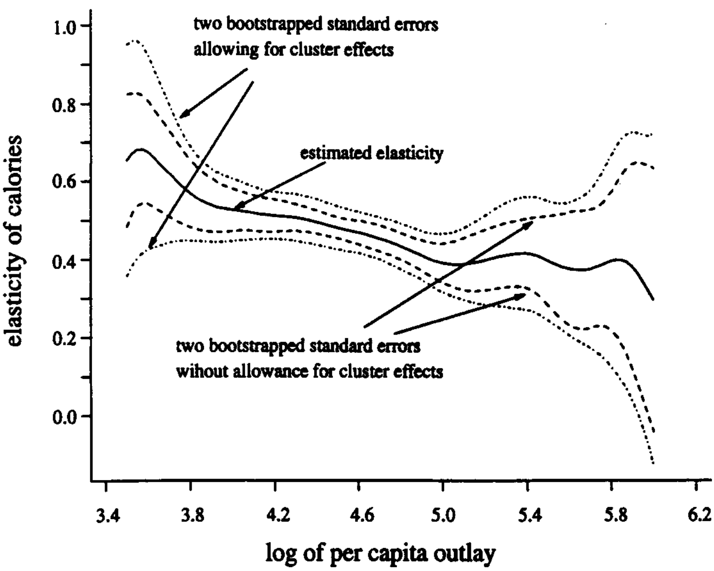


FIG. 3.—Elasticity of per capita calories to per capita expenditure, Maharashtra, India, 1983.

dence intervals from the two bootstraps. Here we see more clearly that the elasticity does indeed decrease with increases in PCE, but the decline is steady, from around 0.65 to 0.40 over the range shown in the figure. Of course, the slopes at the extremes of the distribution are quite imprecisely estimated. Around median PCE, the slope is near 0.5 and is precisely estimated.

Figures 4 and 5 correspond to figures 2 and 3 but show the price per calorie and the elasticity of the price per calorie. The bandwidth is 0.4, and we show only the more conservative clustered bootstrap standard errors. As was the case for the calorie relationship, the curve is close to being linear, and once again there is some evidence that the elasticity is higher for poorer households. Figure 5 confirms the decline in the elasticity: if one ignores the very poorest households, the slope declines from 0.4 to 0.3 within fairly wide confidence bands. These results do not suggest that the preliminary log-linear regressions of the previous section are likely to be misleading.

It is also possible to use nonparametric regressions to examine the extent to which the elasticity of the calorie price results from within-group substitution as opposed to between-group substitution. The inverse of the price per calorie, calories per rupee, is given by the identity

$$\frac{c}{x_f} = \sum_G w_G \pi_G = \sum_G w_G \sum_{i \in G} \frac{k_i q_i}{x_G}, \tag{9}$$

where, as before, $w_G$ is the share of the food budget devoted to good $G$, $x_f$ is expenditure on food, and $\pi_G$, which is defined by the second equality, is calories per rupee devoted to good $G$. As we move from poor to rich, the ratio of $c$ to $x_f$, and thus the price per calorie, is influenced both by changes in the budget shares $w_G$ as households substitute between groups and by changes in the group-specific calorie price inverses $\pi_G$ as quality and nutrient substitution takes place within groups. We can neutralize the within-group effect by calculating the values of the $\pi$'s at the sample mean and using these numbers in place of the actual $\pi$'s to recalculate (9) and, thus, to get an adjusted log calorie price, ln $P$, that excludes within-group substitution:

$$\ln P = -\ln\left(\sum_G w_G \bar{\pi}_G\right). \tag{10}$$

The variation in ln $P$ with per capita outlay reflects the changing allocation of the budget over groups of goods, but with the price per calorie within the groups held constant.

The results are shown in figures 6 and 7. The pictures are in clearest focus for a bandwidth of 0.4, and figure 6 shows the regression

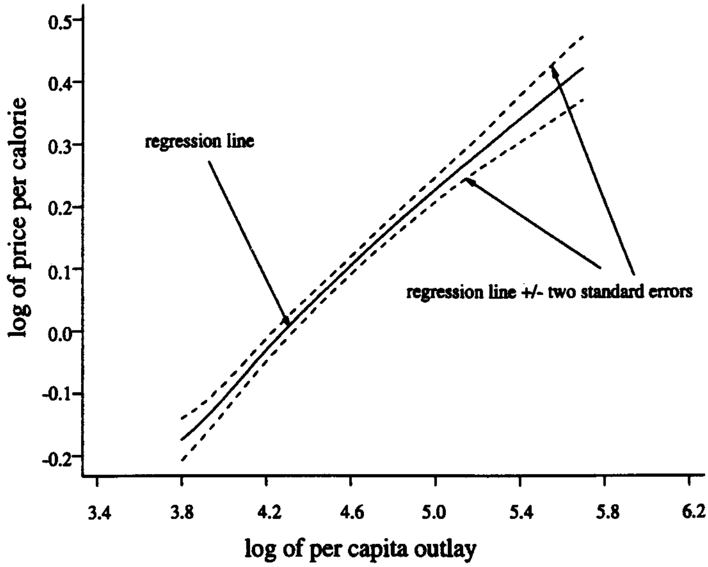


FIG. 4.—Log of price per calorie and log of per capita expenditure, Maharashtra, India, 1983.

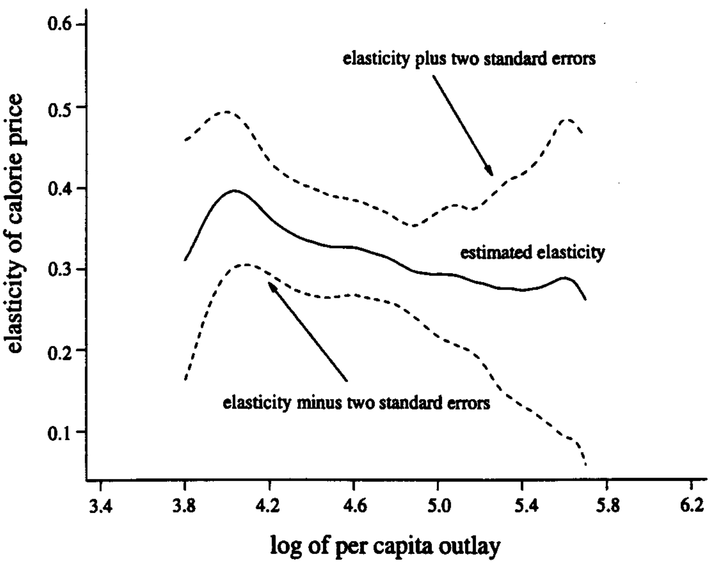


FIG. 5.—Elasticity of calorie price with respect to per capita expenditure, Maharashtra, India, 1983.

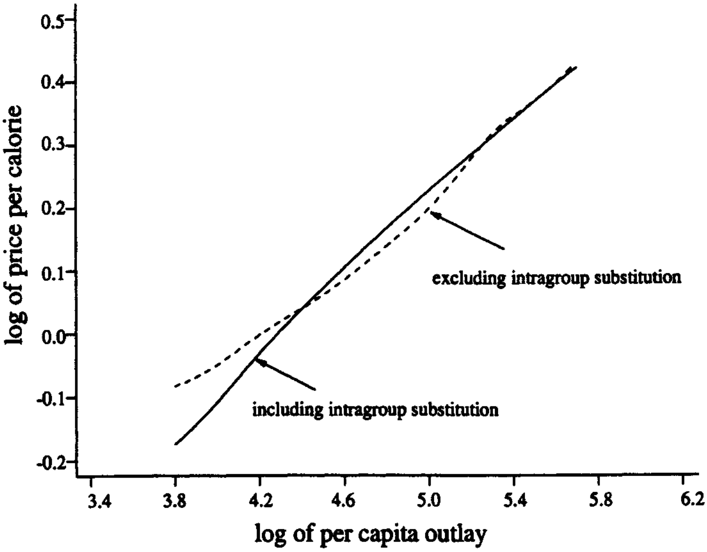


FIG. 6.—Calorie price, per capita expenditure, and intragroup substitution, Maharashtra, India, 1983.

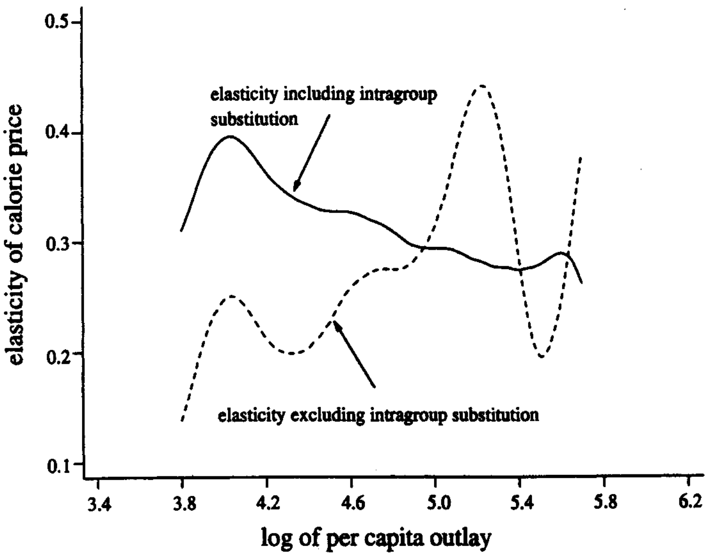


FIG. 7.—Elasticity of calorie price and intragroup substitution, Maharashtra, India, 1983.

functions for the price of calories, $\ln(x_f/c)$, shown as the solid line, and ln $P$ from (10), shown as the broken line. In contrast to figure 4, we use the calorie total unadjusted for meals given and taken, since it is only this figure that can be disaggregated across groups in order to implement (10). Figure 6 is largely what we would expect, given the results in Section III, as well as the simple mechanics of (10). When the fixed values of group calorie prices are used, the poor pay more per calorie, so that the calorie price to PCE line is, at least on average, rotated clockwise. Note also, however, that the close to straight line relationship for the actual price, which we have already seen in figure 4, is replaced by something rather more complex for the calorie price without intragroup substitution. The difference is more starkly portrayed in figure 7, where we can see not only that the elasticity of the calorie price is (typically) lower for the adjusted data but that the pattern of the elasticity with PCE is changed. The solid line in figure 7 is essentially the same as the line in figure 5 and shows an elasticity of around 0.35, declining gently with PCE. The broken line, by contrast, shows that the contribution of intergroup substitution to the elasticity of the calorie price is quite low for poorer households. The difference between the solid and broken lines is a measure of the contribution of intragroup substitution to the elasticity of the price, and this measure is high, 0.20 or so, for the poorest households, decreasing to become zero or negative for households in the top half of the PCE distribution. This result is an important modifier to the preliminary estimates in Section III, where within-group substitution was seen to be less important than between-group substitution. For the poorest households, which consume large amounts of the coarse cereals, quality upgrading within sorghum, between a hybrid and a preferred but more expensive local variety, makes a quantitatively important contribution to the increase in the cost of their calories.

The nonparametric regression techniques of this section are useful for exploring bivariate relationships but become more problematic when we recognize the potential role of other variables. Prices vary over time through the survey year, and they vary from place to place within rural Maharashtra. Individuals who do heavy labor need more calories than clerks or schoolteachers. But perhaps the single most important influence ignored so far is that of the size and demographic composition of the household. Holding constant PCE, we should expect lower per capita calorie consumption for those households that contain a high ratio of children to adults. Even if we compare only all-adult households, it is not obvious that in a household with twice as many members and twice the resources, household members will each choose to consume twice as many calories. Shared household public goods, or economies of scale within the household, might release resources that would permit *more* consumption of food. All these factors cannot simultaneously be examined using nonparametric methods. A sample of 5,630 households is more than adequate for bivariate relationships but is less than what would be required for the list of variables given above. Indeed, a case can be made for allowing for separate effects for each of the 563 villages in the sample; with no parametric structure, we would have to treat each village separately, so that there would be a mere 10 observations with which to examine all the other effects. Clearly, it is necessary to impose more structure.

Household size is a good issue with which to start. Figure 8 shows the relationships between log calorie availability and log PCE for each household size from 1 through 8. The most precisely estimated of these relationships uses data on 1,100 households with five members; for the others the sample size varies from 283 at size 1, up to 1,100 for size 5, and falling to only 306 at size 8. Although some of the curves cross and touch, the relationship is close to monotonic; the highest curve pertains to one-person households and the lowest to those with eight persons. The fact that per capita calorie consumption declines with household size reflects the facts that larger households are those with larger fractions of children and that children eat less than adults. However, there is another feature of figure 8 that should be noted and will reappear in the subsequent analysis. Compared with figure 2, in which household size and composition are projected onto log PCE, the curves are consistently flatter. Over a difference in log PCE of 2.0, from 3.8 to 5.8, log per capita calories grows by 0.8, from 7.2 to 8.0, in figure 2, but by about 0.7 in figure 8. Hence, and as far as can be told from the graphs, the inclusion of household size has the consequence of reducing the calorie elasticity, from about 0.4 to 0.35. This is an entirely reasonable result. In rural Maharashtra, as is usually the case, total household expenditure rises with household size, but not in full proportion, so that there is a negative correlation in the data between household size and *per capita* household expenditure. Since larger household size depresses per capita calories at any given level of PCE, the regression of calories on PCE with no allowance for household size will cause the latter to be projected onto the former and will bias upward the slope of the regression. As is to be expected from the approximately linear shape of the regression lines, the result can be replicated in a simple linear regression. When household size is excluded, the least-squares regression of log per capita calories on log PCE yields a coefficient of 0.44, which falls to 0.38 when the logarithm of household size is added.

Once we control for household size, the inclusion of other variables

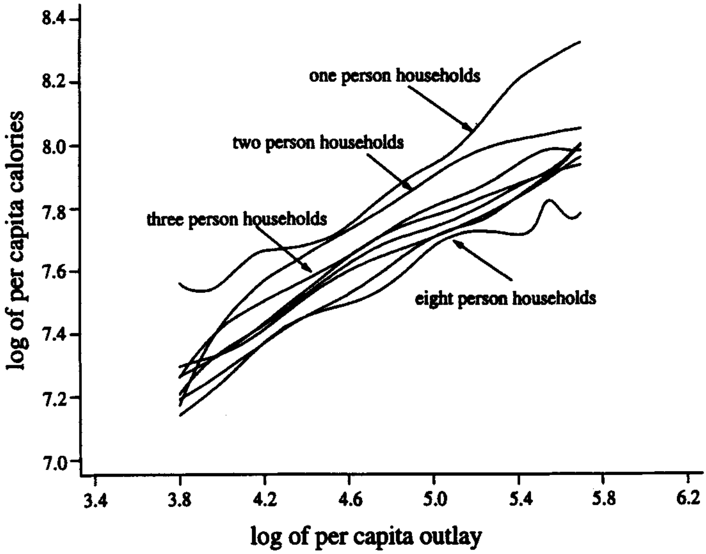


FIG. 8.—Calories and per capita expenditure for households of different sizes, Maharashtra, India, 1983.

has relatively little further effect on the estimate of the elasticity of calories with respect to PCE. Table 2 shows the results of four further regressions. They all contain demographic, caste and religion, and labor type variables. As before, the log of total household size is included, but we now add nine variables for demographic structure in the form of the fraction of members in various age and sex groups. We also add various religion and caste variables, as well as dummies for whether the head is or is not self-employed and is or is not in agriculture. The coefficients in column 1 are those from an ordinary least squares (OLS) regression of log per capita calories on the covariates, and those in column 2 are those obtained when dummy variables are included for each of the 563 villages in the sample. Columns 3 and 4 repeat the exercise for the log of the price paid per calorie. Many of the variables have well-determined effects on calories and on the calorie price. Adults clearly consume more than children, those who work in agriculture consume more calories than those who do not, and it is generally the case that factors that are associated with higher calories are associated with lower calorie price. Even conditional on PCE, quality and quantity are negatively associated: the more one needs the calories, the less one is willing to spend on them. The village effects do not attract very large $F$-statistics, nor does their inclusion change the coefficients in any major way. Indeed, the coefficients on the logarithm of PCE are much as they are in either the

TABLE 2

OLS Estimates of Double Log Calorie and Calorie Price Regressions with Other Covariates

| | Log Calorie Availability | | | | Log Price per Calorie | | | |
|---|---|---|---|---|---|---|---|---|
| | All Data (1) | | Within Village (2) | | All Data (3) | | Within Village (4) | |
| | β | \|t\| | β | \|t\| | β | \|t\| | β | \|t\| |
| Constant | 6.028 | (78) | | | −1.5934 | (18) | | |
| ln PCE | .3655 | (29) | .3407 | (27) | .3799 | (25) | .3217 | (23) |
| ln household size | −.1572 | (14) | −.1630 | (21) | .0839 | (6.8) | .0661 | (8.4) |
| rm04 | −.0967 | (2.2) | −.1461 | (4.1) | .1024 | (2.3) | .1008 | (3.3) |
| rm59 | .0488 | (1.2) | .0321 | (1.0) | −.0467 | (1.2) | −.0331 | (1.2) |
| rm1014 | .0891 | (1.9) | .0612 | (1.9) | −.1120 | (2.3) | −.0842 | (2.9) |
| rm1555 | .1636 | (5.1) | .1634 | (5.9) | −.1700 | (4.3) | −.1347 | (5.0) |
| rm55+ | .1406 | (3.0) | .1213 | (2.8) | −.1565 | (3.6) | −.1074 | (2.9) |
| rf04 | −.1359 | (3.1) | −.1869 | (4.9) | .0460 | (1.1) | .0742 | (2.2) |
| rf59 | .0176 | (.4) | −.0040 | (.1) | −.0643 | (1.4) | −.0476 | (1.4) |
| rf1014 | .1140 | (2.8) | .0679 | (2.0) | −.1108 | (2.7) | −.0873 | (3.0) |
| rf1555 | .0420 | (1.6) | .0514 | (2.1) | .0085 | (.3) | −.0021 | (.1) |
| Scheduled caste | −.0083 | (.8) | −.0179 | (2.0) | .0020 | (.2) | −.0071 | (.8) |
| Hindu | .0114 | (.7) | .0302 | (2.1) | −.0562 | (2.6) | −.0605 | (4.4) |
| Buddhist | .0237 | (1.1) | .0400 | (2.0) | −.1080 | (4.0) | −.0760 | (4.0) |
| Self-employed nonagriculture | .0187 | (1.0) | .0064 | (.4) | −.0270 | (1.1) | .0079 | (.5) |
| Agricultural labor | .0433 | (2.2) | .0222 | (1.4) | −.0837 | (3.4) | −.0418 | (2.7) |
| Nonagricultural labor | .0275 | (1.1) | .0293 | (1.5) | −.0210 | (.8) | −.0315 | (1.7) |
| Self-employed agriculture | .0618 | (3.5) | .0389 | (2.7) | −.0610 | (2.8) | −.0118 | (.8) |
| $R^2$ | .5532 | | .6706 | | .4254 | | .6414 | |

Note.—Variables beginning with *r* are demographic ratios, so that, e.g., rf59 is the ratio of females aged 5–9 to total household members, and rm55+ is the ratio of males older than 55. There are four labor type dummies, self-employed or employed, in agriculture or not. The omitted category is "other labor." The omitted religion/caste variable is Jain and other. The within regressions contain 563 dummy variables for the villages and do not contain a constant term. The (uncorrected) *F*-tests for the exclusion of the village effects are 3.19 with 562 and 5,043 degrees of freedom for the calorie regression and 5.39 with 562 and 5,042 degrees of freedom for the calorie price regression. The reported absolute *t*-values are corrected for heteroskedasticity and, in the case of the all-data regressions, for the cluster structure of the sample.

linear or nonparametric regressions. The total expenditure elasticity of expenditure on food is close to 0.75, and it is about evenly split between the elasticity of calories and the elasticity of the calorie price.

## V. Caveats, Conclusions, and Relation to the Broader Literature

In this final section, we try to summarize what we have learned and to put our results in the context of the broader literature and other related evidence. We also discuss a number of econometric and interpretational issues that have been raised in the literature and that have not been dealt with so far.

Our main objective has been to use the data from rural Maharashtra to characterize the relationship between per capita calorie availability and per capita total household expenditures. We have done so by modeling the expectation of the logarithm of per capita calories conditional on the logarithm of PCE, sometimes conditioning on other covariates, particularly household size and composition. The logarithmic transformation was chosen because the empirical joint density of the logarithms is a good deal closer to joint bivariate normality than for any other obvious data transformation, so that the regression functions are close to being linear. Furthermore, if normality is accepted, at least as an approximation, other functions can be calculated, most obviously the regression function of calories on expenditure. When we conditioned only on ln PCE, we found an average elasticity around 0.45, with some evidence of decline with PCE from over 0.55 for the poorest groups to around 0.40 at the top of the range. This curvature, although real, is much less sharp than that found by Strauss and Thomas (1995) for Brazil, for one of the two Filipino data sets they examine, as well as for the Indian ICRISAT data (to be discussed below). The regression function of the logarithm of price paid per calorie and ln PCE is also close to being linear, with an elasticity of around 0.30. If household size is added to the conditioning variables, the expenditure elasticity of calories falls somewhat, to a little below 0.40, and that of the price of calories rises, to much the same figure. Conditional on household size, the total expenditure elasticity of food expenditures is about equally split between the elasticity of calories and the elasticity of the price of calories. A 10 percent increase in food expenditure is associated with a 5 percent increase in calorie consumption and a 5 percent increase in the price paid per calorie. While these figures are far from the naive position that the elasticity should be one, they are close to what was once the standard view and are far from the extreme revisionist position of Behrman and Deolalikar (1987), Bouis and Haddad (1992), and Bouis (1994) that the elasticity is zero or close to it.

We have interpreted the relationship between nutrition and expenditure as a demand function for calories and have ignored the possibility of reverse causation: that higher income earning opportunities are open to those who are better nourished, the relationship that is at the center of efficiency wage accounts of unemployment and destitution in the subcontinent, particularly Dasgupta (1993, chap. 16). We are skeptical that the feedback from nutrition to wages can be dealt with by standard simultaneous equation techniques, such as instrumentation. The models of Mirrlees (1975) and Stiglitz (1976) are inherently nonlinear, and the behavior that they predict is not accommodated by the routine application of instrumental variables. As argued by Bliss and Stern (1978), it is far from obvious that the model is identified: income affects nutrition, and nutrition (supposedly) affects income, and any variable that enters one equation can arguably be included in the other. Nevertheless, the Maharashtran data are not easily reconciled with the existence of a nutrition trap. In this part of India in 1983, it was possible to buy 2,000 calories for a good deal less than 2 rupees. Since most calories, perhaps 70 percent, are required to maintain the body even when resting, an extra 600 calories, say, will provide enough energy for substantial additional effort, indeed a full day's farming activities according to Dasgupta (1993, p. 422). These calories are available for 60 paise, about 4 percent of the daily wage; see also Srinivasan (1994) for a similar argument. There is also a body of empirical evidence that the labor market in India does not behave according to the predictions of the theory (see, e.g., Rosenzweig [1984] and Dasgupta [1993, chap. 16] for some counterarguments).

One immediate question is whether it is possible to reconcile our results with those for the Indian ICRISAT villages used by Behrman and Deolalikar (1987). The preferred estimate of the elasticity from that study is 0.37, which is obtained from a differenced regression and has an estimated standard error of 0.37. Although this estimate is clearly in accord with our own much more precise estimates, the agreement actually conceals quite large differences between the ICRISAT and Maharashtran state data. Figure 9 shows the calorie expenditure regression functions for both sets of data. The higher curves are those in figure 2, and those (mostly) below come from the application of the same local regression and bootstrapping techniques to the 253 observations that come from pooling the two years of the ICRISAT data. The log of per capita outlay has been converted to 1983 prices using the official consumer price index for agricultural laborers; this index has problems, and the revaluation is at best ap-

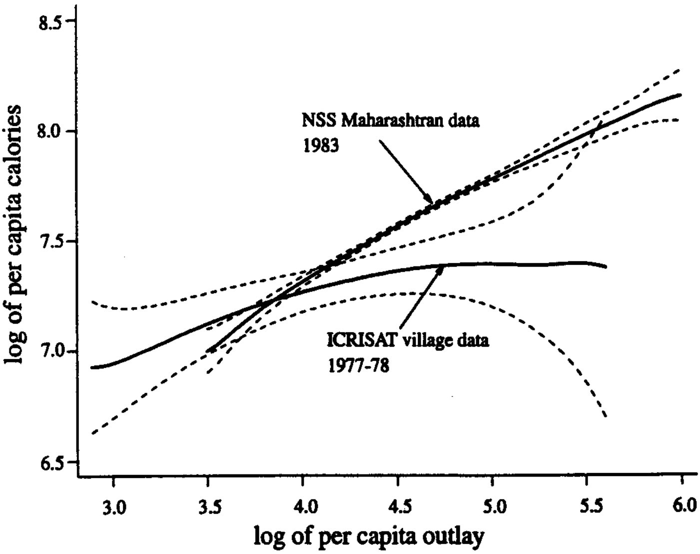


FIG. 9.—Calorie regression functions, NSS Maharashtran data and ICRISAT village data.

proximate. Even so, the fact that the ICRISAT villagers are poorer than the Maharashtran average is consistent with other evidence. The ICRISAT data give a lower slope, and even with the bandwidth of 1.5 that is required by the relatively few observations, there appears to be more curvature than in the Maharashtran data, though the wide confidence bands would permit a linear regression. There are two leading explanations for the differences in slope. The ICRISAT calorie data come from 24-hour recall of actual consumption, whereas the NSS data are the total calories available from food purchases. If richer households buy more foods for others, even after adjustment, waste more, or give more food to pets and animals, the availability data will overstate the elasticity compared with the intake data. The other explanation invokes measurement error in per capita outlay in the ICRISAT data. The NSS has a long and well-validated track record in collecting data on expenditures, and the accurate estimation of per capita expenditure is one of the main aims of the survey. By contrast, expenditure data can be indirectly inferred for the ICRISAT households only by tracking purchases and disappearance of crops and food, and the procedure is known to be subject to great uncertainty. Standard attenuation bias will artificially flatten the ICRISAT curve, and if the variance of measurement error is larger for better-off households, the estimated regression function will be subject to progressive attenuation bias, which will induce a spurious concavity.

There are a number of economic and econometric reasons why the regressions on the NSS data might overstate the calorie elasticity. First, there is the issue of whether we should prefer the lower estimates that come from conditioning on household size. The answer depends on the view we wish to take about the determinants of fertility or, more generally, since migration, fostering, and family restructuring can also be important, about what determines household size. If household size depends in an important way on income and calorie consumption, the regression of calories on expenditure, unconditional on household size, can be regarded as an estimate of the reduced-form or long-run relationship between calories and total expenditure once fertility and migration effects have worked themselves out. While we certainly recognize that fertility and family formation are subject to economic influences, we doubt that they are sufficiently well predicted by nutrition and expenditure to render the standard simultaneous equations econometric model a very useful framework for thinking about the issue. We find it more useful to treat household size and structure as exogenous variables and to accept the rather lower estimate. For those who prefer the other view, the slope of the regression of total calories on *total* rather than *per capita* expenditure is higher even than the higher of our two estimates, around 0.65. There is certainly nothing in the endogenous structure story that would suggest that our estimates of the calorie elasticity are too low.

Second, there are questions about correlations between total expenditure and the error term. This could arise if nutrition constrained earnings, and we have already explained our skepticism about the importance of this effect and whether, if it were important, it could be dealt with by instrumentation. A more pressing problem is measurement error in total expenditure and in calorie availability. Total household expenditure is the sum of food expenditure and nonfood expenditure, each of which is certainly measured with some error. Food expenditure is the sum of a large number of components, the same components that, appropriately scaled, make up the estimate of total calorie availability. Total expenditure is therefore measured with error, and the error of measurement is positively correlated with the composite error term in the regression, itself partly determined by the measurement error in calories. Note that the correlation between the measurement errors in the dependent and independent variables means that this is *not* a standard errors-in-variables problem. Bouis and Haddad (1992) have examined the issue in the context of a linear model and argue that the upward bias from the correlated errors will typically outweigh the standard downward attenuation bias from the measurement error in total expenditure, leaving a net upward bias. Nonlinearity complicates the issue a good deal, so we explore the issue within a log-linear structure; given our previous results, we can at least hope that this is a reasonable approximation.

Suppose that the true model is that food is a log-linear function of total expenditure, so that

$$\ln \tilde{f} = \alpha + \beta \ln \tilde{x} + u, \tag{11}$$

where $\alpha$ and $\beta$ are the parameters, and tildes indicate unobserved true values; we suppose that $\beta$ is the parameter of primary interest. For the purposes of the current exercise, we may temporarily imagine that food expenditure converts into calories at a fixed proportional rate. To the extent that calories rise less rapidly than food expenditure and that the measurement error in calories is less than perfectly correlated with the measurement error in food, the bias will be less; food proportional to calories is the worst case. Suppose that both food and nonfood are measured with proportionate measurement errors so that

$$\ln f = \ln \tilde{f} + \epsilon_f; \; \ln g = \ln \tilde{g} + \epsilon_g, \tag{12}$$

where $g$ is expenditure on nonfood. We assume that the measurement errors $\epsilon_f$ and $\epsilon_g$ are mutually independent and that each is independent of the regression error $u$ in (11).

Total expenditure $x$ is defined as the sum of food and nonfood, so that, with minimal manipulation, we reach

$$x = f + g = \tilde{x}[(1 - w)e^{\epsilon_f} + we^{\epsilon_g}], \tag{13}$$

where $w$ is the budget share of nonfood. Taking logarithms, we have

$$\ln x = \ln \tilde{x} + (1 - \theta)\epsilon_f + \theta\epsilon_g, \tag{14}$$

where we have used the first-order approximation

$$\ln[1 - w(1 - e^{\epsilon_g - \epsilon_f})] \simeq \theta(\epsilon_g - \epsilon_f), \tag{15}$$

and $\theta$ is the budget share of nonfood evaluated at the sample mean, $\bar{w}$. If (14) is substituted into the food equation to set up a standard errors-in-variables analysis, we have

$$\ln f = \alpha + \beta \ln x + u + [1 - \beta(1 - \theta)]\epsilon_f - \beta\theta\epsilon_g. \tag{16}$$

Hence, the asymptotic bias of $b$, the OLS estimate of $\beta$, is given by

$$\text{plim}\, b - \beta = m^{-1}\{(1 - \theta)[1 - \beta(1 - \theta)]\sigma_f^2 - \beta\theta^2\sigma_g^2\}, \tag{17}$$

where $m$ is the asymptotic variance of $\ln x$. If the variance of measurement error in nonfood were large enough, it is possible for this bias to be negative. This seems unlikely, especially given the fact that $\theta$ is less than a third.

The model as written above is not identified without further assumptions, so that there is no way to correct the bias and obtain a consistent estimate of $\beta$. However, consider the result of estimating the food equation by instrumental variables, using the logarithm of nonfood expenditure as an instrument. Denoting this estimator by $\tilde{\beta}$, we have

$$\text{plim}\,\tilde{\beta} - \beta = \frac{\text{plim}\,n^{-1}\sum u_i \ln g_i - \beta\theta\sigma_g^2}{\text{plim}\,n^{-1}\sum \ln x_i \ln g_i}. \tag{18}$$

The first term in the numerator is negative. Conditional on the true value of total expenditure, a positive value of $u$ implies that food expenditure is above its predicted value so that nonfood expenditure must be below its predicted value. In consequence, this instrumental variable estimator is biased downward, so that the probability limits of the OLS and instrumental variable estimators provide upper and lower bounds for the true value. In practice, the comparison of the two estimators can give us some indication of the maximum possible size of the upward bias that results from the common measurement error.

Results are given in table 3 for the three most important linear regressions: the logarithm of total calories on the logarithm of total household expenditure, and the logarithm of per capita total calories on the logarithm of per capita household expenditure, with and without the addition of the logarithm of household size. Note that while the upper bounds correspond to the previous estimates ignoring the effects of correlated measurement error and will be consistent in the absence of such error, the lower bounds are biased down whether or not measurement error is present. The upper bounds may or may not be too high, but the lower bounds are certainly too low. For all three cases, the lower bounds to the elasticities are approximately 0.1 less than the figures previously estimated. While we have no way of knowing where within the band we should look, it is clear that not *all* of the estimated elasticity can be accounted for by the effects of measurement error. Recognition of the errors should make us cautious in claiming that the per capita calorie elasticity with respect to per capita total expenditure is 0.5 rather than 0.4 or is 0.4 rather than 0.3 conditional on household size, but we can resist further revision downward.

Finally, we note an important issue about the measurement of calories. As is often the case, the calorie data from the Maharashtran sample come from a *consumption* survey, where the weights of foods reported as having been consumed are multiplied by their presumed calorie content so as to provide an estimate of calorie *availability*. An

TABLE 3

UPPER AND LOWER BOUNDS FOR VARIOUS ELASTICITY ESTIMATES

| Dependent Variable | Independent Variable | Instruments | Upper Bound | Lower Bound |
|---|---|---|---|---|
| ln calories | ln THE | ln *g* | .6489 (.00643) | .5413 (.00957) |
| ln per capita calories | ln PCE | ln *g* | .4393 (.00629) | .3343 (.00866) |
| ln per capita calories | ln PCE | ln *g*, ln *n* | .3780 (.00623) | .2805 (.00831) |

NOTE.—THE is total household expenditures, PCE is per capita total household expenditures, and *g* is total nonfood expenditures. Standard errors are in parentheses.

alternative method is direct nutritional monitoring, where enumerators observe actual consumption, weighing and measuring the cooked and prepared foods or, less intrusively, where the respondent is asked to recall the previous day's meals, including the preparation technique, the recipe, and the allocation between individuals. These methods provide estimates of calorie *intake*. Bouis and Haddad (1992) obtain their results from direct recall data, and Bouis (1994) confirms using both Filipino and Kenyan evidence that direct recall data yield lower elasticities than data based on consumption surveys. Behrman and Deolalikar's (1987) estimates are also based on directly collected nutritional data. The evidence summarized in Strauss and Thomas (1995) also provides at least some support for the idea that estimates based on intake are lower than those based on availability. Bouis and Haddad attribute the difference between availability and intake to meals provided for servants and workers, to common measurement error in calories and total expenditure, and to wastage of calories in preparation, through "plate wastage," or to food fed to pets or domestic animals. In this paper, we have removed the calories from servants' meals from the total, and by not removing the corresponding expenditures from total expenditure, we have biased our estimates down, not up. We do not deny the potential importance of correlated measurement errors, but we have demonstrated that they cannot be responsible for generating our results if the true elasticity is close to zero. There is nothing we can do about cooking and plate wastage, although we find it an implausible source for a large overestimation of the elasticity. If the intake elasticity were truly zero, then we would have to believe that, as incomes rise, households buy more food and more calories for the pleasure of throwing them away. The hypothesis that intake data are less accurate than availability data is no more plausible than the hypothesis that household incomes or expenditures are less well measured in special nutritional surveys, so that the lower elasticities from the latter are attributed not to poor calorie measurement in the consumption surveys, but to attenuation bias when income or outlay data from the nutrition surveys are used.

There are two other types of evidence that are relevant to this debate. Bouis (1994) argues that, in the long run, body weight is proportional to calorie consumption. As a result, even if the calorie to per capita expenditure elasticity were as high as 0.3, people in the top per capita expenditure decile, who might be six times better off than people in the bottom decile, would weigh 71 percent more on average, something that we do not observe. Like Strauss and Thomas (1995), who discuss some of the relevant nutritional literature, we are skeptical that the weight to calorie relationship is simple enough to permit this sort of calculation. Indeed, the recent experimental evidence of Leibel, Rosenbaum, and Hirsch (1995) shows that metabolism increases when the body is above its "natural" weight, so that changes in calorie consumption are offset and do not automatically lead to changes in body weight. Applied to decreases in calorie consumption, the phenomenon is familiar to many unsuccessful dieters. The second source of evidence is aggregate time series. According to the Food and Agriculture Organization, per capita calorie consumption in India in 1992 was 2,394 calories, compared with 2,006 in 1971. Comparing this with the 54 percent increase in per capita real income over the same period reported by the World Bank, we get an elasticity of 0.41. The OLS regression of annual data from 1971 to 1992 gives an elasticity of 0.38, so that the Indian time-series evidence is very close to our estimates from the survey data. Similar calculations for some other countries yield a mixed set of estimates: 0.13 for Bangladesh and Pakistan, 0.25 for China, 0.26 for Indonesia, 0.41 for Kenya, and 0.65 for the Philippines. Even so, there is nothing in these numbers to suggest that the elasticity is zero. Our conclusion is that the range of estimates that we have established for the expenditure elasticity of calories, from 0.3 to 0.5, is the right one for this part of rural India.

## References

Behrman, Jere R., and Deolalikar, Anil B. "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India." *J.P.E.* 95 (June 1987): 108–38.

Bliss, Christopher J., and Stern, Nicholas H. "Productivity, Wages and Nutrition." 2 parts. *J. Development Econ.* 5 (December 1978): 331–98.

Bouis, Howarth E. "The Effect of Income on Demand for Food in Poor Countries: Are Our Food Consumption Databases Giving Us Reliable Estimates?" *J. Development Econ.* 44 (June 1994): 199–226.

Bouis, Howarth E., and Haddad, Lawrence J. "Are Estimates of Calorie-Income Elasticities Too High? A Recalibration of the Plausible Range." *J. Development Econ.* 39 (October 1992): 333–64.

Cleveland, William S. "Robust Locally Weighted Regression and Smoothing Scatterplots." *J. American Statis. Assoc.* 74 (December 1979): 829–36.

Dasgupta, Partha. *An Inquiry into Well-Being and Destitution.* Oxford: Clarendon, 1993.

Efron, Bradley, and Tibshirani, Robert J. *An Introduction to the Bootstrap.* London: Chapman and Hall, 1993.

Fan, Jianqing. "Design-Adaptive Nonparametric Regression." *J. American Statis. Assoc.* 87 (December 1992): 998–1004.

———. "Local Linear Regression Smoothers and Their Minimax Efficiencies." *Ann. Statis.* 21 (March 1993): 196–216.

Fogel, Robert W. "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy." *A.E.R.* 84 (June 1994): 369–95.

Gopalan, C.; Sastri, B. V. Rama; and Balasubramanian, S. C. *Nutritive Value of Indian Foods.* Hyderabad: Nat. Inst. Nutrition, 1971.

Leibel, Rudolph L.; Rosenbaum, Michael; and Hirsch, Jules. "Changes in Energy Expenditure Resulting from Altered Body Weight." *New England J. Medicine* 332 (March 9, 1995): 621–28.

Leibenstein, Harvey A. *Economic Backwardness and Economic Growth: Studies in the Theory of Economic Development.* New York: Wiley, 1957.

Mirrlees, James A. "A Pure Theory of Underdeveloped Economies." In *Agriculture in Development Theory,* edited by Lloyd G. Reynolds. New Haven, Conn.: Yale Univ. Press, 1975.

Ravallion, Martin. "Income Effects on Undernutrition." *Econ. Development and Cultural Change* 38 (April 1990): 489–515.

Reutlinger, Shlomo, and Selowsky, Marcelo. *Malnutrition and Poverty: Magnitude and Policy Options.* Baltimore: Johns Hopkins Univ. Press (for World Bank), 1976.

Rosenzweig, Mark R. "Determinants of Wage Rates and Labor Supply Behavior in the Rural Sector of a Developing Country." In *Contractual Arrangements, Employment, and Wages in Rural Labor Markets in Asia,* edited by Hans P. Binswanger and Mark R. Rosenzweig. New Haven, Conn.: Yale Univ. Press, 1984.

Silverman, Bernard W. *Density Estimation for Statistics and Data Analysis.* London: Chapman and Hall, 1986.

Srinivasan, T. N. "Destitution: A Discourse." *J. Econ. Literature* 32 (December 1994): 1842–55.

Stiglitz, Joseph E. "The Efficiency Wage Hypothesis, Surplus Labour, and the Distribution of Income in L.D.C.s." *Oxford Econ. Papers* 28 (July 1976): 185–207.

Strauss, John, and Thomas, Duncan. "The Shape of the Calorie-Expenditure Curve." Manuscript. Santa Monica, Calif.: Rand Corp.; New Haven, Conn.: Yale Univ., June 1990.

———. "Human Resources: Empirical Modeling of Household and Family Decisions." In *The Handbook of Development Economics,* vol. 3, edited by Jere R. Behrman and T. N. Srinivasan. Amsterdam: North-Holland, 1995.