



**AgEcon** SEARCH  
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search  
<http://ageconsearch.umn.edu>  
[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

**Does Measurement Error Explain a Paradox about  
Household Size and Food Demand?  
Evidence from Variation in Household Survey Methods**

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

*Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting,  
Montreal, Canada, July 27-30, 2003*

*Copyright 2003 by John Gibson. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.*

**Does Measurement Error Explain a Paradox about  
Household Size and Food Demand?  
Evidence from Variation in Household Survey Methods**

John Gibson<sup>†</sup>

**Abstract**

*Several recent papers report a puzzling pattern of food demand falling as household size rises at constant per capita expenditure, especially in poorer countries. This pattern is contrary to a widely used model of scale economies. This paper exploits within-country differences in household survey methods and interviewer practices to provide a measurement error interpretation of this puzzle. A comparison of household surveys in Cambodia and Indonesia with the results from Monte Carlo experiments suggest that food expenditure estimates from shorter, less detailed recall surveys have measurement errors that are correlated with household size. These correlated measurement errors contribute to the negative effect of household size on food demand and cause upward bias in Engel estimates of household scale economies.*

**JEL: D12, O12, Q11**

Keywords: Food demand, Economies of scale, Household surveys, Measurement error

**Acknowledgements:**

I am grateful to Andy Horowitz and Scott Rozelle for helpful discussions and to Trinh Le and Susan Olivia for outstanding research assistance. All remaining errors in this paper are those of the author. The financial support from the AARES-AAEA Young Professionals Exchange Travel Award – Heading North is gratefully acknowledged.

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

## I. Introduction

The relationship between household size and food demand is important to our understanding of economies of scale within households, which in turn affects empirical analyses of poverty and food security. In a widely cited paper, Deaton and Paxson (1998) report the puzzling result that at constant per capita expenditure (PCE), the budget share for food falls with increases in household size, especially in poorer countries. A unit increase in the logarithm of household size decreases the food budget share by up to 10 percentage points in a group of poor countries (Thailand, Pakistan, and African households in South Africa), holding outlay per person constant.<sup>1</sup> The food share falls by 1-2 percentage points in Taiwan and the U.S., and by less in France and Britain. These empirical results have been confirmed by Gardes and Starzec (2000), although the negative effect of household size is somewhat weakened if household economic resources are measured with an equivalence scale rather than in *per capita* terms.<sup>2</sup> Confirmation is also provided by Gan and Vernon (2003), using alternative countries and different definitions of food shares.

This pattern of declining food shares is exactly the opposite to what basic household demand theory predicts. According to the two-good Barten (1964) model of scale economies, larger households should have higher food demand; holding PCE constant, increasing household size allows resources released by the wider sharing of public goods (e.g., light) to be spent on both public and private goods, giving a positive income effect. While substitution effects favour public goods, which are effectively cheaper in larger households, the income effect should

---

<sup>1</sup> With an average food share of 0.5, a 20 percent decrease in per capita food expenditures is implied.

<sup>2</sup> Perali (2001) also makes this point.

prevail for private goods like food where the (absolute) own-price elasticity is lower than the income elasticity. Hence, food expenditure per head is predicted to rise with household size, particularly in poor countries where the income elasticity of food demand is highest.

Several possible explanations are listed by Deaton and Paxson (1998) for the puzzling evidence on food demand, ranging from economies of scale in food preparation and calorie overheads through to measurement error. While none are considered convincing they suggest that further research be carried out on at least some of them. Several papers in the economics literature have taken up this challenge and attempt to explain what has become known as the “Deaton and Paxson paradox” (Horowitz, 2002). But despite the professional interest that agricultural economists have in food demand, there has been no attempt to look at the puzzling relationship between household size and food demand in the agricultural economics literature. This lack of interest is surprising because household scale economies affect many practical questions, such as whether widow-headed households are poorer than others (Dreze and Srinivasan, 1997), whether interventions in transition economies should be aimed at children or at the elderly (Lanjouw, Milanovic and Paternostro, 1998), and whether subsidized food rations should be targeted to small households or large ones (Olken, 2002).

This paper pursues a measurement error interpretation of the Deaton and Paxson paradox. Random, within-country variation in household survey methods and interviewer practices is exploited to provide evidence on the possible role of reporting errors. These effects may matter because of the wide range of methods used to gather expenditure data in different countries. However, it has previously proved hard to isolate these effects because factors associated with

the other explanations also differ across countries. Thus, by focusing on exogenous variation in survey methods within a country, the role of measurement error may become clearer. A measurement error explanation may also be timely because according to Chesher and Schluter (2002: 377) “measurement error is an ever-present, generally significant, but usually neglected, feature of survey based income and expenditure data.”

The next section of the paper reviews the basic demand theory that underlies Deaton and Paxson’s prediction of a positive effect of household size on food demand. In Section III, Monte Carlo evidence on measurement error is reported and the test procedure is outlined. The empirical results from household surveys in Cambodia and Indonesia are then compared with the results from the Monte Carlo experiments (Sections IV and V). This comparison suggests that food expenditure estimates from shorter, less detailed recall surveys have measurement errors that are correlated with household size. These correlated errors cause a negative bias in the coefficient on household size in regression models of food budget shares and in this way may contribute to the paradoxical results found by Deaton and Paxson (1998). These measurement errors also cause an upward bias in Engel estimates of household scale economies, and thus may interfere with poverty measurement and food policy analysis.

## **II. The Deaton and Paxson ‘Paradox’**

To show the necessary conditions for increases in household size to raise food demand, Deaton and Paxson (1998) use the two-good model in Barten (1964). A household where everything is shared equally among  $n$  members allocates consumption between food,  $q_f$  and a non-food good, such as housing,  $q_h$ , in order to maximise utility,  $u$ :

$$\begin{aligned} \max_{q_f, q_h} u &= n \mathbf{u} \left( \frac{q_f}{\mathbf{f}_f(n)}, \frac{q_h}{\mathbf{f}_h(n)} \right) \\ \text{subject to : } p_f \left( \frac{q_f}{n} \right) + \left( \frac{p_h}{n} \right) q_h &= \frac{x}{n} \end{aligned} \quad (1)$$

where  $x$  is total household expenditures,  $p_f$  and  $p_h$  are the price of food and non-food, and  $\mathbf{f}_i(n)$  (where  $i=f, h$ ) is the scaling function that transforms the number of members,  $n$  into ‘effective’ size. The commodity-specific degree of economies of household scale are:

$$\mathbf{s}_i = 1 - \frac{\partial \ln \mathbf{f}_i(n)}{\partial \ln n}. \quad (2)$$

The per capita food demand function is:

$$\frac{q_f}{n} = \frac{\mathbf{f}_f(n)}{n} g_f \left( \frac{x}{n}, \frac{p_f \mathbf{f}_f(n)}{n}, \frac{p_h \mathbf{f}_h(n)}{n} \right) \quad (3)$$

where  $g_f(x, p_f, p_h)$  is the food demand function for a single person household. When the logarithm of equation (3) is differentiated with respect to  $\ln n$ , the condition for per capita food consumption to increase with household size, holding  $x/n$  constant becomes apparent:

$$\frac{\partial \ln(q_f/n)}{\partial \ln n} > 0 \Leftrightarrow \mathbf{s}_h(\mathbf{e}_{fx} + \mathbf{e}_{ff}) - \mathbf{s}_f(1 + \mathbf{e}_{ff}) > 0 \quad (4)$$

where  $\mathbf{e}_{ff}$  and  $\mathbf{e}_{fx}$  are the own-price and income elasticities of demand for food. If non-food contains some public goods, so that  $\mathbf{s}_h \neq 0$ , while food is a pure private good ( $\mathbf{s}_f = 0$ ), and if the (absolute) own-price elasticity is less than the income elasticity of food demand, per capita food consumption will increase with household size at constant  $x/n$ . This condition is most likely to hold for poor consumers, so the positive effect of household size on per capita food consumption and food budget shares is predicted to be strongest in poor countries.

To test whether the empirical evidence is consistent with this pattern, Deaton and Paxson (1998) use the following food share model:

$$\frac{p_f q_f}{x} = w_f = \mathbf{a} + \mathbf{b} \ln \left( \frac{x}{n} \right) + \mathbf{g} \ln n + \sum_{j=1}^{J-1} \mathbf{h}_j r_j + \mathbf{d} \cdot \mathbf{z} + u \quad (5)$$

where  $r_j = n_j/n$  is the proportion of persons in the household in demographic group  $j$ ,  $\mathbf{z}$  is a vector of other household characteristics,  $u$  is a disturbance term, and  $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{g}$ ,  $\mathbf{h}$ , and  $\mathbf{d}$  are parameters to be estimated. If the condition in equation (4) holds,  $\mathbf{g}$  should be positive, with the largest values estimated from household data in poor countries. In fact, the empirical results of Deaton and Paxson show exactly the opposite pattern, with  $\mathbf{g}$  most negative in poor countries.

One possible reason for the conflict between theory and evidence may be that the two-good model used to generate the theoretical predictions is too restrictive. Horowitz (2002) shows that in a three-good model, food demand rises with household size at constant PCE only if food and the public good are complements. Thus, the condition in equation (4) on the size of the own-price and income elasticities of food demand may not be the relevant one and there may be no basis for predicting that food demand should rise as household size increases at constant PCE.

But even if a two-good model is inappropriate, Deaton and Paxson's results still imply another unresolved puzzle about economies of scale. Their food Engel curve, when reparameterized, gives estimates of economies of scale based on the Engel method used by Lanjouw and Ravallion (1995). This alternative Engel curve is:

$$w_f = \mathbf{a} + \mathbf{b} \ln \left( \frac{x}{n^{1-s}} \right) + \sum_{j=1}^{J-1} \mathbf{h}_j r_j + \mathbf{d} \cdot \mathbf{z} + u, \quad (6)$$



which is identical to equation (5) because  $\mathbf{g}=\mathbf{b}\mathbf{s}$ .<sup>3</sup> The Engel estimates of scale economies from equation (6) are given by the ratio of  $\hat{\mathbf{g}}$  to  $\hat{\mathbf{b}}$ . The large negative  $\hat{\mathbf{g}}$  found in poor countries contributes to large Engel estimates of scale economies. These large scale economy estimates imply improbable reductions in the per capita food expenditures of larger households (Deaton, 1997). Thus, even if Horowitz (2002) is correct in arguing that the Deaton and Paxson ‘paradox’ occurs only because it uses a restrictive two-good model, there is still a puzzle to solve about why  $\hat{\mathbf{g}}$  is estimated to be so negative in poor countries.

### III. Measurement Error and the Testing Procedure

To see how measurement errors might affect the estimates of  $\hat{\mathbf{g}}$  from equation (5), Gibson (2002) carried out Monte Carlo experiments on a simplified version of the model:

$$w_f = \mathbf{a} + \mathbf{b} \ln \left( \frac{x}{n} \right) + \mathbf{g} \ln n + u. \quad (7)$$

Such experiments are needed because  $\ln(x/n)$  and  $\ln n$  are negatively correlated by construction, so errors in  $\ln(x/n)$  are likely to bias  $\hat{\mathbf{g}}$ , but in an unpredictable direction (Deaton and Paxson, 1998). Bias in  $\hat{\mathbf{g}}$  is even more likely if the errors are correlated with household size or with the true value of expenditures. To implement the experiments, total expenditure,  $x$  was partitioned into food expenditures,  $x_f = x \cdot w_f$  and non-food expenditures,  $x_{nf} = x - x_f$ . A proportionate error was added to true food expenditures, so that the observed variable was  $\ln \tilde{x}_f = \ln x_f + v$ . In the first experiment the measurement error was independent of any of the variables in the model:  $v \sim N(0, \mathbf{s}_v^2)$ , with three values of  $\mathbf{s}_v$  used; 0.1, 0.2, and 0.3. In the second experiment errors

---

<sup>3</sup> By rewriting  $\mathbf{b} \ln(x/n^{1-s})$  as  $\mathbf{b} \ln x - (1-s)\mathbf{b} \ln n$  it is clear that  $\mathbf{b} \ln(x/n) + \mathbf{b}\mathbf{s} \ln n = \mathbf{b} \ln(x/n^{1-s})$ .

were correlated with true values,  $v = \mathbf{j} \ln x_f + \mathbf{e}$ , where  $\mathbf{e} \sim N(0, \mathbf{S_e}^2)$  and  $E(\mathbf{e}, x_f) = 0$ .<sup>4</sup> In the third experiment errors were correlated with household size,  $v = \mathbf{l} \ln n + \mathbf{e}$ , where  $\mathbf{e} \sim N(0, \mathbf{S_e}^2)$  and  $E(\mathbf{e}, n) = 0$ . The values used for  $\mathbf{j}$  and  $\mathbf{l}$  were -0.3, -0.2, and -0.1. The error-ridden total expenditure and food share variables were reconstructed as  $\tilde{x} = \tilde{x}_f + x_{nf}$  and  $\tilde{w}_f = \tilde{x}_f / \tilde{x}$ , and equation (7) and the Engel scale elasticity,  $\mathbf{s} = \mathbf{g}/\mathbf{b}$  were estimated.

The results of the Monte Carlo experiments in Table 1 suggest that errors in measuring food expenditures that are negatively correlated with either household size (row 3b) or with the true value of food expenditures (row 2b) could cause negative bias in estimates of  $\mathbf{g}$ . Also, if measurement errors in food expenditures are correlated with the true value of expenditures, the coefficient on  $\ln(x/n)$ ,  $\hat{\mathbf{b}}$  will suffer attenuation bias (i.e., towards zero) but if errors are correlated with household size, there will be no effect on  $\hat{\mathbf{b}}$  (see row 2a and 3a). It is also apparent that errors in measuring food expenditures that are negatively correlated with either true values (row 2c) or with household size (row 3c) can cause  $\hat{\mathbf{s}}$  to be biased upwards.

The results of the Monte Carlo experiments suggest that one way to observe the effect of correlated measurement errors is to estimate a food Engel curve with an interaction term between household size and a dummy variable,  $D$  indicating differences in household survey methods. For example, if it is assumed that reporting errors are less likely when households have their expenditures measured with a long, detailed recall questionnaire rather than with a shorter recall, the effect of errors correlated with household size may be observed from:

---

<sup>4</sup> Bound, Brown and Mathiowetz (2001) point out that survey errors are often negatively correlated with true values.

$$w_f = \mathbf{a} + \mathbf{d}_0 D + \mathbf{b} \ln \left( \frac{x}{n} \right) + \mathbf{g} \ln n + \mathbf{d}_1 [\ln n \times D] + \sum_{j=1}^{J-1} \mathbf{h}_j r_j + \mathbf{l} \cdot \mathbf{z} + u \quad (8)$$

where  $D=1$  if the household expenditures are measured with the long recall and  $D=0$  if the short recall is used. If  $\hat{\mathbf{d}}_1 > 0$  it would imply that reporting errors in shorter, less detailed surveys are correlated with household size, where such a correlation could occur because of the greater number of food purchases to recall in larger households (Gibson, 2002).

In contrast, if errors are negatively correlated with the true value of food expenditures, the bias will affect not only  $\hat{\mathbf{g}}$  but also  $\hat{\mathbf{b}}$  (see row 2a, Table 1). Consequently, other variables may also need to be interacted with  $D$ , giving the more general model:

$$w_f = \mathbf{a} + \mathbf{k}_0 D + \mathbf{b} \ln \left( \frac{x}{n} \right) + \mathbf{k}_1 \left[ \ln \left( \frac{x}{n} \right) \times D \right] + \mathbf{g} \ln n + \mathbf{k}_2 [\ln n \times D] \\ + \sum_{j=1}^{J-1} \mathbf{h}_j r_j + \sum_{j=3}^{J+1} \mathbf{k}_j [r_j \times D] + \mathbf{l} \cdot \mathbf{z} + \mathbf{j} \cdot [\mathbf{z} \times D] + u. \quad (9)$$

In equation (9),  $\mathbf{k}_2 > 0$  and  $\mathbf{k}_1 > 0$  would be consistent with reporting errors that are negatively correlated with the true value of food expenditures. On the other hand, errors that are correlated with household size would imply  $\mathbf{k}_2 > 0$  and  $\mathbf{k}_1 = 0$ .

#### IV. Data

To estimate equations (8) and (9), data from household surveys carried out in 1999 in two developing countries, Cambodia and Indonesia, are used. Both of these surveys feature random variation in the methods and practices used *within* each country. By relying on within- rather than between-country variation, most of the other factors listed as possible explanations by Deaton and Paxson should be held constant. If estimated food demand parameters then differ

between two randomly selected groups of households whose expenditures were measured in different ways, measurement error emerges as a much more plausible explanation.

### *Indonesia*

The annual SUSENAS (National Socio-Economic Household Survey) has a short consumption recall, where respondents provide details on their household's expenditures on 15 food groups over the previous week and eight non-food groups over the previous month and year. The sample of 32,000 households used here is restricted to urban areas on the island of Java, because household wage income is used as an instrument for total expenditures and wage earning is much more prevalent in urban areas. In 1999 a random subset of almost 13,000 households from the SUSENAS sample in urban Java were given a much longer consumption recall questionnaire with over 300 items specified.<sup>5</sup>

Households who receive the long recall questionnaire have measured per capita expenditures that are almost one-quarter higher than the average for the households receiving the short recall (Table 2). The food budget share is also lower, suggesting that non-food expenditures are raised most by using the more detailed recall, corroborating results reported by Deaton (1997). Except for these recall questionnaire effects, there is no evidence that the two samples of households differ in any significant way. Variables measuring literacy and gender composition, which may affect household income, show no difference in means between the two samples.

---

<sup>5</sup> In principle, each household received both the short and the long recall questionnaires so the comparison of different survey methods could use the *same* households. But in practice there appears to be widespread copying from the long recall survey form to the short recall form (Sumarto, *et al.* 2002) so such a comparison is unlikely to find anything.

Dwelling attributes, which may proxy for wealth, show few differences (the difference in the prevalence of earthen floors is weakly significant, at  $p<0.10$ ).

### *Cambodia*

The Cambodia Socio-Economic Survey (CSES) did not aim to apply different procedures to different groups in the population but variation in interviewer practice appears to have produced the same effects. This variation is apparent because the sample was randomly split, with half of the households interviewed between January and March (Round 1), and the remainder interviewed between June and September (Round 2). Between the two rounds, interviewers were retrained, where it was emphasised that estimates of household consumption should be ‘reasonable’ given the estimate of household income. To facilitate these income-expenditure comparisons the questionnaires included a Household Income and Expenditure Balance Sheet. Consistent with a greater effort made to reconcile household total income,  $y$  and total expenditure,  $x$  there is a much closer relationship between the two variables in Round 2 of the survey than there was in Round 1:

Round 1	Round 2
$\ln x = 3.25 + 0.777 \ln y$	$\ln x = 2.01 + 0.862 \ln y$
$R^2=0.60$	$R^2=0.80$

The rise in the estimated income elasticity of expenditure between the two survey rounds is statistically significant ( $p<0.02$ ). As a result of the extra effort to match expenditure and income, there is a 20 percent rise in measured expenditures between the two survey rounds (Table 2).<sup>6</sup>

---

<sup>6</sup> Just because the expenditure data match the income data more closely, it does not necessarily indicate greater accuracy in Round 2 of the survey. Because the income estimates come from the survey as well, they cannot serve as independent validations of the expenditure estimates.

Another possible cause for growth in the expenditure estimates between the survey rounds is that the sample splitting was not random. However, comparisons between the two groups of households in terms of dwelling characteristics (as proxies for wealth) and literacy (as a proxy for income) reveal no evidence that the sub-samples differ in any systematic way (Table 3). Also, if one sub-sample was significantly better off, it would also be expected to alter the food budget share (according to Engel's Law) but the average food share is almost the same across survey rounds, and if anything, indicates that the households in Round 2 are worse off.

Seasonality also can be ruled out as an explanation for the 20 percent jump in reported expenditures in the June-September round of the survey. In a previous survey in Cambodia in 1993/94, these months had lower expenditures than the January-March period (Gibson, 2000). Similarly, other surveys in the region do not show any jump in expenditures in June-September; in the 1998 Vietnam Living Standards Survey (VLSS) expenditures declined by one percent between January-March and June-September. Moreover, the lower inequality in Round 2 of the 1999 Cambodian survey (the Gini fell from 0.40 to 0.29) is surprising if seasonality is operating. Greater dispersion would usually occur in the abundant season (so the mean and variance rise together) because it is then, rather than in the 'hungry season', that heterogeneity in preferences can play a large role in consumption decisions (Behrman, Foster and Rosenzweig, 1997).

In contrast to the unconvincing evidence about seasonality, several indicators suggest a more diligent interviewer performance, with greater probing in Round 2 of the survey. The share of households requiring re-interviews, due to incomplete and/or inconsistent questionnaires, fell from 40 percent to 28 percent in Round 2 (Table 4). The average proportion of households

reporting zero expenditures fell from 48 percent to 43 percent, while the proportion reporting zero own-production also fell.<sup>7</sup> While these falls could be due to seasonality, the zero response rates would normally go in opposite directions for purchases and own-production, as producer-households exhaust their stocks and switch to market purchases. Thus, it seems plausible that the data in Round 2 of the CSES reflect a more probing interview style, so variation across the survey rounds in the estimated food Engel curve may indicate something about measurement error effects in food demand models.

## V. Results

The results of estimating equations (8) and (9) are reported in Table 5 for Indonesia and Table 6 for Cambodia. The regression model in each case is based on the specification used by Deaton and Paxson for Thailand, which is the closest country in their sample to the countries studied here. In addition to (log) PCE and (log) household size, the variables include 11 demographic ratios, the fraction of adults in each household working in agricultural employment, agricultural self-employment, and non-agricultural employment, and dummy variables for farm households and for each province (sector rather than province in Cambodia).

The two equations are estimated by both OLS and Instrumental Variables (IV), which are two of the four estimation methods used by Deaton and Paxson. The justification for using IV is that random measurement errors in  $\ln(x/n)$  might bias the  $g$  coefficient because of the correlation between  $\ln(x/n)$  and  $\ln n$ . The instrument used by Deaton and Paxson is household income, excluding imputed items that are common with expenditures. This variable is not

---

<sup>7</sup> On an item basis, 20 out of the 23 food items in the consumption recall list had a statistically significant fall in the proportion of households reporting zero consumption between survey rounds.

available for the annual SUSENAS survey, so wage and salary income is used instead. Because only 60 percent of the sample have wage and salary earnings, the OLS equation is run twice – once on all households in the sample for urban Java and once on just those with earnings. While the point estimates change between these two samples, the pattern of results is qualitatively the same.

### *Indonesia*

Questionnaire design has a significant effect on the estimated relationship between household size and food demand. When the long recall list is used to measure expenditures, the negative effect of household size on the food budget share (at constant PCE) is significantly smaller for all samples and all estimators in Table 5. This recall effect is shown by the coefficient on the interacted dummy variable term,  $[\ln n \times D]$  being positive (ranging from 0.010 to 0.016) and statistically significant in all columns. In other words, when the household survey uses a longer list of foods for collecting recalled expenditures, the negative effect of household size on the food budget share is less apparent.

Similarly, the difference between the short recall and the long recall samples in the estimated elasticity of per capita food demand with respect to household size,  $g/\bar{w}_f$  is statistically significant in most cases. The other apparent questionnaire effect is that the Engel estimates of economies of scale are about one-quarter larger when household expenditures are measured with a short recall, and this difference is always statistically significant.



Comparing the columns on the right of Table 5 with those on the left suggests that using the more general model (equation (9)) makes little difference to the results. Thus, even when all coefficients are allowed to vary between the short and the long recall samples, it is usually only the interaction between household size and the dummy variable for the long recall questionnaire that attracts a significant coefficient. This result is consistent with the pattern that would be expected if expenditure reporting errors are correlated with household size. Similarly, the use of IV estimation does not alter the basic pattern. Even though the IV estimates of equation (9) are significantly different from the OLS estimates, the gap between the short-recall and long-recall estimates of the Engel elasticity of household scale is almost identical to the gap with the OLS estimates (0.51 for short recall and 0.42 for short recall).<sup>8</sup>

### *Cambodia*

The effect of variation in interview practice in the Cambodian survey (probing in order to reconcile income and expenditure estimates in Round 2) appears to have an even stronger effect on the food Engel curve than did the use of short rather than long recall in Indonesia. When interviewers appear to adopt a more probing interview style in Round 2, the puzzling negative relationship between  $\ln n$  and  $w_f$  almost disappears. The difference in  $\hat{g}$  between survey rounds varies from 0.031 to 0.053, depending on the estimation method and whether the fully interacted model (equation (9)) is used. In other words, *within* the Cambodian survey, the difference in the effect on the food share of a unit increase in the logarithm of household size is greater than many of the between country differences explored by Deaton and Paxson. Because nothing else seems

---

<sup>8</sup> The added variable form of the Hausman test for equation (8) gives  $t=0.28$ , while for equation (9) it gives  $F_{(2,1430)}=4.57$ .

to differ between the two groups of households in Round 1 and Round 2, measurement error emerges as a prime suspect and it is therefore reasonable to suppose that similar errors affect the between country comparisons.

In terms of the Engel estimates of economies of scale, there appear to be significant scale economies available in Round 1, with  $\sigma$  ranging from 0.37 to 0.40. This range is very close to the estimate reported for Pakistan by Lanjouw and Ravallion (1995). In contrast, in Round 2 of the survey, where interviewers appear to have probed more, scale economies appear to have evaporated. The estimates of  $\sigma$  range from 0.04 to 0.08, and are always statistically insignificant.

Replicating the analysis using survey data from neighbouring Vietnam suggests that the results for Cambodia do not just reflect seasonality. When the sample from the 1998 VLSS was split into a January-March ‘round’ and a June-September ‘round’ there was no significant difference in the Engel curve coefficients across the two rounds.<sup>9</sup> Similarly, the elasticity of per capita food expenditures and the Engel scale elasticity did not differ between the periods that correspond to the two rounds of the Cambodian survey. Thus, it seems highly unlikely that the results in Table 6 are being driven by seasonal changes in household behaviour.

One interpretation of this evidence, which is consistent with the Monte Carlo results in Table 1, is that food expenditures collected with a recall questionnaire have measurement errors that are correlated with household size. In the absence of probing, a respondent in a recall survey is likely to forget expenditures. As household size increases it becomes increasingly harder for

---

<sup>9</sup> These results are available from the author.

the respondent to accurately recall all food expenditures, because the number of transactions to remember is likely to grow with the number of residents in the household. The Engel method may mistake this underestimate in the food expenditures of large households for genuine scale economies.

## **VI. Conclusions**

The puzzling finding of Deaton and Paxson (1998) and others, that food shares fall as household size rises at constant per capita expenditures, especially in poor countries, has been examined. By exploiting differences in questionnaire design within the Indonesian SUSENAS (long recall versus short recall) and interviewer behaviour in the Cambodian CSES (more probing versus less), the possible effect of measurement error has been observed. The evidence conforms with the hypothesis that food expenditures collected with a shorter, less probing, recall questionnaire have reporting errors that are correlated with household size. In the absence of prompting from the more detailed recall list, a respondent in a recall survey is likely to forget expenditures, especially in larger households where there are more transactions to remember. These reporting errors may matter to the Deaton and Paxson puzzle because of the wide range of methods (diaries, long recall, short recall) used to gather expenditure data in the countries they study. The results also suggest that underreporting of food expenditures in large households may inflate Engel estimates of scale economies, which might explain why the estimates reported by Lanjouw and Ravallion (1995) are so large.

However, even when the presumably more accurate long recall questionnaire is used, there is a significant negative effect of household size on per capita food demand in Indonesia

(but not in Round 2 of the Cambodian survey). This negative effect conflicts with the theoretical predictions of the Barten (1964) two-good model of scale economies. Thus it is likely that measurement error is only one part of the explanation for the puzzle raised by Deaton and Paxson about the effect of household size on food demand.

In addition to possibly answering a puzzle about food demand raised in the economics literature, the main implication of the results is in highlighting the importance of measurement error in household expenditure data. It appears that in the absence of prompting from either a more detailed recall list or a more probing interview style, a respondent in a recall expenditure survey is likely to forget food expenditures, especially in larger households where there are more transactions to remember. This underreporting of food expenditures in large households may inflate Engel estimates of scale economies, and in this manner interfere with empirical measurement of poverty and food insecurity. Because of the heavy reliance by economists and agricultural economists on recall surveys of household expenditures, such evidence of reporting error and its effects on econometric estimates and policy parameters is highly concerning.

## References

- Barten, A. (1964) 'Family composition, prices, and expenditure patterns,' in *Economic Analysis for National Economic Planning*, P. Hart, G. Mills, and J. Whitaker (eds). Butterworth: London.
- Behrman, J., Foster, A. and Rosenzweig, M. (1997). 'The dynamics of agricultural production and the calorie-income relationship: Evidence from Pakistan.' *Journal of Econometrics* 77(1): 187-207.
- Bound, J., Brown, C., and Mathiowetz, N. (2001). 'Measurement error in survey data.' In *Handbook of Econometrics* Vol. 5. J. Heckman and E. Leamer (eds), Elsevier, pp. 3705-3843.
- Chesher, A., and Schluter, C. (2002). 'Welfare measurement and measurement error.' *Review of Economic Studies* 69(2): 357-378.
- Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore: Johns Hopkins University Press.
- Deaton, A., and Paxson, C. (1998). 'Economies of scale, household size, and the demand for food.' *Journal of Political Economy*, 106(5): 897-930.
- Dreze, J., and Srinivasan, P. (1997) 'Widowhood and poverty in rural India: some inferences from household survey data.' *Journal of Development Economics* 54 (2): 217-234.
- Gan, L., and Vernon, V. (2001) 'Testing the Barten model of economies of scale in household consumption: toward resolving a paradox of Deaton and Paxson' *Journal of Political Economy*, (forthcoming).
- Gardes, F., and Starzec, C. (2000). 'Economies of scale and food consumption: a reappraisal of the Deaton-Paxson paradox.' Working Paper no 2000-08, TEAM, University of Paris I – Pantheon-Sorbonne.
- Gibson, J. (2000) A Poverty Profile of Cambodia, 1999. A Report to the World Bank and the Ministry of Planning, Phnom Penh, Cambodia.
- Gibson, J. (2002) 'Why does the Engel method work? Food demand, economies of size and household survey methods.' *Oxford Bulletin of Economics and Statistics* 64(4): 341-360.
- Horowitz, A. (2002) 'Household size and the demand for food: a puzzle resolved?' Paper presented at the Northeast Universities Development Consortium Conference, Williams College, October, 2002.
- Lanjouw, P. and Ravallion, M. (1995). 'Poverty and household size.' *The Economic Journal*, 105(Nov): 1415-1434.

- Lanjouw, P., Milanovic, B., and Paternostro, S. (1998). 'Poverty in the transition countries: a case of children pitted against the elderly?' *mimeo* Development Economics Research Group, The World Bank.
- Olken, B. (2002) 'Revealed community equivalence scales' Paper presented at the Northeast Universities Development Consortium Conference, Williams College, October, 2002.
- Perali, F. (2001). 'The second Engel law and economies of scale: an empirical puzzle to be resolved' Working Paper no. 28, Centre for Household, Income, Labour, and Demographic Economics, University of Torino.
- Sumarto, S., Suryahadi, A., and Widyanti, W. (2002). 'Designs and implementation of Indonesian social safety net programs.' *The Developing Economies* 40(1): 3-31.

Table 1: Monte Carlo Results for Food Share Model

		Independent measurement errors in food expenditures			
		$v \sim N(0, \mathbf{s}_v^2)$			
		No error	$\mathbf{s}_v = 0.1$	$\mathbf{s}_v = 0.2$	$\mathbf{s}_v = 0.3$
1a.	$E(\hat{\mathbf{b}})$	-0.1379	-0.1344	-0.1241	-0.1082
1b.	$E(\hat{\mathbf{g}})$	-0.0073	-0.0047	0.0030	0.0146
1c.	$E(\hat{\mathbf{s}})$	0.0518	0.0339	-0.0254	-0.1377
		Food expenditure errors correlated with true values			
		$v = \mathbf{j} \ln x_f + \mathbf{e}$ , $\mathbf{e} \sim N(0, 0.4)$			
		no error	$\mathbf{j} = -0.1$	$\mathbf{j} = -0.2$	$\mathbf{j} = -0.3$
2a.	$E(\hat{\mathbf{b}})$	-0.1379	-0.1282	-0.0940	-0.0560
2b.	$E(\hat{\mathbf{g}})$	-0.0073	-0.0383	-0.0448	-0.0331
2c.	$E(\hat{\mathbf{s}})$	0.0518	0.2986	0.4763	0.5904
		Food expenditure errors correlated with household size			
		$v = \mathbf{I} \ln n + \mathbf{e}$ , $\mathbf{e} \sim N(0, 0.4)$			
		no error	$\mathbf{I} = -0.1$	$\mathbf{I} = -0.2$	$\mathbf{I} = -0.3$
3a.	$E(\hat{\mathbf{b}})$	-0.1379	-0.1263	-0.1262	-0.1242
3b.	$E(\hat{\mathbf{g}})$	-0.0073	-0.0289	-0.0582	-0.0844
3c.	$E(\hat{\mathbf{s}})$	0.0518	0.2282	0.4603	0.6792

Source: Gibson (2002).

Note:

Results based on 10,000 replications of the model:  $w_f = \mathbf{a} + \mathbf{b} \ln(x/n) + \mathbf{g} \ln n + u$ .

The true values are  $\mathbf{a}=1.6$ ,  $\mathbf{b}=-0.14$ ,  $\mathbf{g}=-0.007$ , and  $\mathbf{s}=\mathbf{g}/\mathbf{b} = 0.05$

Each series is 1000 observations.

Table 2: Comparison of the Short Recall and Long Recall SUSENAS Samples, Urban Java

	Short Recall	Long Recall	<i>t</i> -statistic on difference <sup>a</sup>
Per capita expenditures (Rupiah per month) <sup>b</sup>	169,373	209,739	4.56**
Food Budget share	64.5	60.3	7.76**
% of households with main source of lighting from electricity	98.8	98.8	0.22
% of households whose dwelling has earthen floor	7.5	5.9	1.75+
Average floor area of dwelling	74.0	71.9	1.20
% of the household who are male	48.8	49.0	0.56
% of the households whose head is literate	90.7	91.2	0.79
Sample size	19,161	12,876	

<sup>a</sup>Corrected for cluster structure of the samples.

<sup>b</sup>Rp 8730 per US\$ at the time of the survey.

\*\*=significant at 1% level, \*=significant at 5% level, +=significant at 10% level.



Table 3: Comparison of Expenditure Estimates and Sample Characteristics for Two Rounds of the Cambodian Socio-Economic Survey

	Round 1	Round 2	<i>t</i> -statistic on difference <sup>a</sup>
Per capita expenditures (Riel per month)	61,944	74,029	3.48**
Food Budget share	66.0	67.3	1.50
% of households with main source of lighting from electricity	16.1	17.3	0.50
% of households whose dwelling has earthen floor	14.8	13.5	0.65
Average floor area of dwelling	42.3	41.0	0.63
% of the household who are male	47.2	47.6	0.70
% of the households whose head is literate	70.8	73.2	1.19
Sample size	3,000	3,000	

<sup>a</sup>Corrected for cluster structure of the samples.

<sup>b</sup>The exchange rate in Round 1 was R3780 per US\$, and in Round 2 R3840 per US\$.

\*\*=significant at 1% level, \*=significant at 5% level, +=significant at 10% level.

Table 4: Comparison of Proxies for Interviewer Performance in the Two Rounds of the 1999 Cambodia Socio-Economic Survey

	Round 1	Round 2
% of re-interviews due to incomplete, incorrect or doubtful entries	40%	28%
Average % of households recording zero purchases of foods	48.1%	42.8%
Average % of households recording zero own-produce/gifts of foods	69.8%	57.8%

*Source:* Author's calculations.

Table 5: Regression Estimates of Food Demand Models With Explanatory Variables Interacted With Indicators of Survey Recall Type, SUSENAS urban Java

		Equation (8)			Equation (9)		
		Wage Income > 0			Wage Income > 0		
		OLS	OLS	IV <sup>a</sup>	OLS	OLS	IV
<i>Engel curve coefficient estimates</i>							
ln (x/n)	( $\hat{b}$ )	-0.1332** (0.0029)	-0.1308** (0.0033)	-0.1322** (0.0056)	-0.1290** (0.0041)	-0.1259** (0.0046)	-0.1355** (0.0083)
[ln (x/n)]×D		...	...	...	-0.0092+ (0.0052)	-0.0104+ (0.0060)	0.0081 (0.0107)
ln n	( $\hat{g}$ )	-0.0625** (0.0029)	-0.0667** (0.0037)	-0.0672** (0.0042)	-0.0609** (0.0034)	-0.0652** (0.0042)	-0.0689** (0.0052)
[ln n]×D		0.0158** (0.0031)	0.0120** (0.0041)	0.0122** (0.0042)	0.0131** (0.0044)	0.0100+ (0.0055)	0.0156* (0.0064)
R <sup>2</sup>		0.379	0.388	0.388	0.382	0.390	0.389
Sample size		32,037	19,594	19,594	32,037	19,594	19,594
<i>Elasticity of per capita food expenditure w.r.t household size</i> $\partial \ln (q_f/n)/\partial \ln n = \mathbf{g}/\overline{w}_f$							
Short recall		-0.1045** (0.0048)	-0.1119** (0.0063)	-0.1127** (0.0070)	-0.1018** (0.0056)	-0.1093** (0.0071)	-0.1155** (0.0088)
Long recall		-0.0835** (0.0049)	-0.0977** (0.0059)	-0.0983** (0.0062)	-0.0855** (0.0053)	-0.0986** (0.0065)	-0.0952** (0.0067)
H <sub>0</sub> : equal elasticities		F=15.24**	F=4.01*	F=4.02*	F=4.68*	F=1.28	F=3.46+
<i>Engel estimate of household scale economies: <math>\mathbf{s} = \mathbf{g}/\mathbf{b}</math></i>							
Short recall		0.469** (0.018)	0.510** (0.026)	0.508** (0.026)	0.472** (0.020)	0.518** (0.029)	0.508** (0.027)
Long recall		0.350** (0.018)	0.418** (0.023)	0.416** (0.024)	0.346** (0.019)	0.405** (0.024)	0.418** (0.027)
H <sub>0</sub> : $\mathbf{s}_{short} = \mathbf{s}_{long}$		F=26.05**	F=8.57**	F=8.74**	F=22.21**	F=8.98**	F=5.44*

Note: Standard error estimates and hypothesis tests correct for cluster structure of the samples.

Standard errors in ( ); \*\*=significant at 1% level, \*=significant at 5% level, +=significant at 10% level.

The degrees of freedom for the  $F$ -tests are 1 and 1431.

All regressions contain the logarithm of per capita total expenditure, the logarithm of household size, the ratios to household size of the number of males and females in the 0-4, 5-9, 10-14, 15-29, 30-54 and (males only) 55+ age groups, the fraction of adults whose main occupation is agricultural self-employment, agricultural employment and non-agricultural work, a dummy variable for farm households, and dummy variables for province.

<sup>a</sup>The log of per capita wage and salary income is used as the instrument. The partial  $R^2$  for the instrument in the first stage regression is 0.175 and the  $F$ -test for excluding the instrument is 4154.

Table 6: Regression Estimates of Food Demand Models With Explanatory Variables Interacted With Indicators for Survey Round, CSES Cambodia

		Equation (8)		Equation (9)	
		OLS	IV <sup>a</sup>	OLS	IV <sup>a</sup>
<i>Engel curve coefficient estimates</i>					
$\ln(x/n)$	$(\hat{b})$	-0.1162** (0.0069)	-0.1523** (0.0094)	-0.1020** (0.0082)	-0.1630** (0.0117)
$[\ln(x/n)] \times D$		...	...	-0.0519** (0.0153)	0.0180 (0.0195)
$\ln n$	$(\hat{g})$	-0.0434** (0.0082)	-0.0566** (0.0088)	-0.0409** (0.0102)	-0.0607** (0.0112)
$[\ln n] \times D$		0.0386** (0.0085)	0.0447** (0.0088)	0.0308* (0.0141)	0.0527** (0.0148)
$R^2$		0.418	0.405	0.431	0.407
Sample size		6,000	5,997	6,000	5,997
<i>Elasticity of per capita food expenditure w.r.t household size <math>\partial \ln(q_f/n) / \partial \ln n = g/\bar{w}_f</math></i>					
Round 1 (less probing)		-0.0657** (0.0125)	-0.0858** (0.0134)	-0.0620** (0.0155)	-0.0920** (0.0169)
Round 2 (more probing)		-0.0072 (0.0123)	-0.0178 (0.0125)	-0.0151 (0.0144)	-0.0119 (0.0144)
H <sub>0</sub> : equal elasticities		$F=20.88^{**}$	$F=26.72^{**}$	$F=4.91^*$	$F=12.99^{**}$
<i>Engel estimate of household scale economies: <math>s = g/b</math></i>					
Round 1 (less probing)		0.373** (0.067)	0.372** (0.052)	0.401** (0.095)	0.373** (0.061)
Round 2 (more probing)		0.041 (0.071)	0.079 (0.054)	0.066 (0.061)	0.055 (0.065)
H <sub>0</sub> : $s_{Round 1} = s_{Round 2}$		$F=19.43^{**}$	$F=26.64^{**}$	$F=8.77^{**}$	$F=12.70^{**}$

Note: Standard error estimates and hypothesis tests correct for cluster structure of the samples.

Standard errors in ( ); \*\*=significant at 1% level, \*=significant at 5% level, +=significant at 10% level.

The degrees of freedom for the  $F$ -tests are 1 and 590.

All regressions contain the logarithm of per capita total expenditure, the logarithm of household size, the ratios to household size of the number of males and females in the 0-4, 5-9, 10-14, 15-29, 30-54 and (males only) 55+ age groups, the fraction of adults whose main occupation is agricultural self-employment, agricultural employment and non-agricultural work, a dummy variable for farm households, and dummy variables for region and sector.

<sup>a</sup>The log of per capita income (excluding in-kind items that are used in the calculation of both income and consumption) is used as the instrument. The partial  $R^2$  for the instrument in the first stage regression is 0.329 and the  $F$ -test for excluding the instrument is 2923.