Journal of Development Economics 39 (1992) 333–364. North-Holland

Are estimates of calorie–income elasticities too high?

A recalibration of the plausible range

Howarth E. Bouis and Lawrence J. Haddad*

International Food Policy Research Institute, Washington DC, USA

Received October 1990, final version received July 1991

The wide range of calorie–income elasticities in the literature results, in large part, from the particular calorie and income variables used for estimation. Elasticities across four estimation techniques and four calorie–income variable pairs for a sample of Philippine farm households, ranged from 0.03 to 0.59. Estimates associated with calorie availability are biased upwards, first, because random errors in measuring food purchases are transmitted (by construction) both to calorie availability and total expenditures, and second, because the residual difference between family calorie intake and household calorie availability will often increase with income. The calorie intake–total expenditure variable pair gives the preferred elasticity estimate in the 0.08–0.14 range.

1. Introduction

What effect do increases in income of the poor in developing countries have on their levels of calorie consumption? While accurate measurement of calorie–income elasticities is fundamental to sound nutrition policy analysis, the literature has produced a perplexingly wide range of elasticity estimates for this relationship at the household level (see table 1). This has led to contradictory policy conclusions [for a discussion, see Behrman (1988)]. Initial estimates in the literature tended to be high (often above 0.4), which led to a policy conclusion that increasing income is a sufficient condition for better nutrition. More recently, estimates have appeared which are much lower [in one case not significantly different from zero, Behrman and Deolalikar (1987)], suggesting that increasing income may not be a sufficient, and possibly not even a necessary condition, for nutritional improvement.

We argue in this paper that much of the variation in calorie–income

Correspondence to: Dr. Howarth E. Bouis, International Food Policy Research Institute, 1200 17th St., NW., Washington, DC 20036-3006, USA.

*We would like to thank Harold Alderman, Jere Behrman, Duncan Thomas, and an anonymous referee for many helpful comments. All remaining errors and omissions are ours.

0304-3878/92/$05.00

elasticity estimates can be explained very simply by the particular calorie and income variables that are used in the regression analysis. Specifically, we focus on the expected relationships between four possible pairs of dependent and explanatory variables, calorie *availability* and calorie *intake* (alternative dependent variables), and *total expenditures* and *current income* (alternative explanatory variables).

Two key points will be made. First, any random errors in measuring food purchases are transmitted (by construction) both to calorie availability and total expenditures resulting in correlation between measurement errors for these two variables, which in turn leads to an upwardly biased estimated coefficient in total expenditures using ordinary least squares (OLS). Second, the residual difference between family calorie intake and household calorie availability will often increase as a percentage of total food expenditures as income increases. If this is the case, then systematic underestimates of meals served to non-family members will be positively correlated with any income variable, and this will also result in an overestimate of the true elasticity, which cannot be eliminated (and may even be exacerbated) by use of instrumental variable (IV) or panel estimating techniques.

In earlier studies presented in section 2 of the paper, data availability constrained the regression analyses to one pair of variables, often calorie availability and total expenditures, because of cost considerations (expenditure surveys are less costly to administer than calorie intake and current income surveys) and research considerations (total expenditures are thought to measure permanent income more accurately; Friedman (1957)]. Our data set, collected for rural households in the Philippines, allows for a comparison of estimates obtained by using all four calorie and income variable pairs mentioned above. Depending on the variable pair and estimating technique used, elasticities range from 0.03 to 0.59.

For reasons elaborated in section 3, which develops simple mathematical expressions that highlight the various econometric problems involved in estimating the calorie–income relation, we opt for the calorie intake–total expenditure pair as the most appropriate variables for measuring how household calorie intakes change with income. These elasticity estimates, ranging from 0.08 to 0.14 at the low end of estimates obtained, are presented in section 4. Final conclusions are drawn in section 5.

2. Previous estimates of calorie–income elasticities

Table 1 illustrates the wide range of calorie–income elasticity estimates obtained from several low-income samples. Confusion arises in reaching policy conclusions considering that the studies in table 1 involve: (1) calorie–income relationships derived using all four variable-pairs mentioned earlier;[1] (2) estimations at different levels of food data aggregation; (3) different estimation techniques; and (4) data drawn from diverse populations.

The elasticity estimates in table 1 are presented in descending order of magnitude. There is a tendency for the larger elasticities to be derived from calorie availability and total expenditure data. However, it is difficult to isolate the effect of the calorie data source and the measure of income used upon the magnitude of the estimates obtained, because each data source is associated with a different level of food aggregation.

In one of the first studies to question high calorie elasticity estimates, Strauss (1984) suggests that the high level of food group aggregation is the main culprit. Household expenditure for a food group aggregate may increase substantially in response to higher income, without a proportionate increase in calorie intake because of within-group substitution toward more expensive calorie sources. Behrman and Deolalikar (1987) expand on this aggregation theme by introducing the concept of 'direct' (conversion of food group quantities into aggregate calories before estimation) and 'indirect' (computing a weighted average of food group expenditure elasticities after estimation) calorie–income elasticity estimation.

Behrman and Deolalikar calculate calorie–income elasticities for the same households from two different data sources: (1) calorie intake, from a 24-hour recall of 120 foods estimated as a function of predicted total expenditure; and (2) food group expenditure regressed on predicted total expenditure in a nonsystem framework, for only six aggregate food groups. They find that while the direct nutrient elasticities are not significantly different from zero, the indirect nutrient elasticities are close to one.

The authors attribute the difference in magnitudes to the fact that the food group calorie conversions are fixed across all households and, therefore, do not take account of intra-food group substitution as argued above. While the impression is left that the indirect method is intrinsically unsound, its only real failing appears to be that it has to be implemented with incomplete (very aggregate) data. We will argue below that use of calorie intakes to implement one methodology and use of calorie availability to implement the other is an additional factor explaining the divergence in elasticity estimates, perhaps of greater importance than the aggregate bias.[2]

A variety of econometric techniques have been used in the literature to control for simultaneity bias due to behavioral endogeneity, measurement

[1]The qualitative and quantitative differences between food availablilty and food intake data are well documented [Burk and Pao (1976)] as are the differences in what total expenditure and household income measure in a cross-sectional study [Anand and Harris (1985)].

[2]Price paid per kilogram for rice and corn, the two primary staples consumed, varied only slightly by income group for our Philippines data set. Both because the level of food group disaggregation is greater than 50 for both calorie variables and because quantity information is available for both calorie variables, we do not expect aggregation bias to be large relative to the other sources of bias discussed in this paper.

Table 1

Some calorie–income elasticity estimates from the literature ranked by magnitude at mean income.

Author(s)	Year	Dependent variable: Calories from	Independent variable: Household	Estimation method	Elasticity calculation	Level of food group aggregation for food-to-nutrient conversion factors	Country of study	Elasticity at mean
Behrman and Deolalikar	1987	Food expenditure	Total expenditure	IV, First difference fixed effects	Indirect	6	India	1.18[a]
Strauss	1984	Food expenditure	Total expenditure	Joint production-demand system	Indirect	5	Sierra Leone	0.82
Pitt	1983	Food expenditure	Total expenditure	Tobit demand system	Indirect	9	Bangladesh	0.82–0.78
Behrman and Deolalikar	1987	Food expenditure	Total expenditure	IV	Indirect	6	India	0.77[a]
Sahn	1988	Food expenditure	Total expenditure	Probit, truncated OLS	Indirect	14	Sri Lanka	0.62
Edirisinghe	1987	Food expenditure	Total expenditure	OLS	Direct	182	Sri Lanka	0.56
Chernichovsky and Meesook	1984	7-day recall intake	Total expenditure	OLS(?)	Direct	120	Indonesia[f]	0.54
Kumar and Hotchkiss	1988	Production disposal and food expenditure	Total expenditure	OLS	Direct	12	Nepal	0.51
Pinstrup-Andersen and Caicedo	1978	Food expenditure?	Income?	Demand system	Indirect	22	Colombia	0.51
von Braun, de Haen, and Blanken	1991	7-day recall foods cooked	Total expenditure	OLS	Direct	52	Rwanda	0.50
Timmer and Alderman	1979	Food expenditure	Total expenditure	OLS	Direct	NA[d]	Indonesia[f] (Rural)	0.47[d]
von Braun, Puetz, and Webb	1989	7-day recall foods cooked	Total expenditure	OLS	Direct	46	Gambia	0.48–0.37
Alderman	1987	Food expenditure	Total expenditure	Within fixed effects, OLS	Direct	42	India	0.44–0.41
Ward and Sanders	1980	Intake[e]	Income	OLS, 2SLS	Direct	?	Brazil	0.53–0.24
Behrman and Deolalikar	1987	24-hour recall intake	Total expenditure	IV, First difference fixed effects	Direct	120	India[g]	0.37[b]
Garcia and Pinstrup-Andersen	1987	Food expenditure	Total expenditure	OLS	Direct	124	Philippines	0.34–0.32
Alderman	1989	Food expenditure	Total expenditure	2SLS	Direct	38	Pakistan (Rural)	0.34–0.31

von Braun, Hotchkiss, and Immink	1989	Food expenditure	Total expenditure	OLS	Direct	108	Guatemala	0.31
Trairatvorakul	1984	Food expenditure	Income	OLS	Direct	13	Thailand	0.33–0.27
Timmer and Alderman	1979	Food expenditure	Total expenditure	OLS	Direct	NA[d]	Indonesia[f] (Urban)	0.26[d]
Alderman, Chaudhry, and Garcia	1988	Food expenditure	Total expenditure	OLS?	Direct	38	Pakistan (Urban)	0.22
Williamson-Gray	1982	7-day weighing intake	Total expenditure	OLS	Direct	?	Brazil[h]	0.18
Behrman and Deolalikar	1987	24-hour recall intake	Total expenditure	IV	Direct	120	India[g]	0.17[b]
Ravallion	1990	7-day recall intake	Total expenditure	GLS	Direct	166	Indonesia[f]	0.15
Kennedy	1989	24-hour recall intake	Total expenditure	OLS	Direct	168	Kenya	0.15
Strauss and Thomas	1989	7-day weighing intake	Total expenditure	2SLS	Direct	300	Brazil[h]	0.12[c]
Pitt, Rosenzweig, and Hassan	1990	24-hour weighing intake	Income	2SLS	Direct	?	Bangladesh	0.12
Behrman and Wolfe	1984	24-hour recall intake	Income	OLS	Direct	15	Nicaragua	0.06
Bhargava	1991	24-hour recall intake	Income	Random effects, time-varying endogeneity	Direct	120	India[g]	0.05
Wolfe and Behrman	1983	24-hour recall intake	Income	OLS, 2SLS	Direct	15	Nicaragua	0.01

[a]Elasticity calculated to demonstrate upward bias due to aggregation.
[b]Not significantly different from zero at the 5% level.
[c]It was not the main objective of Strauss and Thomas to estimate the elasticity at mean income levels, but to examine elasticities for the lowest income decile which they found to be much higher using non-parametric techniques.
[d]Calories from rice, shelled maize, and fresh cassava only.
[e]The term 'intake' used but no description provided of how calorie data were collected.
[f]The Chernichovsky and Meesook, Timmer and Alderman, and Ravallion studies makes a particularly interesting comparison in that all three use either the 1978 or 1981 SUSENAS surveys, but give a wide range of elasticities, Ravallion suggests that much of the difference between his and the Chernichovsky and Meesook results is due to choice of functional form; Timmer and Alderman use cell means.
[g]Behrman and Deolalikar and Bhargava use the same ICRISAT surveys.
[h]Strauss and Thomas and Williamson-Gray use the same 1974/75 ENDEF surveys; Williamson-Gray uses cell means.

error, and household heterogeneity (again see table 1), which may explain some of the variation in the estimates. Assuming an appropriate instrument is available (an assumption we later scrutinize), instrumental variables (specifically two-stage least squares) may be employed to wash out this simultaneity bias. In addition, panel techniques can be used to account for household heterogeneity. The panel nature of the Philippine data set, and the availability of a range of instrumental variables, makes possible the implementation of estimation techniques such as within (fixed effects) estimation, and panel-based instrumental variable techniques. These econometric considerations are discussed in detail in section 3.

3. Econometric considerations in estimating calorie–income elasticities

Let C and X represent generic terms for calorie consumption intakes and income, respectively. Assume that these are the true values without measurement errors. Two available proxy survey measures for X are TE, household total expenditures, and Y, household current income. Two available survey measures for C are CI (family calorie intakes measured from a 24-hour recall) and CA (household calorie availability as measured from a food expenditure questionnaire). CA needs to be adjusted for various sources of leakages, L, to be a direct measure of C. If this adjustment is not undertaken, the household calorie availability elasticity will usually differ from the family calorie intake elasticity. To see this:

$$CA = CI + L, \tag{1}$$

where L is the sum of *leakages* due to plate waste, loss in cooking and other food preparation, feeding of animals, and feeding nonhousehold members such as guests, hired farm laborers, and servants.

Differentiating (1) with respect to X and expressing the result in terms of elasticities gives

$$\eta_{\mathrm{CA,X}} - \eta_{\mathrm{CI,X}} = \frac{\partial(L/CA)}{\partial X}\,\frac{CA}{CI}\,X. \tag{2}$$

If L is assumed to be a constant absolute value for all households, then $\eta_{\mathrm{CA,X}} - \eta_{\mathrm{CI,X}} < 0$. However, data to be presented below suggests that food eaten by non-family members increases as proportion of total food purchases, so that $\eta_{\mathrm{CA,X}} - \eta_{\mathrm{CI,X}} > 0$ holds empirically. The studies cited in table 1, which used CA as a proxy for C, are mostly silent on the issue of what types of adjustments were made in the food expenditure surveys to account for the various components of L.

In general, we want to estimate

$$C = a + bX + Z'd + v, \tag{3}$$

where Z is a vector of sociodemographic independent variables which we will assume are measured relatively accurately and so, for convenience, can be dropped in the derivations that follow; a is an intercept term; and v is a random error term such that $\mathrm{E}(v)=0$, and $\mathrm{cov}(v_{ij})=0$ for $i \neq j$; $=\sigma_v^2$ for $i=j$.

3.1. Income variables: The standard case of OLS bias towards zero due to random measurement error

Letting e represent an error term in measuring X, such that $\mathrm{E}(e)=0$, and $\mathrm{Var}\,(e)=\sigma_e^2$, the familiar result can be derived that

$$\operatorname{plim} \hat{b}_{\mathrm{OLS}} = b + \frac{-b\sigma_e^2}{\sigma_x^2 + \sigma_e^2 + 2\sigma_{X,e}}. \tag{4}$$

If X is the only explanatory variable measured with error, the direction of the bias on $\hat{b}_{\mathrm{OLS}}$ is unambiguously toward zero. Because more than one explanatory variable may be measured with error, to conclude that $\hat{b}_{\mathrm{OLS}}$ is biased towards zero relies on the assumption that Y and TE are the only variables that are measured with any significant error.

For all estimations presented in this paper, the set of explanatory variables used (in addition to TE or Y) are the number of household members, percentage of total household members falling into various age and gender categories, survey round dummies, population density of the municipality in which the household resides, dummies for mother's and father's years of schooling, father's age, and real barrio rice and corn prices. This group of explanatory variables used in our estimations represents a fairly standard set of Engle curve regressors and the estimated coefficients (reported in Appendix A) on the income variables (TE or Y) are robust in magnitude, sign, and significance to variation of 2SLS exclusion restrictions. These regressors are, by nature, less susceptible to measurement error than the income variables, and exogenous to the household in the short run. Hence, we are confident that the direction of the biases on $\hat{b}_{\mathrm{OLS}}$ due to measurement error will be fairly well defined.

3.2. Commonality of measurement error: OLS bias upwards?

The downward bias on $\hat{b}_{\mathrm{OLS}}$ holds when C is calorie availability or calorie intake, and X represents current income. It is the combination of calorie availability on the left-hand side together with total household expenditure on the right-hand side that results in a departure from this general rule of thumb, because of the *commonality* in error terms between the dependent and independent variables.

In deriving a complete expression for the bias due to this commonality in error terms, it is necessary to consider five separate potential sources of measurement error shown below (e_{TE} is decomposed into the first two sources):[3]

e_{FE} = the error in measuring household food expenditures (FE);
e_{NFE} = the error in measuring household non-food expenditures (NFE);
e_{PG} = the error in proportion of household food expenditures being fed to guests (PG);
e_{PW} = the error in proportion of household food expenditures being fed to workers (PW);
e_{OL} = the error in measuring leakages other than food expenditures going to guests and workers.

The simplification assumption is made that leakages due to meals fed to guests and workers dominate other leakages so that e_{OL} may be ignored. It is necessary to distinguish between meals fed to guests and meals fed to workers because meals fed to workers are a production expense which need to be subtracted from TE.

Letting k equal a multiplicative factor which converts total food expenditures to calories and denoting an observed variable with an asterisk, we can write

$$CA^* - L^* = a + b \cdot TE^* + v^*, \tag{5}$$

where

$$v^* = [k(1 - PG^* - PW^*) - b]e_{FE} - be_{NFE} - CA(e_{PG} + e_{PW})$$
$$+ b(FE)e_{PW} + v.$$

For illustrative purposes, using terms that dominate the magnitude of the bias empirically, we can write

$$\operatorname{plim} \hat{b}_{OLS} = b + \frac{[k(1 - PG - PW) - b]\sigma^2_{eFE}}{\sigma^2_{TE} + \sigma^2_{eFE} + 2\sigma_{TE,eFE}}. \tag{6}$$

The measurement error component of the bias due to OLS estimation of b in (6) will be upwards provided $[k(1 - PG - PW) - b] > 0$. The multiplicative

[3]Full expressions for the potential biases are developed in an appendix available from the authors. This appendix also contains the results of various specification tests; see section 4.2.

factor k represents the calories made available *on average* for each peso spent on *food*, whereas b represents the *marginal* change in family calorie availability for every extra peso of *total expenditure*. Both because (PG^*+PW^*) is relatively low (5.8% at the mean of the data; table 3), but more importantly because households purchase more expensive sources of calories at the margin, we are confident that $[k(1-PG-PW)-b]>0$ for most of our households.[4] The same exercise for the CI, TE pair drops the first term in the brackets of the numerator of (6), so that the direction of the bias is negative as in (4).

3.3. Upward bias due to imperfectly measured leakages

Based on data to be presented later, we expect $\mathrm{cov}(e_{PG}, TE)$, $\mathrm{cov}(e_{PW}, TE)$, $\mathrm{cov}(e_{PG}, Y)$, and $\mathrm{cov}(e_{PW}, Y)$ to be negative, which is an additional source of bias using OLS. This leads to upwardly biased coefficients where the dependent variable is derived from calorie availability, but to downwardly biased coefficients where the total expenditure variable is paired with calorie intake. Intuitively, e_{PG} and e_{PW} are more important determinants of CA^* than is e_{PW} for TE^*, so that this bias is more important for CA, Y than for CI, TE.

3.4. Correcting for measurement errors and taking account of simultaneity using instrumental variables

Instrumental variable (IV) estimation is a potentially useful technique for correcting for contemporaneous correlation between an explanatory variable and the disturbance term. For example, for the CA, TE pair the instrumental variable estimator is unbiased if an instrument, z, can be found such that $\mathrm{cov}(z, e_{FE}) = \mathrm{cov}(z, e_{NFE}) = \mathrm{cov}(z, e_{PG}) = \mathrm{cov}(z, e_{PW}) = \mathrm{cov}(z, v) = 0$, while having a reasonably high correlation with TE^*.

Bearing this in mind, we expect our instrument set to be able to neutralize commonality covariances involving the random error components – e_{FE}, e_{NFE}, and v – but *not* the covariance terms involving the error in measuring leakages due to guests and workers – e_{PG} and e_{PW}. Because the latter two error terms are so strongly linked to income (as will be shown), they very likely will be linked to any otherwise appropriate instrument for income.

[4]Suggestive magnitudes for k and b are given in Bouis and Haddad (1990). Table 7.7. indicates that at mean income levels, an average peso spent on food increases calorie availability by 428 calories (an estimate of k). Econometric estimates of the relationship between calorie intakes and food expenditures presented in table 7.9 indicate that calorie intakes increase by about 90 calories at the margin for each extra peso spent on *food* (evaluated at mean income levels). To derive an empirical estimate of b, this figure of 90 calories still needs to be reduced by the marginal propensity to spend for food out of increases in total expenditures.

Because it is so difficult (if not impossible) to find instruments correlated with X but uncorrelated with e_{PG} and e_{PW}, the IV estimator *may* be more biased than the OLS estimator. To see this, (4) is reformulated for the OLS bias for the $CA-L, Y$ pair.

$$\operatorname{plim} \hat{b}_{OLS} = b + \frac{-b\sigma^2_{eY} - \operatorname{cov}(Y, e_{PG} + e_{PW})}{\sigma^2_Y + \sigma^2_{eY} + 2\sigma_{Y,eY}}. \tag{7}$$

The first term in the numerator of (7) is associated with the familiar bias towards zero of measurement errors in Y. The second term is associated with the guest and worker meals that have not been purged from the dependent variable. These two terms have opposite signs so that it is impossible, a priori, to determine the sign of the overall bias. What is more pertinent for the purposes of this discussion, however, is the fact that if the second term in the numerator dominates, the overall positive bias is *dampened* both by the presence of the first term in the numerator and by the fact that Y is inaccurately measured, which, ceteris paribus, implies a relatively large denominator.

When the 2SLS IV estimator is used: (i) the first term in the numerator of (7) disappears (assuming the instrument has the desired standard properties); (ii) the denominator is replaced by $\operatorname{cov}(Y^*, \hat{Y}^*)$ where $\hat{Y}^*$, the predicted value of Y^*, serves as the instrument; and (iii) the second term in the numerator is replaced by $\operatorname{cov}(\hat{Y}^*, e_{PG} + e_{PW})$. Because $\operatorname{var}(Y^*)$, the denominator of (7), will always be greater than $\operatorname{cov}(Y^*, \hat{Y}^*)$, the denominator for the expression for the bias for the IV estimate will always be smaller.[5] These first two factors, ceteris paribus, will lead to a higher positive bias. It is impossible, however, a priori, to determine the relative magnitudes of $\operatorname{cov}(Y, e_{PG} + e_{PW})$ and $\operatorname{cov}(\hat{Y}^*, e_{PG} + e_{PW})$; depend on the components of the instrument set which combine linearly to create $\hat{Y}^*$.

3.5. *Accounting for unobservable household effects*

Household-specific effects, which are not included as variables in the

[5]This may be demonstrated as shown below. Letting r designate a coefficient of correlation, SE a standard error, and e the standard error term for the first stage of 2SLS, we may write

$$r_{Y^*,\hat{Y}^*} = \frac{\Sigma(Y^* - \bar{Y})(\hat{Y}^* - \bar{Y})}{SE_{Y^*} \cdot SE_{\hat{Y}^*}} = \frac{A}{B} < 1,$$

$$r_{Y^*,Y^*} = \frac{\Sigma(Y^* - \bar{Y})(Y^* - \bar{Y})}{SE_{Y^*} \cdot SE_{Y^*}} = \frac{C}{D} = 1.$$

Because $\operatorname{var}(Y^*) = \operatorname{var}(\hat{Y}^*) + \operatorname{var}(e)$, therefore $SE_{Y^*} > SE_{\hat{Y}^*}$. It follows, then, that $A < B < D = C$, so that $A < C$.

regression estimations, but which affect the demand for calories, may be correlated with both X (TE or Y) and its associated measurement error (e_{TE} or e_Y). It is difficult to come up with specific (unmeasurable) variables that are intuitively appealing, so that it is difficult to speculate as to the direction of this potential bias. The unobserved household-specific effects can be represented as an additional error component which is time-invariant.

Under the hypothesis that our measure of income is exogenous, a benchmark panel estimator that is consistent irrespective of whether or not the unobserved effects are correlated with included right-hand-side variables is the 'within' estimator which regresses $(C_{it}-\bar{C}_i)$ on $(X_{it}-\bar{X}_i)$, and thus removes the unobserved effects. However, in the presence of other forms of simultaneity bias (due to common measurement error or behavioral endogeneity), the within estimator is biased. Moreover, when $\mathrm{cov}(X,e_x)$ is not equal to 0 (see section 3.3), the within estimator is possibly more biased than the OLS estimator.[6]

Hausman and Taylor (1981) have developed a random-effects estimator for use with panel data that treats unobserved effects as randomly distributed across households. Unlike conventional random-effects GLS estimators, however, the Hausman–Taylor (HT) estimator uses additional information concerning a priori assumptions on which included right-hand-side variables will be correlated (endogenous) and uncorrelated (exogenous) with the unobserved effects. If these assumptions are correct (and they can be tested if a consistent benchmark exists, e.g., a 'within' estimate under the null hypothesis of $\mathrm{cov}(X,v^*)$ equal to zero) the HT estimator is more efficient than the within estimator. Appealing features of the HT estimator are that: (i) if the restrictions are correct it deals with potentially endogenous, time-

[6]Using differences from the mean for all four rounds ($T=4$) gives the 'within' estimator, where for the standard measurement error case we have

$$\mathrm{plim}\,\hat{b}_W = b + \frac{T-1}{T}\,\frac{-b\sigma_e^2}{\mathrm{var}(X_t^*-\bar{X}^*)}.$$

Comparing the denominator in the above expression with that of (4) in the text allows us to state that if

$$(T)\mathrm{var}(X_t^*-\bar{X}^*)<(T-1)(\sigma_X^2+\sigma_e^2+2\mathrm{cov}(X,e)),$$

or

$$\sigma_X^2+\sigma_e^2<(4T-2)\mathrm{cov}(X,e),$$

then the downward bias due to measurement error is larger, ceteris paribus, for $\hat{b}_W$ than for $\hat{b}_{OLS}$ [Hsiao (1986)].

Analogous to the discussion for the $CA-L, Y$ pair for IV estimation [see (7) in the text], it is unclear, a priori, whether the direction of the bias will be negative (random measurement error predominates) or positive (the presence of the non-purged guest and worker meals which are correlated with X predominates). In addition, for the within estimate it is also unclear, a priori, whether whatever bias may be present will be magnified (the above conditions hold) as is certain with the IV estimates. Until the empirical estimates are presented, it is enough for now to establish that there is the potential for a magnified upward bias.

invariant unobserved household effects *as well as* the time-varying endogeneity of X; and (ii) allows time-invariant explanatory variables (which necessarily are omitted from the 'within' estimations) to be included in the regression estimations. However, if the restrictions are incorrect the HT estimator is inconsistent.

3.6. Summary of expectations of direction of bias by source

Table 2 summarizes our expectations as to the direction of biases due to the individual effects discussed above for various calorie–income variable pairs and estimating techniques. The empirical evidence presented below suggests that the downward bias due to e_{PW} for the CI, TE pair is quite negligible. Thus, restricting ourselves to the specific sources of bias discussed above, table 2 suggests that with enough external instruments the Hausman–Taylor estimate for the CI, TE pair will provide the least biased estimate. However, the empirical evidence will also suggest that biases due to random measurement error in TE, the endogeneity of X, and unobserved household effects are relatively small, so that all the estimates involving CI, X pairs fall within a fairly narrow range, with the exception of the OLS estimate for the CI, Y pair for which the random measurement error on Y is apparently a problem.

4. Empirical estimations

While the discussion in the previous section indicates the direction of various biases which may be problems in theory, the primary objective of this paper is to show that such biases may be important enough empirically that they can lead to incorrect policy conclusions under a wide range of circumstances. The Philippine data set is somewhat unique in that all four proxy variables discussed above are available for the same population, in addition to information on food expenditures for guests and hired farm workers, so that some attempt may be made to measure the magnitudes of these biases.

4.1. Description of the Philippine data set

The data to be used for the regressions estimations are taken from surveys of rural households residing in Bukidon province in the Philippines. Households were surveyed four times at four-month intervals and data were collected on a wide range of topics, including landholdings, income sources,

Table 2

Summary of directions of possible estimated bias by calorie–income variable pair, by source of bias, and by estimating technique.

Calorie–income variable pair/ Source of bias	Estimating technique OLS	IV (2SLS)	Within	Panel HT[g]
CA − L, TE				
Random ME[a]	$\left.\begin{matrix}-\\+\end{matrix}\right\}+$		$\left.\begin{matrix}-\\+\end{matrix}\right\}+$	
Common ME[b]				
Food to G & W (LHS)[c]	+	+	+	+
IV Magnfd bias G & W[d]		+		+
X Endogenous[e]	?		?	
Unobserved *HE*[f]	?	?		
CI, TE				
Random ME[a]	−			
Common ME[b]				
Food to G & W (RHS)[c]	−	−	−	−
IV Magnfd bias G & W[d]				
X Endogenous[e]	?		?	
Unobserved *HE*[f]	?	?		
CA − L, Y				
Random ME[a]	−		Not	
Common ME[b]			estimated	
Food to G & W (LHS)[c]	+	+	using	
IV Magnfd bias G & W[d]		+	panel	
X Endogenous[e]	?		techniques	
Unobserved *HE*[f]	?	?		
CI, Y				
Random ME[a]	−		Not	
Common ME[b]			estimated	
Food to G & W[c]			using	
IV Magnfd bias G & W[d]			panel	
X Endogenous[e]	?		techniques	
Unobserved *HE*[f]	?	?		

[a]Random measurement error (section 3.1); it is expected that this will be a more serious problem for *Y* than for *TE* using OLS.

[b]Common measurement error (section 3.2); it is possible that the upward bias due to cov(e_{CA}, e_{FE}) will outweigh the downward bias due to the random error in measuring *NFE*; if this is the case, then the bias may be exacerbated by use of the within technique as compared with OLS.

[c]Transfers of food expenditures to guests and workers which have not been purged from $CA^* - L^*$ (section 3.3); it is expected that this will be a more serious problem for *CA*, *X* pairs (LHS) than for *C*, *TE* pairs (RHS).

[d]Instrumental variables may more strongly 'reveal' the potential bias due to expenditures for guests and workers which have not been purged from $CA^* - L^*$ than uninstrumented *X* in OLS which is measured with error (section 3.4); this is a particular problem for the $CA - L, Y$ pair, although IV will magnify any biases present for the remaining three pairs as well, in particular for the $CA - L, TE$ pair.

[e]*C* and *X* may be determined simultaneously (section 3.4).

[f]Unobserved household effects (section 3.5).

[g]Biases assuming that categorization of regressors into those correlated and not correlated with unobserved household effects is correct.

expenditure patterns, calorie intakes, and heights and weights [see Bouis and Haddad (1990) for a more detailed description of how the data were collected]. Our analysis here uses data for 406 households, which were present for all survey rounds and whose livelihood depended primarily on production of either corn or sugarcane.

Table 3 presents data on a per capita basis for current income, total expenditures, and total food expenditures (the estimated food budget shares for food eaten by family members is also shown), and on a per adult equivalent basis for family calorie availability and family calorie intake. The calorie availability data have already been adjusted for meals served to guests and hired farm laborers reported by our respondent households.

Respondents were asked to recall food expenditures (by item from each of five possible sources: purchases, own-production, in-kind wages, gifts, borrowed) for the past month and non-food expenditures (by item) for the past four months, using a variable time period recall method for items purchased more frequently than once a month (for food expenditures) or once every four months (for non-food expenditures). The food expenditure questionnaire also asked for quantity and/or price data by item, depending on the source of acquisition. A simple average of real corn and rice prices across individual households in each of twenty-two barrios in each survey round was computed for use in the regression estimations.

Calorie availability is derived from the expenditure data for fifty food categories using aggregate conversion factors constructed from actual food intake patterns of the component foods within each category (i.e., these aggregate conversion factors were computed from information available from the 24-hour recall of food intakes). Total *household* available calories were reduced by a proportional factor, $(n/nf+n)$, to give *family* calorie availability, where n is the number of family members reported for a particular household (expressed in adult equivalents), and nf is the number of non-family members reported (expressed in adult equivalents; all guests and workers were assumed to be adults) computed as a function of the number of meals that respondents estimated that they served to guests and hired laborers over a period of one month.

Family calorie intakes are derived from a 24-hour recall by the wife of foods consumed by individual household members, for which 120 individual food commodities were identified. The survey technique involved asking mothers (with the assitance of props such as pictures showing various sized vegetables/fruits, and various sized containers; the capacities of pots and other household utensils were measured as necessary) about amounts of ingredients of each recipe for each meal. She was then asked to estimate how much of each recipe each household member consumed. This not only gave a measure of intra-household distribution of food, but more importantly for the purposes of this paper, gave a measure of how much of each recipe was

Table 3

Per capita income, total expenditures, and total food expenditures, family shares of food budget, family calorie availability and intakes, and per capita food expenditures for guests and hired workers, by income quintile, total expenditure quintile, and by crop-tenancy group.

	Per capita income (pesos/week)[c]	Per capita total expenditures (pesos/week)[c]	Per capita total food expenditures (pesos/week)[c]	Family share of food budget (%)	Family calorie availability[a] per day	Family calorie intake[b] per day	Per capita food expenditures for non-family members (pesos/week)[c]: Guests[e]	Per capita food expenditures for non-family members (pesos/week)[c]: Hired workers[f]
Income quintile[d]								
1	13.1	30.0	24.6	97.6	2,170	2,266	0.59	0.04
2	21.9	36.6	27.4	96.9	2,237	2,313	0.84	0.17
3	29.8	39.7	29.8	95.4	2,321	2,336	1.37	0.10
4	41.4	48.1	35.0	95.2	2,639	2,433	1.68	0.26
5	101.7	76.2	47.1	91.7	2,826	2,443	3.90	0.65
All	41.7	46.3	32.8	94.9	2,439	2,358	1.68	0.25
Expenditure quintile[d]								
1	21.9	21.8	17.6	98.0	1,790	2,107	0.35	0.01
2	25.4	29.8	24.0	97.4	2,143	2,288	0.62	0.04
3	28.5	38.0	29.6	97.3	2,411	2,384	0.81	0.06
4	45.8	50.0	37.2	93.6	2,666	2,439	2.37	0.19
5	87.6	91.9	55.9	92.3	3,193	2,575	4.28	0.93
All	41.7	46.3	32.8	94.9	2,439	2,358	1.68	0.25
Crop-tenancy group								
Corn	35.2	41.4	31.0	94.7	2,375	2,372	1.63	0.18
Corn owners	47.7	49.2	35.3	92.7	2,445	2,387	2.59	0.35
Corn owners/tenants	46.4	46.6	32.3	94.2	2,368	2,329	1.86	0.26
Corn share tenants	28.4	40.0	30.9	95.6	2,405	2,412	1.35	0.14
Corn laborers	26.6	32.3	26.0	96.0	2,266	2,326	1.05	0.04
Sugar	52.2	53.5	35.6	94.8	2,534	2,343	1.85	0.36
Sugar owners	70.1	64.4	41.8	94.0	2,655	2,386	2.50	0.45
Sugar owners/renters	83.3	89.9	50.8	93.4	3,148	2,447	3.37	1.17
Sugar renters	43.0	43.5	31.7	94.7	2,350	2,371	1.67	0.08
Sugar laborers	26.5	30.8	24.7	97.5	2,208	2,237	0.62	0.00

Source: International Food Policy Research Institute–Research Institute for Mindanao Culture survey, 1984/85.

[a] Per adult-equivalent, derived from food expenditures.

[b] Per adult-equivalent, derived from 24-hour recall of food consumed.

[c] 1984 pesos.

[d] Quintile 1 is the lowest rank and 5 the highest.

[e] A component of non-food expenditures, but not family calorie availability.

[f] A production expense; neither a component of food or non-food expenditures nor of family calorie availability.

not consumed which could be subtracted from the total prepared. If a family member was not present for any meal during the previous 24-hours he/she simply was not included in the calculation of household intakes. If a meal was consumed by an individual household member away from home, the wife was asked to provide an estimate of what was eaten. The survey area is a relatively remote, undeveloped region with minimal opportunities for purchases of prepared meals. These meals, then, tended to be taken as guests in a neighbor's or relative's home, or as partial payment for work performed on another farm.

An extremely involved calculation utilizing various components of the questionnaire gave current income, which included profit from own-farm production (values were imputed to food consumed out of own production), off-farm employment, money earned from small businesses, and miscellaneous transfers.

For the sample as a whole, the difference between family calorie intake and family calorie availability is between 3 and 4%. However, disagregating by expenditure quintile, there is a clear pattern for family calorie availability to be well below family calorie intake for low-income households and for family calorie availability to be well above calorie intake for high-income households. For six out of a total of eight crop-tenancy groups, for all four corn groups and for the two poorest sugar groups, the values for the two variables are within three percent of each other. However, for the two wealthiest sugar groups, family calorie availability is substantially greater than calorie intake. This may be due to an understatement of meals fed to hired laborers, although the information provided in table 3 on the relative magnitudes of meals served to guests and workers suggests the possibility that the guest meals may be the more important phenomenon. The understatement of meals served to non-household members occurred even though recorded meals for guests and laborers are already highest for these two high income sugar groups. The discrepancy in the family calorie availability and calorie intake figures for these two sugar groups was particularly pronounced for the third round surveys, which were undertaken at the height of the sugar harvest.

The differences in calorie consumption patterns between calorie availability and calorie intakes across income groups in table 3 are consistent with (indeed, motivated investigation of) the potential biases summarized in table 2. Random overstatements (understatements) of food purchases tend incorrectly to place households in too high (too low) an income bracket when total expenditures are used as a stratifying variable. Large positive (negative) 'errors' in measuring calorie availability (deviations from actual family nutrient intake) are concentrated at the upper (lower) end of the income distribution. The range of average consumption from first to fifth quintile is narrower for both calorie intakes and availability when using income as a

Table 4

Calorie–income elasticity estimates, by calorie-income variable pair and by estimating technique.[a]

Calorie–income variable pair		Estimating technique: OLS (1)	IV(2SLS) (2)	Within (no IV) (3)	Panel HT (4)
(1)[b]	$CA-L, TE$	0.43 (0.02)	0.32 (0.03)	0.59 (0.02)	0.55 (0.03)
(2)[b]	CI, TE	0.12 (0.01)	0.08 (0.03)	0.14 (0.03)	0.09 (0.03)
(3)[b]	$CA-L, Y$	0.11 (0.01)	0.28 (0.03)	-	-
(4)[b]	CI, Y	0.03 (0.01)	0.09 (0.02)	-	-
(5)[c,d]	CA, XTE	0.52 (0.02)	0.46 (0.04)	0.65 (0.03)	0.62 (0.04)
(6)[c]	CI, XTE	0.12 (0.01)	0.08 (0.02)	0.14 (0.03)	0.09 (0.03)
(7)[d]	CA, Y	0.17 (0.02)	0.40 (0.04)	-	-

[a]Standard errors for elasticity estimates are reported in parentheses.
[b]Detailed results for rows (1)–(4) are presented in Appendix A.
[c]XTE equals TE plus food expenditures for surveyed worker meals.
[d]Calories from surveyed guest and worker meals have not been subtracted from calorie availability.

stratifying variable as compared with total expenditures since incomes are measured less precisely.

4.2. Estimates of calorie–income elasticities

Table 4 presents a seven-by-four matrix of estimated calorie–income elasticities. The four columns are for the four estimators used. This first four rows correspond to the four combinations possible for calorie availability, calorie intake, total expenditures, and current income as defined above. In view of the fact that data may not have been available for other studies to make the downward adjustment for meals eaten by guests and hired workers, for purposes of comparison, two new variables were computed that added back the guest and worker calories to calorie availability to give CA (as opposed to $CA-L$), and food expenditures for worker meals to total expenditures to give XTE (expenditures for guest meals are already included as a 'non-food' item in total expenditures). This resulted in three new combinations, which correspond to the last three rows in table 4.

A semi-logarithmic specification was used in the results presented in table

4. Various alternative specifications were tried; the elasticity estimates at mean income varied little between these specifications (see also footnote 3). Full estimation results are reported in Appendix A at the end of the paper.

4.2.1. Ordinary least squares estimates

Given the patterns of calorie consumption across income groups shown in table 3, the relative magnitudes of the estimates in the first four rows of column one should come as no surprise. As predicted in table 2, the CI, Y pair yields the lowest estimate which is biased downward (for the moment setting aside the issue of correcting for simultaneity and household unobserved effects) because of the difficulty of measuring Y. The $CA-L, TE$ pair gives the highest estimate because of the bias introduced by non-purged guest and worker meals and random, but common measurement errors between CA and TE. The $CA-L, Y$ pair avoids the problem of common measurement errors, but combines the inaccurately measured income problem and the non-purged guest and worker meal problem, which work in opposite directions in terms of bias.

It is useful at this point to establish the estimate for the remaining CI, TE pair as something of a benchmark to facilitate discussion of comparisons with estimates in the remaining three columns and last three rows. It is only coincidental that the estimates for the $CA-L, Y$ and CI, TE pairs are nearly identical. For CI, TE, table 2 indicates two sources of downward bias, random measurement error on TE and inclusion of worker meal expenditures in TE. A comparison of row two and row six, in which *observed worker meal expenditures* have been added to TE, shows that this latter source of potential bias can be ignored empirically; the calorie–income elasticity estimate remains virtually unchanged. By contrast, note from rows five and seven that inclusion of observed guest and worker meals in calorie availability results in significant increases in the estimated elasticities, indicating that this is an important source of bias when introduced to the dependent variable.

Thus, the remaining source of bias for CI, TE (again abstracting from the effects of the endogeneity of X and unobserved household effects) is only the random measurement error in TE, which is small relative to the bias due to the random measurement error in Y. This can be seen by comparing the estimates in rows two and four. A comparison of rows two and one, then, shows the substantial upward bias due to the combined effects of common measurement error and the non-purged guest and worker meals.

4.2.2. Instrumental variable estimates

Turning to the IV estimates, table 2 suggests that the row two and row four estimates will be similar, since the right-hand-side bias of non-purged worker meals is no longer an empirical consideration in row two. In fact, in table 4 these two IV estimates are nearly identical. For both rows, the IV estimates correct for measurement error and endogeneity of X as compared with OLS. In row two, the correction for simultaneity apparently dominates the correction for random measurement error. If the bias had been due to measurement error alone, the estimate should have increased. This suggests that the direction of the bias in OLS due to endogeneity is positive. By contrast, in row four the estimate increases for instrumental variables, indicating that the correction for measurement error dominates.

For the $CA-L, Y$ pair the IV estimate increases substantially for two reasons. First, the correction for random measurement error in Y, net of the correction for endogeneity of X, is positive as suggested by the row four results. Second and unfortunately, the potential bias due to non-purged guest and worker meals, which was masked to some extent by the measurement error in Y for OLS, is now exacerbated by use of the IV estimator as outlined in section 3.4.

For the $CA-L, TE$ pair, the IV estimate declines as compared with OLS. This is because the elasticity is now purged of the bias due to commonality in (random) measurement error, which apparently dominates the tendency of the IV estimator to exacerbate (as compared with OLS) the bias due to non-purged guest and worker meals.[7] Table 2 suggests that the row one and row three IV estimates will be similar; and they are closer together as compared with OLS. The remaining difference in column two can be attributed to the commonality in measurement error caused by the (non-random) non-purged guest and worker meals in row one.

Throughout the paper, we have assumed that errors in measuring calorie intakes are orthogonal to income, leading us to conclude that the CI, TE pair (row two in table 4) gives the most reasonable estimates of the calorie–income relationship. Our maximum estimate of the calorie–income elasticity for the CI, TE pair across all estimation methods even evaluated at mean values for the lowest expenditure quintile is only 0.16. Does our range of estimates for the CI, TE pair across estimation techniques as shown in table 4 (0.08 to 0.14) represent a lower bound which substantially understates the true elasticity? For example, do low-income groups persistently over-estimate food intakes and/or high-income groups persistently underestimate food intakes [households reporting idealized or stylized diets; Lipton (1983)],

[7]2SLS estimations were undertaken for all four variable pairs, but excluding all sugar-growing households with access to land from the sample ($n=1216$). Calorie–income elasticities for the CI, X pairs remained virtually unchanged from those reported in table 4. For the $CA-L, TE$ pair the elasticity declined from 0.32 to 0.28 and for the $CA-L, Y$ pair the elasticity declined from 0.28 to 0.22. The lower elasticities for $CA-L, X$ pairs are due to the exclusion from the sample of households for which underestimation of guest and worker meals is apparently the most serious. However, the estimates remain upwardly biased since the problem of under-estimated food leakages remains (although to a lesser degree) for the corn households with access to land and landless laborer households.

leading to downwardly biased estimates? What evidence can we offer that $\mathrm{cov}(X, e_{CI})$ is close to zero?

Analysis presented by Bliss and Stern (1978) on the functional relationship between calorie intakes, energy expenditures, and body weight, can be used to show that apart from extra calories needed for work and other strenuous activities, calorie requirements increase less than proportionately with body weight. Metabolic rates may vary between two individuals, with the consequence that energy requirments may not be roughly proportional to weights for these two individuals (holding activity levels constant and assuming that each individual is maintaining a constant weight). However, comparing group averages, these individual-specific differences in metabolism may be ignored, assuming that they are randomly distributed.

For our survey population, adult weights remained nearly constant over the twelve months of surveys *within* expenditure quintiles. Furthermore, adult energy expenditures (activity levels) were somewhat lower for the highest expenditure quintile as compared with the lowest expenditure quintile. The ratio of calorie intakes between adults in high and low income groups, then, should increase somewhat less than the ratio of their weights. If percentage increases in adult calorie intakes do not correspond to percentage increases in adult weights, then we have some empirical basis for claiming that $\mathrm{cov}(X, e_{CI})$ is substantially greater than or substantially less than zero.

Adult females (who were never pregnant or breastfeeding during the four survey rounds) in the highest expenditure quintile consume 13% more calories and weigh 12% more than their counterparts in the lowest expenditure quintile. The same comparison for adult males gives a calorie intake increase of 16% and a weight increase of 12%. Because activity patterns on average are apparently somewhat higher for the lowest expenditure quintile adults than for the highest expenditure quintile adults, a smaller difference in calorie intakes than the 12% difference in body weights might have been expected, as opposed to the observed difference of 13–16%.

This evidence would suggest, then, that if our row two (table 4) estimates are biased at all due to deficiencies in the methodology used to recall calorie intakes, they are biased upwards; these represent *upper bound* estimates. Nevertheless, the discrepancies in the ratios of intakes and weights are reasonably small.

A simple example using this indirect calculation of calorie-income elasticities demonstrates the implausibility of 'high' calorie–income elasticities shown in table 1 under a reasonable set of assumptions. Assume: (1) stable weights over time *within* expenditure quintiles; (2) constant activity patterns *across* expenditure quintiles; and (3) an *average* per adult equivalent calorie–income elasticity of 0.2 across expenditure quintiles ranging from $50 per capita to $300 per capita. Under these assumptions, one would be forced to conclude that the average person in the highest expenditure quintile weighs roughly twice as much as the average person in the lowest expenditure quintile, which of course is highly improbable.

4.2.3. Panel estimates

The panel allows us to correct our row one and row two estimates for household unobserved effects. In the case of the row one OLS and IV estimates, because they are already substantially biased for other reasons, this hardly seems worthwhile. However, that the row one (and row five) panel estimates are even higher than the OLS and IV estimates, is consistent with the observation made in section 3.5 that when there is commonality in measurement error, this bias may be exacerbated by use of panel techniques.[8]

4.2.4. Effects of other explanatory variables

The detailed regression results for rows one through four are reported in Appendix A. The estimated coefficients on the survey round dummy variables suggest quite different patterns of seasonal consumption of calories. The CA, X regressions indicate that calorie consumption is highest in round 3, while the CI, X regressions suggest that calorie consumption is lowest in round 3 and highest in round 1. As already indicated, the discrepancy in the calorie availability and intake data is particularly pronounced for the third round surveys which were undertaken at the height of the sugar harvest when bias due to non-purged worker meals would be a particular problem. As with the income elasticities for the CA, X pairs, the estimated seasonal pattern of consumption likewise should be considered incorrect.

Seasonal fluctuations in calorie intakes with respect to the CI, X pairs appears to be the result of a combination of a trend decline in economic

[8]A consequence of the fact that $\operatorname{cov}(X, v^*) \neq 0$, is that the 'within' estimate cannot be used as a consistent benchmark with which to test the over-identifying restrictions necessary for consistent and efficient Hausman–Taylor (HT) estimates. This is important because our HT estimates are based on an ad hoc classification of explanatory variables where:

X_1 = a vector of time-varying variables not correlated with unobserved household effects: the barrio prices of corn and rice ($k_1 = 2$);

X_2 = a vector of time-varying variables correlated with unobserved household effects: log of total expenditure per capita ($k_2 = 1$);

Z_1 = a vector of time-invariant variables not correlated with unobserved household effects: municipal population density, seven demographic variables ($g_1 = 8$);

Z_2 = a vector of time invariant variables correlated with unobserved household effects: household size ($g_2 = 1$).

For identification of an efficient estimator, it is necessary that $k_1 > g_2$. Intuitively, the HT estimator employs the means of the time-varying exogenous variables (X_1) as instruments for the time-invariant endogenous variables (Z_2). Because we cannot test the ad hoc classification of explanatory variables, we have no reason to believe the HT estimates represent an improvement upon the 'within' estimates; they may even be worse. Griliches and Hausman (1986) do outline a method for recovering consistent fixed effects estimates in the presence of measurement error, but not, as far as we are aware, in the presence of behavioral endogeneity.

activity, normal intra-year fluctuations in energy expenditures, and changes in the real prices of non-staple foods between rounds, which is much too complex to unravel here. This does, however, raise the issue of price effects of the two main staple foods, corn and rice. The fact that calorie–income elasticities are close to zero, combined with the reasonable assumption that consumers are aware of fluctuations in their calorie intakes, suggests that consumers view calorie consumption as an important necessity. If this is the case, it is illogical to expect that *total* calorie consumption will be responsive to changes in the prices of inexpensive sources of calories. If the price of corn goes up, for example, and poor consumers are willing consciously to trade substantial amounts of calories at the margin for other food characteristics or non-foods, then they should also be interested in acquiring substantial amounts of calories at the margin as their income increases.

In accordance with above view, the coefficients on the prices of corn and rice are not significantly different from zero for the CI, X regressions. However, in some of the CA, X regressions these price coefficients are significant (and negative). Again, this result may be due to negative correlations between staple food prices and non-purged guest and worker meals.

5. Conclusions

Our estimates add to the growing evidence that calorie–income elasticities are much lower than previously thought. Initial reaction to the low estimates was to reject the previous policy conclusion that increasing income alone could improve calorie intakes [e.g., Behrman and Deolalikar (1987)], with the implication that more direct government intervention would be necessary to improve nutrition in the short-to-medium run. More recently, Ravallion (1990) and Strauss and Thomas (1989) have presented empirical evidence to show that the calorie response to income of the lowest income households can be considerably higher than that of households at average income levels. Ravallion has made the further point that relatively small fluctuations in income could move high percentages of households above or below critical nutrition threshold points, and thus could have a major nutritional impact. Increases in income alone may be very important after all for the nutritional well-being of very low-income households.

While these types of issues can be investigated with our data set, the empirical evidence presented in this paper does not address their arguments directly. Having said that, however, we do think that the policy debate has focussed too much on elasticities – percentage changes in income – and not enough on absolute changes in income. Very large percentage increases in income for the very poor do very little to change their income status, while relatively small changes in absolute income can have a very large effect. For the bottom 40% of our sample population, a $175 increases in per capital annual total expenditure (a 300% increase in income for the lowest expenditure quintile) would have a very substantial effect in raising the (presently deficient) intakes of calories and other important nutrients. At the household level, average intakes of most nutrients for our survey population reach recommended levels at a total expenditure of about $250 per capita per year. Considering the very wide rage that absolute incomes can take, adequacy is achieved at a relatively low level.

Other analysis suggests that better nutrient intakes alone will not solve a serious morbidity problem in the survey area [see Bouis and Haddad (1990)], so that we are reticent to state that income alone will improve health and anthropometry, which are often implicitly considered important aspects of 'nutrition'. Nevertheless, our results suggest that a crucial policy issue with respect to achieving nutrient adequacy is the extent to which the poor in the Philippines will share in future absolute increments in real per capita gross national product. If the economy experiences sustained economic growth and the share of the poor in this growth is disproportionately large, then income alone can contribute significantly to the nutrient adequacy problem.[9] We do not mean to imply by this that the present economic situation or policies give us any reason to be optimistic that sustained growth can be realized or that increments in income would be so distributed.

Our paper, however, has focussed primarily on methodology. In estimating the calorie–income relationship for a particular population, we have shown that the selection of calorie availability as dependent variable leads to seriously upwardly biased elasticity estimates.[10] The extent to which consumption based on household food purchases diverges from family nutrient intakes and is a potential source of estimation bias will, of course, vary by economic and cultural setting. For example, for urban populations for which meals provided for hired workers are presumably not an important phenomenon, it may (or may not) be possible to identify instruments which are not correlated with unobserved leakages and so to purge calorie–income elasticity estimates of the potential biases outlined in this paper.

Nevertheless, our Philippine data demonstrate that the divergence between calorie availability and calorie intakes can be empirically important for rural populations and that care should be taken to accurately measure this phenomenon in food expenditure surveys. Our results suggest that collection

[9]Nationwide surveys (NCSB, 1988) indicate that households in the bottom 40% of the income distribution earn 15% of total income. If the economy were to double in per capita terms (a 3.5% per capita growth rate over twenty years), and if each person's income were to increase by the same absolute amount, this implies a 267% increase in the incomes of the bottom 40% of the population, or the $175 in total expenditures referred to in the previous paragraph in the text.

[10]Caution also may be warranted in deriving estimates of demand for individual foods from food expenditure survey data. For example, these expenditure data generate much higher staple food income elasticities for the Bukidnon population than do the 24-hour food recall data.

of calorie intake data should be seriously considered for future studies of the calorie–income relationship and of demand relationships for individual foods. This paper leaves unresolved the practical queston ot whether it will be more cost-effective to collect such intake information, or to conduct food expenditure surveys which undertake a careful accounting of food purchased but not eaten by household members.

Appendix A: Detailed regression results

Table A.1

Calorie availabilty–total expenditure estimations.[a]

Variable	Estimating technique: OLS Coeff.	OLS t-stat	2SLS Coeff.	2SLS t-stat.	Fixed effects Coeff.	Fixed effects t-stat.	Hausman–Taylor Coeff.	Hausman–Taylor t-stat.
Intercept	241.67	0.86	1,059.31	3.12			−745.12	−1.88
LNTEPC	752.46	24.77	570.08	10.02	1,029.81	24.68	972.57	16.87
HHTOT	−41.98	−6.20	−48.75	−6.89			−33.81	−3.81
DWED	27.97	0.95	68.66	2.21			−21.14	−0.58
DFED	−76.11	−2.74	−34.22	−1.06			−126.67	−3.59
FATAGE	0.08	0.45	0.23	1.25			−0.10	−0.52
DEMM05	−859.74	−5.00	−996.98	−5.51			−694.10	−3.27
DEMF05	−1,042.31	−5.62	−1,181.80	−6.15			−873.97	−3.90
DEMM511	−572.53	−3.46	−659.10	−3.79			−468.06	−2.33
DEMF511	−626.56	−3.73	−771.48	−4.39			−451.67	−2.27
DEMM1117	−953.49	−4.51	−966.26	−4.49			−938.08	−3.96
DEMF1117	−619.69	−2.85	−589.45	−2.70			−656.18	−2.64
DEMFGT17	−910.66	−3.56	−810.28	−3.09			−1,031.82	−3.43
RD1	70.24	1.97	80.79	2.26			57.50	1.44
RD2	122.78	2.61	92.11	1.91			159.79	2.91
RD3	267.52	4.33	276.59	4.37			256.56	3.40
BRPRCORN	−85.69	−2.18	−69.63	−1.73	46.96	2.43	−105.08	−2.26
BRPRRICE	−13.23	−0.37	−58.48	−1.57	−18.73	−0.64	41.38	0.96
POPDEN	0.44	1.63	0.42	1.57			0.46	1.48
Adjusted R^2	0.47		0.45		0.58		0.29	
n	1,624		1,624		1,624		1,624	

[a]Heteroscedasticity-consistent standard errors; see table A.6 for definitions of variables.

Table A.2

Calorie intake–total expenditure estimations.[a]

	Estimating technique							
	OLS		2SLS		Fixed effects		Hausman–Taylor	
Variable	Coeff.	*t*-stat	Coeff.	*t*-stat.	Coeff.	*t*-stat.	Coeff.	*t*-stat.
Intercept	1,658.88	5.36	1,927.62	5.41			1,865.27	5.05
LNTEPC	202.78	8.06	142.83	3.32	236.31	4.87	156.74	3.14
HHTOT	−24.99	−3.80	−27.22	−4.01			−26.70	−3.82
DWED	−11.58	−0.28	1.79	0.04			−1.31	−0.03
DFED	−58.72	−1.72	−44.95	−1.27			−48.14	−1.31
FATAGE	−0.20	−1.04	−0.15	−0.77			−0.16	−0.80
DEMM05	−1,034.45	−5.06	−1,079.56	−5.22			−1,069.09	−5.09
DEMF05	−1,056.80	−5.03	−1,102.64	−5.21			−1,092.01	−5.06
DEMM511	−664.26	−3.41	−692.71	−3.53			−686.11	−3.44
DEMF511	−825.37	−4.17	−872.99	−4.35			−861.95	−4.17
DEMM1117	−227.00	−0.93	−231.20	−0.95			−230.23	−0.93
DEMF1117	−545.07	−2.45	−535.13	−2.41			−537.43	−2.39
DEMFGT17	−779.79	−2.75	−746.79	−2.63			−754.45	−2.63
RD1	232.23	5.73	235.70	5.82			234.90	5.76
RD2	54.67	1.03	44.59	0.83			46.92	0.87
RD3	−149.85	−2.12	−146.86	−2.07			−147.55	−2.04
BRPRCORN	32.47	0.71	37.75	0.82	−65.02	−2.89	36.52	0.78
BRPRRICE	54.36	1.41	39.49	0.99	−88.91	−2.61	42.94	1.07
POPDEN	−1.47	−4.86	−1.48	−4.86			−1.48	−4.78
Adjusted R^2	0.13		0.12		0.24		0.10	
n	1,624		1,624		1,624		1,624	

[a]Heteroscedasticity-consistent standard errors; see table A.6 for definitions of variables.

Table A.3

Calorie availability–income and calorie intake–income estimations[a]

	C−X variable pair and estimating technique							
	Calorie availability–income				Calorie intake–income			
	OLS		2SLS		OLS		2SLS	
Variable	Coeff.	*t*-stat	Coeff.	*t*-stat.	Coeff.	*t*-stat.	Coeff.	*t*-stat.
Intercept	2,647.79	7.48	1,225.30	2.94	2,297.49	7.23	1,815.42	5.04
LNYPC	198.65	7.89	490.78	9.13	55.55	2.90	154.55	4.06
HHTOT	−68.25	−7.91	−65.81	−7.71	−32.06	−4.77	−31.23	−4.64
DWED	161.60	4.09	111.25	2.66	24.08	0.57	7.02	0.17
DFED	28.63	0.76	−71.52	−1.67	−31.18	−0.88	−65.12	−1.75
FATAGE	0.45	2.06	0.10	0.43	−0.10	−0.50	−0.22	−1.08
DEMM05	−1,317.06	−5.64	−1,156.88	−4.78	−1,156.59	−5.60	−1,102.30	−5.24
DEMF05	−1,496.27	−6.20	−1,317.55	−5.20	−1,177.90	−5.52	−1,117.34	−5.16
DEMM511	−870.25	−3.93	−782.85	−3.37	−743.88	−3.73	−714.26	−3.54
DEMF511	−1,017.77	−4.74	−713.82	−3.06	−928.69	−4.60	−825.69	−4.00
DEMM1117	−990.67	−3.68	−967.84	−3.40	−236.86	−0.96	−229.13	−0.92
DEMF1117	−573.82	−2.12	−689.83	−2.39	−533.50	−2.35	−572.82	−2.47
DEMFGT17	−810.07	−2.44	−1,271.22	−3.43	−755.86	−2.59	−912.14	−2.97
RD1	127.00	3.00	146.46	3.25	247.66	6.08	254.26	6.16
RD2	8.23	0.14	25.85	0.44	23.92	0.45	29.89	0.55
RD3	231.68	2.96	123.89	1.54	−160.25	−2.21	−196.77	−2.69
BRPRCORN	2.53	0.05	34.81	0.69	56.46	1.20	67.40	1.43
BRPRRICE	−130.83	−2.91	−29.24	−0.61	23.37	0.59	57.80	1.38
POPDEN	0.24	0.73	0.04	0.11	−1.53	−4.98	−1.60	−5.19
Adjusted R^2	0.17		0.09		0.10		0.08	
n	1,624		1,624		1,624		1,624	

[a]Heteroscedasticity-consistent standard errors; see table A.6 for definitions of variables.

Table A.4

First-stage estimations for two-stage least squares.[a]

	Dependent variable			
	LNTEPC		*LNYPC*	
Variable	Coefficient	t-stat.	Coefficient	t-stat.
Intercept	3.576981	14.72	3.457142	9.90
DWED	0.069349	2.04	−0.005397	−0.14
DFED	0.128606	4.89	0.253890	6.13
CULTARPC	0.324051	7.30	0.463644	7.52
NETWTHPC	0.000027	5.72	0.000034	6.76
OWNARPC	0.063819	1.93	0.196306	4.10
FATAGE	0.000364	2.36	0.000625	3.28
DROOF	0.066238	2.90	0.092413	3.07
DFLOOR	0.232010	5.58	0.234021	4.16
DWALL	−0.034008	−0.69	−0.491091	−6.42
DELECT	0.084243	3.05	−0.033090	−0.79
HHTOT	−0.049112	−8.22	−0.010631	−1.50
DEMMO5	−0.325760	−2.03	−0.055476	−0.28
DEMFO5	−0.402474	−2.51	−0.222984	−1.06
DEMM511	−0.429236	−2.88	−0.141843	−0.78
DEMF511	−0.349096	−2.27	−0.415233	−1.93
DEMM1117	−0.094617	−0.52	−0.187179	−0.70
DEMF1117	0.079416	0.41	0.274114	1.16
DEMFGT17	0.267600	1.08	1.527683	3.80
RD1	0.063173	2.11	−0.054124	−1.36
RD2	−0.106064	−2.59	−0.001341	−0.03
RD3	−0.033142	−0.60	0.222817	3.13
BRPRCORN	0.073791	2.24	−0.095874	−2.18
BRPRRICE	−0.078980	−2.49	−0.140642	−3.46
POPDEN	0.000441	1.85	0.001472	4.70
Adjusted R^2	0.47		0.41	
n	1,624		1,624	

[a]See table A.6 for definitions of variables.

Table A.5

Descriptions on all variables used in regression estimations.[a]

Variable	N	Minimum	Maximum	Mean	Std. dev.
BRPRCORN	1,624	3.20	6.30	4.49	0.69
BRPRRICE	1,624	4.65	7.31	5.79	0.52
CULTARPC	1,624	0	2.60	0.38	0.39
DELECT	1,624	0	1.00	0.29	0.45
DEMF05	1,624	0	0.60	0.13	0.13
DEMF1117	1,624	0	0.50	0.06	0.09
DEMF511	1,624	0	0.50	0.10	0.11
DEMFGT17	1,624	0	0.50	0.19	0.08
DEMM05	1,624	0	0.67	0.15	0.14
DEMM1117	1,624	0	0.44	0.06	0.09
DEMM511	1,624	0	0.43	0.10	0.11
DEMMGT17	1,624	0	0.60	0.20	0.08
DFLOOR	1,624	0	1.00	0.14	0.34
DFED	1,624	0	1.00	0.79	0.41
DROOF	1,624	0	1.00	0.71	0.45
DWALL	1,624	0	1.00	0.10	0.29
DWED	1,624	0	1.00	0.87	0.33
FATAGE	1,624	250.80	751.30	446.46	100.16
HCALPC	1,624	268.00	4,583.27	1,714.30	563.09
HHTOT	1,624	3.00	19.00	7.16	2.68
LNTEPC	1,624	2.07	6.07	3.66	0.55
LNYPC	1,624	−0.50	6.06	3.44	0.73
MCALPC	1,624	309.65	5,457.72	1,754.75	653.06
NETWTHPC	1,624	0	63,221.43	2,916.58	5,512.63
OWNARPC	1,624	0	3.43	0.28	0.48
POPDEN	1,624	51.00	223.00	150.06	44.90
RD1	1,624	0	1.00	0.25	0.43
RD2	1,624	0	1.00	0.25	0.43
RD3	1,624	0	1.00	0.25	0.43
TEPC	1,624	7.89	432.98	46.29	35.13
TEPCSQ	1,624	62.18	187,467.90	3,375.68	9,285.26
XMCALPC	1,624	322.17	7,476.07	1,889.49	792.81
XTEPC	1,624	7.89	440.34	46.53	35.74
YPC	1,624	0.61	429.11	41.76	42.25
YPCSQ	1,624	0.37	184,134.60	3,527.75	12,255.16

[a] See table A.6 for definitions of variables.

Table A.6

Variable definitions.

BRPRCORN	= retail price of shelled corn per kilogram (1984 pesos; simple average of respondents for barrio)
BRPRRICE	= retail price of milled rice per kilogram (1984 pesos; simple average of respondents for barrio)
CULTARPC	= cultivated area per capita (average of four survey rounds)
DELECT	= zero–one dummy for presence of electricity for house
DEMF05	= percent of *HHTOT* that are females less than or equal to 5 years of age
DEMF1117	= percent of *HHTOT* that are females greater than 11 years and less than or equal to 17 years of age
DEMF511	= percent of *HHTOT* that are females greater than 5 years and less than or equal to 11 years of age
DEMFGT17	= percent of *HHTOT* that are females greater than 17 years of age
DEMM05	= percent of *HHTOT* that are males less than or equal to 5 years of age
DEMM1117	= percent of *HHTOT* that are males greater than 11 years and less than or equal to 17 years of age
DEMM511	= percent of *HHTOT* that are males greater than 5 years and less than or equal to 11 years of age
DEMMGT17	= percent of *HHTOT* that are males greater than 17 years of age
DFLOOR	= zero–one dummy for quality of flooring materials for house
DFED	= zero–one dummy for years of husband's education, 1 if years in school are greater than 3
DROOF	= zero–one dummy for quality of roofing materials for house
DWALL	= zero–one dummy for quality of materials used for house framing and walls
DWED	= zero–one dummy for years of wife's education, 1 if years in school are greater than 3
FATAGE	= age of head of household (in months)
HCALPC	= household calorie intake per capita per day (from 24-hour recall)
HHTOT	= number of household members
LNTEPC	= logarithm of household total expenditures per week per capita (1984 pesos; varies by round)
LNYPC	= logarithm of current household income per week per capita (1984 pesos; average over four surveys)
MCALPC	= household calorie availability per capita per day (net of surveyed meals served to guests and hired workers; from food expenditure survey)
NETWTHPC	= value of all assets (1984 pesos; average of rounds 1 and 4)
OWNARPC	= owned area per capita (average of four survey rounds)
POPDEN	= municipal population density (persons per square kilometer)
RD1	= zero–one dummy for first round survey
RD2	= zero–one dummy for second round survey
RD3	= zero–one dummy for third round survey
TEPC	= total household expenditures per week per capita (1984 pesos; varies by round)
TEPCSQ	= *TEPC* squared
XMCALPC	= *MCALPC* plus calories fed to guests and hired workers
XTEPC	= *TECP* plus expenditures for food fed to hired workers
YPC	= current household income per week per capita (1984 pesos; average over four surveys)
YPCSQ	= *YPC* squared.

References

Alderman, Harold, 1987, Cooperative dairy development in Karnataka, India: An assessment, Research report 64 (International Food Policy Research Institute, Washington, DC).

Alderman, Harold, 1989, The impact of changes in income and schooling on the demand for food quantity and quality in rural Pakistan (International Food Policy Research Institute, Washington, DC).

Alderman, Harold, M. Ghaffar Chaudry and Marito Garcia, 1988, Household food security in Pakistan: The ration shop system, Working papers on food subsidies no. 4 (International Food Policy Research Institute, Washington, DC).

Anand, Sudhir and Christopher Harris, 1985, Living standards in Sri Lanka, 1973–1981/82: An analysis of consumer finance survey data, Report prepared for the World Bank-Central Bank of Ceylon project on evolution of living standards in Sri Lanka (Oxford University, Oxford).

Behrman, Jere R., 1988, Nutrition and incomes: Tightly wedded or loosely meshed? Pew/Cornell lecture series on food and nutrition policy (Cornell Food and Nutrition Policy Program, Ithaca, NY).

Behrman, Jere R. and Anil B. Deolalikar, 1987, Will developing country nutrition improve with income? A case study for rural South India, Journal of Political Economy 95, no. 3, 492–507.

Behrman, Jere R. and Barbara L. Wolfe, 1984, More evidence on nutrition demand: Income seems overrated and women's schooling underemphasized, Journal of Development Economics 14, nos. 1/2, 105–128.

Bhargava, Alok, 1991, Identification and panel data models with endogenous regressors, Review of Economic Studies 58, 129–140.

Bliss, Christopher and Nicholas Stern, 1978, Productivity, wages, and nutrition. Part II: Some observations, Journal of Development Economics 5, no. 4, 363–398.

Bouis, Howarth E. and Lawrence J. Haddad, 1990, Agricultural commercialization, nutrition, and the rural poor: A study of Philippine farm households (Lynne Rienner Publishers, Boulder, CO).

von Braun, Joachim, David Hotchkiss and Maarten Immink, 1989, Nontraditional export crops in Guatemala: Effects on production, income, and nutrition, Research report 73 (International Food Policy Research Institute, Washington, DC).

von Braun, Joachim, Detlev Puetz and Patrick Webb, 1989, Irrigation technology and commercialization of rice in the Gambia: Effects on income and nutrition, Research report 75 (International Food Policy Research Institute, Washington, DC).

von Braun, Joachim, Hartwig de Haen and Juergen Blanken, 1991, Commercialization of agriculture under population pressure: Effects on production, consumption, and nutrition in Rwanda, Research report 85 (International Food Policy Research Institute, Washington, DC).

Burk, Marguerite C. and Eleanor M. Pao, 1976, Methodology for large-scale surveys of household and individual diets, Home Economics Research Report No. 40 (United States Department of Agriculture, Washington, DC).

Chernichovsky, Dov and Oey Astra Meesook, 1984, Patterns of food consumption and nutrition in Indonesia: An analysis of the national socio-economic survey, 1978, World Bank staff working paper no. 670 (The World Bank, Washington, DC).

Edirisinghe, Neville, 1987, The food stamp scheme in Sri Lanka: Costs, benefits, and options for modification, Research report 58 (International Food Policy Research Institute, Washington, DC).

Friedman, Milton, 1957, A theory of the consumption function (Princeton University Press, Princeton, NJ).

Garcia, Marito H. and Per Pinstrup-Andersen, 1987, The pilot food price subsidy scheme in the Philippines: Its impact on income, food consumption, and nutritional status, Research report 61 (International Food Policy Research Institute, Washington, DC).

Griliches, Zvi and Jerry A. Hausman, 1986, Errors in variables in panel data, Journal of Econometrics 31, 93–118.

Hausman, Jerry A. and William E. Taylor, 1981, Panel data and unobservable individual effects, Econometrica 49, Nov., 1377–1398.

Hsiao, Cheng, 1986, Analysis of panel data (Cambridge University Press, New York).
Kennedy, Eileen, 1989, The effects of sugarcane production on food security, health, and nutrition in Kenya: A longitudinal analysis, Research report 78 (International Food Policy Research Institute, Washington, DC).
Kumar, Shubh K. and David Hotchkiss, 1988, Consequences of deforestation for women's time allocation, agricultural production, and nutrition in hill areas of Nepal, Research report 69 (International Food Policy Research Institute, Washington, DC).
Lipton, Michael, 1983, Poverty, undernutrition, and hunger, Working paper no. 597 (World Bank, Washington, DC).
National Statistical Coordination Board (NSCB), 1988, Philippine Statistical Yearbook 1988 (National Economic Development Authority, Republic of the Philippines, Manila).
Pinstrup-Andersen, Per and Elizabeth Caicedo, 1978, The potential impact of changes in income distribution on food demand and human nutrition, American Journal of Agricultural Economics 60, Aug., 402–415.
Pitt, Mark M., 1983, Food preferences and nutrition in rural Bangladesh, The Review of Economics and Statistics 65, Feb., 105–114.
Pitt, Mark M., Mark R. Rosenzweig and Md Nazmul Hassan, 1990, Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries, American Economic Review 80, no. 5, 1139–1156.
Ravallion, Martin, 1990, Income effects on undernutrition, Economic Development and Cultural Change 38, no. 3, 489–515.
Sahn, David, 1988, The effect of price and income changes on food-energy intake in Sri Lanka, Economic Development and Cultural Change 36, no. 2, 315–340.
Strauss, John, 1984, Joint determination of food consumption and production in rural Sierra Leone: Estimates of a household-firm model, Journal of Development Economics 14, nos. 1/2, 77–103.
Strauss, John and Duncan Thomas, 1989, The shape of the expenditure-calorie curve, Economic Growth Center discussion paper 595 (Yale University, New Haven, CT).
Timmer, C. Peter, and Harold Alderman, 1979, Estimating consumption parameters for food policy analysis, American Journal of Agricultural Economics 61, Dec., 982–987.
Trairatvorakul, Prasarn, 1984, The effects on income distribution and nutrition of alternative rice price policies in Thailand, Research report 46 (International Food Policy Research Institute, Washington, DC).
Ward, John O. and John H. Sanders, 1980, Nutrition determinants and migration in the Brazilian northeast: A case study of rural and urban Ceara, Economic Development and Cultural Change 29, Oct., 141–163.
Williamson-Gray, Cheryl, 1982, Food consumption parameters for Brazil and their application for food policy, Research report 32 (International Food Policy Research Institute, Washington, DC).
Wolfe, Barbara L. and Jere R. Behrman, 1983, Is income overrated in determining adequate nutrition?, Economic Development and Cultural Change 31, April, 525–549.