Is a Picture Worth a Thousand Unit Values?
Price Collection Methods, Poverty Lines and Price Elasticities in Papua New Guinea

John Gibson*
University of Waikato

Scott Rozelle
University of California, Davis

Abstract

Researchers often use *unit values* (household expenditures on a commodity divided by the quantity purchased) as proxies for market prices when calculating poverty lines and estimating consumer demand equations. Such proxies are often needed because community price surveys in developing countries are either absent or suffer quality problems. However, biases may result from using unit values, due to measurement error and quality effects. In this paper, we report evidence on a household survey experiment where information on prices was obtained in three ways: from unit values, from a market price survey, and from the opinions of householders who were shown pictures of various items and asked to report the local price. These three sets of price data are used to calculate poverty lines and to estimate systems of demand equations and price elasticities. Our results demonstrate substantial biases when unit values are used as a proxy for market price, even when sophisticated correction methods are applied. In contrast, the performance of the price opinions obtained from householders on the basis of the pictures was much better. Hence, a picture-based methodology appears attractive because it may have lower bias than unit values and be less expensive and easier to manage than community price surveys.

JEL: C81, D12, H21, I32

Keywords: Demand, Measurement error, Poverty, Prices, Survey data, Unit values

Acknowledgements
Data used in this paper were originally collected as part of a World Bank poverty assessment for Papua New Guinea, for which financial support from the governments of Australia (TF-032753), Japan (TF-029460), and New Zealand (TF-033936) is gratefully acknowledged.

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

Is a Picture Worth a Thousand Unit Values?
Price Collection Methods, Poverty Lines and Price Elasticities in Papua New Guinea

I. Introduction

Prices are important. Economists need good measures of prices to conduct studies for a broad number of applications in developing countries. For example, when constructing computable equilibrium models for policy and trade analysis, researchers need to have matrices of own- and cross price elasticities of demand (Minot and Goletti, 2000). Similarly, the effective reform of indirect taxation and subsidy regimes requires accurately estimated price elasticities in order to predict the change in the demand for goods and in tax revenues as tax rates change (Ahmad and Stern, 1991). Poverty analysts also need accurate and timely price data to ensure that poverty lines correspond to the actual change in the cost of living for poor people; this issue has affected recent debates about poverty reduction in India (Deaton and Tarozzi, 2000).

Surprisingly, despite being important for so many analyses, few studies systematically collect price data. State statistical bureaus in countries such as China and Indonesia do not collect market price data that can be matched to their rural household income and expenditure surveys. Research-driven surveys also suffer from a lack of price data. For example, the Indonesia Family Life Survey (IFLS2) collected a tremendous amount of data from households and communities, including expenditures on 37 food items, but market price surveys were carried out for only nine foods. This incomplete information on prices makes it difficult to reliably measure the inflation rate that Indonesian households faced during the economic crisis, and may contribute to the large discrepancy between the poverty increases implied by the IFLS price data and those implied by the official (urban) inflation rates (Beegle, *et. al.* 1999). Even in the well-funded and comprehensive Living Standards Measurement Study (LSMS) surveys, there have been problems in gathering prices:

> "In most previous LSMS surveys, interviewers have collected price data by visiting markets and vendors and asking the price of particular goods. … Another possible way to collect prices would be to ask community informants or a sub-sample of household informants about prices. *Given how little is known about how to collect data on community-level prices and how many problems there have been in past LSMS studies* [emphasis ours], it is recommended that both methods be used (Frankenberg, 2000, p. 329)."

Community-level prices, of the type collected in most LSMS surveys, may be unreliable either because they are gathered from the wrong market, or are for the wrong specification of goods, or the prices quoted are not the prices actually paid by local residents (Deaton and Grosh, 2000). Indeed, in some LSMS surveys, the market price data have either never been released because of quality problems (e.g. Tajikistan), or analysts have been forced to discard some of the prices.[1]

This poor track record for collecting price data may not be surprising. In the rural areas of many developing countries it is hard for outsiders to find, understand, and study markets. Markets may meet intermittently, at different places on different days and often at very early hours. Perhaps because managing the traditional part of the data collection effort (household expenditures) is already logistically difficult, adding another part of the survey (for collecting prices) with its own set of complications may cause a decline in overall survey quality. These problems are likely to be most apparent in countries with poor infrastructure and low population densities, which are the very places where price policy can be an important tool for government because of the high per capita administrative cost of delivering income interventions.

Without good price data, economists have had to turn to imperfect proxy measures, such as *unit values* (the ratio of household expenditure on a particular good to the quantity consumed).[2] The range of applications where unit values have recently been used include the calculation of poverty lines (Deaton and Tarozzi, 2000), the analysis of indirect tax and subsidy reforms (Deaton and Grimard, 1992), and assessments of the distributional and nutritional impacts of devaluation (Minot, 1998). However, in some applications, such as demand studies, the use of unit values is believed to give biased results (Deaton, 1997). The problem with unit values is that, in contrast to market prices, they reflect household-specific quality and reporting error effects, and are subject to sample selection effects because they are unavailable for non-purchasing households. Even procedures developed by Deaton (1990) to correct these biases have been shown to produce inaccurate and imprecise results (Gibson and Rozelle, 2002). Alternative strategies, such as using more readily available *urban* price series as proxies for the prices faced by rural households, also may cause bias (Alderman, 1988).

Recognizing that these types of problems with the gathering of price data in household surveys appear to be pervasive, we devised an experiment during a survey in Papua New Guinea (PNG) to test alternative ways of collecting price data. We use three ways to obtain information on prices: from the unit values implicit in household expenditure data; from a market price survey (which we conducted by making repeated trips to the market and surveying traders); and from the 'opinions' of household respondents that were shown pictures of various items during the survey and asked to report the local price for the product in the picture. This picture-based methodology has several potential advantages over traditional, unit value-based approaches: since it is easy to show pictures to all households and ask for their estimates of the price, there are likely to be fewer missing observations. More importantly, any measurement error in these price opinions should not be correlated with actual demands. Finally, biases due to quality effects should be less, since everyone sees and is responding to the same picture.

We use the prices from the market price survey as the standard against which we judge the two alternative price proxies. Although somewhat innocuous, such a preference for relying on market price surveys is not always apparent in the literature (Deaton and Grosh, 2000). In this paper, we explicitly assume that prices for well-defined items collected from market surveys using certain sampling rules are the appropriate standard. In some cases, there may be reasons to worry about the quality of market prices themselves. In the case of our study, however, two features of the case study country increase the reliability of the market price surveys. First, villages are small and in almost every village that the survey team visited, the market that serves the village is well-defined. Second, for whatever reason, haggling is uncommon in markets in PNG. Both of these features mean that the prices observed by enumerators in the local market are likely to be the prices actually faced by households in the survey.

Although our experiment relates to just a single country, we believe that there may be wider interest in our findings. We appear to have made the only systematic attempt to test an idea that was proposed early in the development of the LSMS surveys, which was to obtain price data by interviewing groups of housewives (Saunders and Grootaert, 1980). In light of the subsequent difficulties that price collection efforts in the LSMS faced, it is surprising that there was not more experimentation along these lines.[3] Also, ours is one of the only papers to empirically demonstrate the magnitude of the bias from using unit values as proxies for market prices. Surprisingly, despite the widespread reliance on unit values and despite the plea by Deaton (1990), there has never been a ‘crucial experiment’ in which results calculated from market price data are compared with the results from either naïve or corrected unit value procedures.

The rest of the paper is organized as follows. The next section describes our three methods of collecting price data, while basic descriptive statistics and comparisons of the various price measures are reported in Section III. The results of using the three sets of price data to calculate and compare poverty lines, poverty indices, and price elasticities are reported in Section IV. The final section concludes.

II. Data Collection

Data used in this paper come from the Papua New Guinea Household Survey (PNGHS), which was designed and supervised by the authors in 1995 and 1996. The survey covered a random sample of 1200 households, residing in 73 rural clusters (each providing 12 households to the sample), 40 clusters from the capital city (providing six households each) and seven clusters from smaller urban areas (12 households each). The survey fieldwork was spread over a 12-month period. The key feature of this survey is that it collected information on prices in three ways: market price surveys, unit values, and price opinions of householders who were shown pictures of various items.

Market prices were collected in each cluster using two different surveys. The prices of 14 commercially produced food items (e.g., rice, sugar, beer) and nine non-food items (e.g., soap, kerosene) were collected from the two main trade stores or supermarkets used by the households in the cluster. These prices typically were for a finely defined specification (e.g., a 1kg bag of "Trukai" brand rice). For four of the foods and one of the non-foods, the prices covered two different specifications of the same commodity (e.g., a bottle of beer and a carton of beer). In these cases, the analyses use a simple average of the prices of the two specifications of the same commodity. The second market survey collected the prices of 11 locally produced foods from the nearest local market, with one food (banana) having prices collected for two different varieties. Enumerators recorded the price and weight of up to six different lots of each commodity (drawing the sample from different sellers). The market price survey was carried out on two different days in each cluster, so potentially, up to 12 observations are available on the price of each of these foods for a given market.

The unit values were obtained from a closed interval consumption recall. After an initial interview to signal the start of the consumption recall period, enumerators revisited the households approximately two weeks later and asked respondents to recall the value and quantity of all purchases, gifts, and own-production made since the initial interview. This recall covered 36 categories of food and 20 categories of other frequent expenses.[4] The unit values are calculated as the ratio of purchase values to purchase quantities. The purchase quantities were recorded in metric units, unlike the production quantities where a variety of measuring units were allowed (e.g., sacks, heaps, and bunches).[5]

The "picture method" data come from price opinions that were gathered from each household for 15 food items (including beverages) and for three tobacco products. Since six of the food items were alternate specifications of a particular food (e.g., a bottle and a can of soft drink), the pictures refer to nine categories of food. On average, these nine foods comprise 30 percent of the household's *total* consumption expenditure, with individual budget shares ranging from 11 percent (sweet potato) to one percent (flour, biscuits, and soft drinks). Central to the enumeration process, respondents were shown a series of 18 high-quality photographs (in A4 format). These photographs had been taken by professionals and showed each of the food items, presented in the typical bundle, pile, or package that is found in markets in PNG. For foods where there could be some confusion about the size of the items shown, a box of matches was included in the photograph so that respondents could put the item into perspective. Examples of these photographs are shown in Figure 1, for the four items with the largest budget shares – sweet potato, banana, betelnut (a mild narcotic, like *pan*), and rice. These photos were shown at the conclusion of the second visit to the household. Interviewers were instructed to ask the following question when showing the photograph:

> "How much does it currently cost to buy a *(Item)* like this, in the main market or store in this village/town?"

The questions about food were directed to the person in the household who typically buys most of the food, and the questions about drinks, betelnut and tobacco to the person who makes most of these purchases.

III. Unit Values, Prices and Pictures

In summary, our data collection effort provides us with three different measures of price (market prices, picture prices and unit values) for nine foods (sweet potato, banana, rice, betelnut, flour, biscuits, canned fish, soft drink, and beer). In this section, we seek to assess the quality of unit values relative to that of picture prices. To do so, we first examine the degree of correspondence of each series with market prices.

To guard against outliers affecting the results of these comparisons, the original survey forms were re-examined and cases of data entry errors and obvious miscoding (e.g., kilograms entered as grams) were removed or rectified. As a further defence against the effect of outliers, we followed the rule of Cox and Wohlgenant (1986) and trimmed the sample by removing unit values and price opinions more than five standard deviations from their respective means. This procedure removed 23 unit values and 25 price opinions, which amounted to proportionately trimming more of the unit values because there were only 4550 of them, compared with 9100 observations on price opinions.

Even after proportionately trimming more outliers from the unit value series, unit values are noisy and biased measures of market prices. The correlations between household-specific unit values and market prices range between 0.38 and 0.59 for sweet potatoes, bananas and rice, the three foods with the largest budget shares.[6] Examining deviations from the 45-degree line in price plots also demonstrates the low correlations for the major food commodities (Figure 2). The correlations for the major food commodities, however, are even higher than those for the six other, more minor food commodities ($\bar{r} = 0.37$ -- results not shown).[7] In addition to the greater variability, unit values also appear to be biased measures of market prices. Using the ratio of the means of the two price series, $\bar{x}_{uv} / \bar{x}_p$, as a measure of bias, the average unit value overstates the average market price by about 30 percent for sweet potato and banana, the two most common locally produced foods.

In contrast to unit values, picture prices provide a better measure of market prices. When using the same households as the unit value analysis, the scatter plots of market prices and picture prices are distributed more symmetrically around the 45-degree line (Figure 2). Moreover, the ratio of means of the two price series, $\bar{x}_{pp} / \bar{x}_p$, is much closer to one, ranging from 0.94 to 1.01. The picture prices also tend to have a higher correlation with market prices. Those for the three major food commodities range from 0.48 to 0.79, higher than those based on unit value-market price correlations. The average picture price-market price correlation coefficients for the six more minor food commodities are also higher, $\bar{r} = 0.64$ (compared with $\bar{r} = 0.37$ for the unit values).

There are several reasons why the picture prices might appear to be better measures of market prices. First, it could be that unit values, for some reason, are subject to more reporting error. Second, it has been shown that unit values contain quality effects, an additional source of variability (Deaton, 1990). Finally, it also may be that the specification of each food shown in the pictures coincides better with that used in the market price surveys. The unit values reported by households, on the other hand, even if they were reported without error, could be referring to commodities that differ in some systematic way from those collected in the market price survey. The differences could arise from differences in brand or package size.[8]

By examining Figure 2, it seems possible that a few households disproportionately generate much of the bias. To see how important this source of bias is, we follow a common practice of much research by replacing household-specific unit values with their cluster averages. The use of cluster-level (or even more aggregated) medians can give even more defence against the effect of outliers (Deaton and Tarozzi, 2000). In fact, when we use aggregated unit values (cluster-level averages), the correlation between unit values and market prices improves, although the unit values still tend to be noisier measures than the picture prices (Table 1, columns 6 and 7). For all nine foods, the correlation with market prices is either as high or higher for the price opinions than it is for the unit values. The average correlation of cluster-level unit values and market prices, across all nine foods, is only 0.63; while the average correlation for picture prices is 0.77.[9]

While improving the correlations with market prices, averaging by cluster does not remove the bias that occurs when unit values are used to calculate average market prices (Table 1). On average, the mean price for each food and the mean of the cluster-level unit values for the same food differs by 14 percent (this is calculated for each food as: $\left|\bar{x}_{uv} - \bar{x}_p\right| / \bar{x}_p$). Moreover, there are large differences among the commodities. For example, for canned fish there is almost no error. In contrast, there is a 40 percent difference for banana.

When compared to the low correspondence between unit values and market prices, the correspondence between picture prices and market price is higher. The average error is only 6 percent (Table 1, columns 1 and 3). The maximum price difference for any commodity is only 18 percent (for betelnut—row 7). Hence, the conclusion that unit values are more biased measures of average market prices holds even for the cluster-level estimates.

In addition to being a biased and noisy measure of market prices, there is a further statistical problem with unit values which becomes apparent when the cluster means are formed. A cluster mean unit value is available only when at least one household in that cluster made a purchase during the recall period. When there are no households making such a purchase, a sample selection problem occurs. In the case of some commodities, this can be a fairly serious problem. For example, in our sample, rather than the expected sample of 120 clusters, there are only 63 clusters with an average unit value for beer and 92 clusters with one for banana.[10] How serious this sample selection problem would be elsewhere is likely to depend on the length of the survey recall period, with longer recalls allowing more households to record a purchase.[11] In contrast to the unit values, the picture prices are much more widely available, with the most for any food being four clusters having all households with missing prices opinions. Thus, the method of obtaining opinions about prices rather than just relying on purchase behaviour can, potentially, capture the full range of spatial price variation in a sample.

IV. The Effects of the Alternative Price Collection Methods

In this section, we seek to measure the impact of using the alternative prices series as proxies for market prices. To do so, in the next subsection, we examine how using unit values (compared to using picture prices) will affect estimates of the poverty line and a number of different aggregate measures of poverty. In the following subsection, we do the same for price elasticity estimates and assess the implications for tax policy analysis.

Price Collection Methods and Poverty

Existing poverty lines for PNG are based on the market prices collected by the survey (World Bank, 1999). Specifically, the cost of buying a basket of foods that provides 2200 calories per day was calculated for five regions: the National Capital District (NCD), the South Coast, the Highlands, the North Coast, and the New Guinea Islands. Rural and urban areas within each region are combined because the sample usually had only one urban cluster per region and there are no rural clusters in the NCD.[12] The regional average prices used to calculate the cost of the poverty line basket of foods were themselves calculated from the cluster-level averages of the market prices, which have been described in Table 1.[13]

In this section of the paper we follow the above procedures used to calculate the food poverty line in PNG, but work instead with the unit values and price opinions. The aim of this replication is to construct alternative poverty lines, to see what impact the use of a different source of price information would have on measured poverty. One constraint is that while the poverty line contains 35 foods, there are only nine foods with data on both price opinions and unit values. These foods, however, contribute almost one-half of the value of the food poverty line. Thus, our experiments are, effectively, varying only one half of the value of the food poverty line, so the measured effect of different price collection methods on estimated poverty may be, if anything, understated.

The regional food poverty lines that result from using the market prices, unit values and price opinions are illustrated in Figure 3. When market prices are used, the food poverty lines range from K235 per year in the North Coast region to K626 in the NCD, and have a population-weighted average of K330.[14] While the existing poverty lines for PNG include a non-food allowance, which is equivalent to between one-third and one-half of the value of the food poverty line, we ignore that here because our different price information is only for foods.

The food poverty line is consistently overstated when unit values are used as the measure of price (Figure 3). In the NCD and South Coast regions, the use of unit values overstates the poverty line by a relatively modest margin, only about 10 percent. However, in the other three regions, areas containing 80 percent of the population, unit value-based analysis overstates the food poverty line by 13 to 27 percent. In contrast, the use of picture prices creates a smaller bias in poverty line estimates. In two regions, the NCD and South Coast, the use of picture prices causes the food poverty line to be understated by about 10 percent. In the other three regions, it is overstated by 4 to 11 percent. On average, the food poverty line has a proportionate error, $\left|z_i - z_p\right| / z_p$ (where z is the food poverty line and p=market prices, and i=unit values or picture prices) of 17 percent with the unit values and only 9 percent with picture prices.

When data collection methods create biased estimates of the poverty line, they also affect measures of poverty rates (Table 2). In particular, the overstatement of the food poverty line when unit values are used causes an upward bias in measured poverty rates. For example, the head-count index is estimated to be 30 percent rather than the actual figure (based on market prices) of 22 percent (rows 1 and 2). The poverty gap index is estimated as 8.9 percent rather than 5.9 percent. Thus, using unit values as a proxy for market prices causes headcount poverty to appear more than 30 percent higher, and the poverty gap and poverty severity measures to be more than 50 percent higher.[15]

In contrast, although there is also an upward bias associated with the use of the picture prices, the discrepancy is significantly smaller (Table 2, rows 1 and 3). Picture price-based estimates overstate the headcount poverty measure by only eight percent. They overstate the other two poverty measures by 15 to 17 percent. Clearly, in this respect, the picture price series provide more accurate measures of poverty in PNG, information needed by both domestic officials and international donor agencies.

Price Collection Methods, Price Elasticity Estimates, and Indirect Tax Analysis

In this section we report the results of using the three different price measures to calculate own- and cross-price elasticities of demand. In developing countries, pricing policy plays the same central role in fiscal policy that income tax and social security plays in developed countries (Deaton, 1989). The matrix of price elasticities, which is needed to estimate the revenue effects of price reforms, can therefore provide fundamental information to governments.[16] Hence, it is important to establish what bias might occur when elasticities are calculated from either unit values or picture prices if estimates from prices based on the preferred data collection method (that is, market price surveys) are not available.

Although we have the three measures of price for nine different foods, we focus attention on the three major staples; sweet potato, banana, and rice.[17] These three foods comprise over one-fifth of total household consumption expenditures and supply about 45 percent of calories to households. In addition to their consumption and nutritional importance, these three foods have some policy significance because until recently rice was imported duty free, whereas all other food imports were subject to tariffs. But following a switch to a Value-Added Tax (VAT), rice is now taxed at the same ten percent rate as other imported goods. In contrast, sweet potato and banana effectively fall outside of the tax net because the farmers and traders who sell them in informal markets are not registered for the VAT.

There are 11 clusters with no market price survey data for either sweet potato or banana, so the demand system is estimated on the remaining 109 clusters (containing 1018 households). This reduced sample highlights one advantage of the picture method, because there would be only two clusters with missing data if only the picture prices were used. It is also notable that of the 109 clusters, only 86 have at least one household making purchases of either sweet potato or banana (the total number of purchasing households is ca. 350). Thus, we are forced to rely on methods of imputing unit values for those households and clusters that do not have any available.

The base model uses market prices and a "share-log" functional form (Deaton, 1989):

$$w_i = a_i + b_i \ln x + \sum q_{ij} \ln p_j + \tilde{\mathbf{a}}'\mathbf{z} + u_i \qquad (1)$$

where w_i is the share of the budget devoted to good i, x is total expenditure, p_j are the prices and **z** is a vector of other household characteristics: (log) household size, the share of the household in seven demographic groups: males and females 0-6 years, 7-14 years, 15-50 years, and over 50 years (males excluded), dummy variables for whether the household head was either female or employed in the formal sector, and regional and quarterly dummy variables. An advantage of the functional form in equation (1) is that it is able to treat zero and non-zero consumption in the same way. While there is a literature on censored demand systems, this is not needed here; the analysis of tax and subsidy reform relies on unconditional demand functions because the revenue effect of a tax increase does not depend on whether demand changes take place at the extensive or intensive margins (Deaton, 1990). The price elasticities for equation (1) are given by:

$$e_{ij} = \left(q_{ij} / w_i\right) - d_{ij}, \qquad (2)$$

where $\ddot{a}_{ij}$ is the Kronecker delta (=1 if $i=j$, 0 otherwise) and budget shares are evaluated at their mean values.

The most common empirical strategy for using unit values is to simply replace the prices in equation (1) with unit values. Most of the variation concerns how analysts deal with the missing unit values and with the choice of leaving unit values at household level or aggregating them to cluster level. We use the following two methods:

UV1: using household-specific unit values, with missing unit values replaced by the mean unit value calculated across other households in the same region and season (following Minot, 1998);

UV2: using cluster median unit values, in place of both household-specific and missing unit values. This follows several studies that use averages, but with the median chosen for its robustness to outliers.

We also apply these same two methods to the picture prices, denoting them PP1 and PP2.

In addition to replacing unobserved prices with some form of unit value (as in UV1 and UV2) and estimating equation (1) and then getting elasticities from equation (2), we also use the procedures developed in Deaton (1990). The Deaton procedure uses a two-equation system of budget shares (w_{Gic}) and unit values (v_{Gic}) that are both functions of the *unobserved* prices, (p_{Hc}):

$$w_{Gic} = a^0_G = b^0_G \ln x_{ic} + g^0_G \cdot z_{ic} + \sum_{H=1}^{N} q_{GH} \ln p_{Hc} + (f_{Gc} + u^0_{Gic}) \qquad (3)$$

$$\ln v_{Gic} = a^1_G = b^1_G \ln x_{ic} + g^1_G \cdot z_{ic} + \sum_{H=1}^{N} y_{GH} \ln p_{Hc} + u^1_{Gic} \qquad (4)$$

In addition to the variables previously defined, f_{Gc} is a cluster fixed-effect in the budget share for good G, u^{0}_{Gic} and u^{1}_{Gic} are idiosyncratic errors, and the i indexes households, the G and H index goods, and the c indexes clusters.

Deaton's method recognises that the data are collected on *clusters* of households that are presumed to face the same market prices. The intra-cluster variation in budget shares and unit values is used to identify the effect of income and other household characteristics on both the quantity and quality demanded. The first-stage, within-cluster regressions are consistent even in the absence of market prices, which are treated as fixed effects. Any residual variation in unit values (and covariance with budget share residuals) is assumed to reflect measurement error, and the first-stage regression residuals give an empirical estimate of these errors. In the second stage of the procedure, a between-clusters errors-in-variables regression is applied to the (adjusted) average budget shares and unit values, which have been purged of household characteristics at the first stage. If it were not for the effect of prices on cluster-wide quality variation, the parameters estimated at the second stage would be sufficient for calculating price elasticities. Instead, a separability theory of quality (Deaton 1988) has to be used to identify the price effects at the third and final stage. An important feature of the procedure is that it depends on a large number of clusters (rather than a large number of households) for its consistency properties.

When comparing the elasticity estimates from the five price proxy series and methods (UV1, UV2, PP1, PP2, and the Deaton method) with those that are based on market prices, both picture price series (PP1 and PP2) create the estimates with the least bias (Figure 4). The point estimates of the elasticities estimated from picture price methods (particularly those using the cluster-medians--PP2) are close to those of the market price-based estimates. Also, the confidence intervals have a high degree of overlap.

There is less overlap for the two simple unit value procedures, UV1 and UV2, and for that of the Deaton method (Figure 4). For example, in the case of the estimates of the own-price elasticity of demand for sweet potato, the market price-based estimate is -1.33±0.09. When household-level unit values are used, however, the estimated elasticity is much lower in an absolute value sense (-1.00±0.08). When cluster median unit values are used (UV2), the absolute value of the estimated elasticities are even lower (-0.77±0.10). Moreover, while the Deaton procedure calculates point estimates of the own-price elasticities for sweet potato and rice that are relatively consistent with the estimates from market prices, it does a poor job of estimating the own-price elasticity for banana (giving a point estimate of -2.2 rather than -1.0). There is also considerable imprecision in the Deaton estimates. The imprecision, however, is not surprising because Deaton's method essentially reduces to a between-clusters regression, and, in our sample, there are rather fewer clusters available.

Estimates of cross price elasticities, also important in indirect taxation analysis, are likewise affected adversely by the use of unit values. Although there are too many cross-price elasticity estimates to display individually, the aggregate bias (AB) can summarize the performance of each method. Let $\mathring{\mathbf{a}}$ be the vector of elasticities calculated from the market price data and $\hat{\mathbf{a}}$ the corresponding elasticity vector from unit values or picture prices, so that the bias is $\hat{\mathbf{a}} - \mathring{\mathbf{a}}$, and $AB = (\hat{\mathbf{a}} - \mathring{\mathbf{a}})'(\hat{\mathbf{a}} - \mathring{\mathbf{a}})$, which is the sum of squared biases. The aggregate bias is calculated for the own-price elasticities alone (AB1) and for the full system of own- and cross-price elasticities (AB2).[18]

According to our results, the aggregate bias in the own-price elasticities is lowest (AB1=0.048) when the estimation uses cluster medians of the picture prices (Table 3, column 1). When the cross-price elasticities are included in the aggregate bias calculation (AB2), the use of household-specific picture prices performs best (AB2=0.904—column 2). It is notable that the bias estimates for either procedure using picture prices are less than 35 percent of those for the similar procedure using unit values. The correlation of the picture price elasticities (PP1 and PP2) with the market price elasticities is also higher (0.94-0.96) than is the correlation for UV1 and UV2 (0.67-0.80—column 3). The Deaton procedure does worst in the aggregate bias calculations, although the correlation between the elasticities from this procedure and those from the market prices is higher than for one of the naïve unit value procedures (UV2). [19]

The bias in the elasticities calculated from naïve unit value procedures could affect public policy decisions. One obvious use of the price elasticities is for deciding on the direction of marginal tax reform (Deaton and Grimard, 1992). The last three columns of Table 3 contain estimates of the social cost-benefit ratios, l_i of a marginal increase in tax on each of the three foods, calculated from:

$$l_i = \frac{w_i^e \big/ \tilde{w}_i}{1 + \frac{t_i}{1+t_i}\left(\frac{q_{ii}}{\tilde{w}_i} - 1\right) + \sum_{k \neq i} \frac{t_k}{1+t_k} \frac{q_{ki}}{\tilde{w}_i}} \tag{5}$$

where t_i is the tax rate on good i (0.1 for rice and 0 for the others), q_{ki} is the log price derivative of the budget share (from equation (1) or (3)), and the average budget shares w_i^e and $\tilde{w}_i$ are defined as:

$$w_i^e = \left[\sum_{m=1}^{M} (x_m / n_m)^{-e} x_m w_{im} \right] \Big/ \sum_{m=1}^{M} x_m \tag{6a}$$

$$\tilde{w}_i = \sum_{m=1}^{M} x_m w_{im} \Big/ \sum_{m=1}^{M} x_m \tag{6b}$$

where x_m and n_m are the total expenditure and size of household m, and e is the coefficient of inequality aversion.[20] According to the calculations in Table 3, when market prices are used to estimate q_{ki}, the highest ratio of social costs to benefits occurs when there is a marginal increase in the tax on sweet potato (λ=1.47), followed by a tax on rice (λ=1.44), while banana looks like the best candidate for a tax increase (λ=1.39). But this ranking is preserved by only two of the other estimation methods: picture prices with missing values replaced by regional and quarterly means (PP1) and the Deaton procedure applied to unit values.[21] The other two unit value procedures rank rice as the best candidate for tax increases. Hence, using unit values as proxies for market prices in an optimal tax reform exercise might lead policy makers in PNG to increase a tax which is not the socially least-cost source of revenue.

Part of the poor performance of the methods that rely on unit values may reflect the sample selection problem of several clusters having no unit value available. While this is an intrinsic disadvantage of unit value methods, in some settings there might be a wider availability of unit values either because households are more reliant on purchased food or because the consumption recall period is longer. In Table 4 we explore the performance of the cluster-median and Deaton estimators on the sub-sample of 86 clusters that have unit values available for all three foods. This change in the sample coverage does, in fact, improve the relative performance of the cluster-median unit values, although the aggregate bias (AB2) is still almost twice as large for unit value-based measures when compared to those using picture prices. The Deaton method also appears to do better on this sub-sample, at least in terms of a higher correlation with the market price elasticities. Thus, unit value methods may not fail as badly as indicated in Table 3 and Figure 4, if the unit values are available for a wider range of clusters than they are in PNG.

V. Conclusions

This paper has presented evidence on the accuracy of poverty lines, poverty rates and price elasticities of demand estimated from household budget surveys. Three different measures of price have been used: average market prices as established from a market price survey, unit values, and the price opinions of householders shown pictures of specified foods. The sort of cross-sectional household survey data studied here are increasingly being used as economists try to exploit one of the few data sources in developing countries that can help provide estimates of the demand responses that are needed for evaluating tax and subsidy reforms.

Our findings suggest that unit values, whether used in naïve or improved estimation procedures, lead to biased estimates of poverty rates and biased estimates of the price elasticities that would be calculated with actual market price data. In contrast, the price opinions perform better, with both poverty estimates and demand elasticities being closer to the values established from market price surveys. Further experiments are needed but it seems worthwhile to pursue the approach of directly asking households about prices, rather than indirectly obtaining price information from unit values. Because of the biases attendant in the use of unit values, a picture could turn out to be worth far more (in terms of accurate econometric estimates) than a thousand unit values. The advantages of the picture-based method are that it provides price estimates for a much wider ranger of households than unit values can, the errors in the estimates are unlikely to be correlated with demands and the price opinions should have less quality variation because everyone sees the same picture.

Of course, based on the assumption of our paper, it would be best to collect good measures of prices by surveying local stores and markets. If one could generate good measures of market prices, then neither unit value-based methods or picture-based methods would be needed. However, for whatever reason, the logistics of collecting market prices appear to be so difficult that many surveys do not attempt this, and of those that do, some end up not using the data. If the additional logistics and expense of carrying out market price surveys have any effect on the quality of data in the rest of the survey, it could be that alternative methods are called for. In this case, our paper's results suggest the preference for collecting prices based on the opinions of respondent households shown pictures of various consumer items.

References

Ahmad, E., and Stern, N. (1984) The theory of reform and Indian indirect taxes. *Journal of Public Economics* 25(3): 259-298.

Ahmad, E., and Stern, N. (1991) *The theory and practice of tax reform in developing countries* Cambridge University Press

Alderman, H. (1988) Estimates of consumer price response in Pakistan using market prices as data. *Pakistan Development Review* 27(2): 89-107.

Beegle, K., Frankenberg, E. and Thomas, D. (1999) Measuring change in Indonesia. *Labour and Population Program Working Paper Series* 99-07. Santa Monica, CA: RAND.

Capéau, B. and Dercon, S. (1998) Prices, local measurement units and subsistence consumption in rural surveys: an econometric approach with an application to Ethiopia. *Working Paper* 98-10, Centre for the Study of African Economies, Oxford University.

Cox, T. and Wohlgenant, M. (1986) Prices and quality effects in cross-sectional demand analysis. *American Journal of Agricultural Economics* 68(4): 908-919.

Deaton, A. (1988) Quality, quantity, and spatial variation of price. *American Economic Review* 78(3): 418-430.

Deaton, A. (1989) Household survey data and pricing policies in developing countries. *The World Bank Economic Review* 3(2): 183-210.

Deaton, A. (1990) Price elasticities from survey data: extensions and Indonesian results. *Journal of Econometrics* 44(3): 281-309.

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

Deaton, A. and Grimard, F. (1992) Demand analysis for tax reform in Pakistan. *LSMS Working Paper* No. 85, Washington D.C., World Bank.

Deaton, A. and Grosh, M. (2000) Consumption. In M. Grosh and P. Glewwe (eds) *Designing Household Survey Questionnaires for Developing Countries* The World Bank, Washington, pp. 91-133.

Deaton, A. and Tarozzi, A. (2000) Prices and poverty in India. *Mimeo* Research Program in Development Studies, Princeton University.

Frankenberg, E. (2000) Community and price data. In M. Grosh and P. Glewwe (eds) *Designing Household Survey Questionnaires for Developing Countries* The World Bank, Washington, pp. 315-338.

Gibson, J. and Rozelle, S. (2002) Demand systems with unit values: comparisons with elasticities from market prices. *Mimeo* Department of Economics, University of Waikato.

Glewwe, P. (1991) Investigating the determinants of household welfare in Côte d'Ivoire. *Journal of Development Studies* 35(2): 307-337.

Le, T., Gibson, J., and Rozelle, S. (2002) Prices, unit values and food policy in Vietnam. *Mimeo* Department of Economics, University of Waikato.

Minot, N. (1998) Distributional and nutritional impact of devaluation in Rwanda. *Economic Development and Cultural Change* 46(2): 379-402.

Minot, N., and Goletti, F. (2000) *Rice Market Liberalization and Poverty in Viet Nam* Research Report 114, International Food Policy Research Institute, Washington.

Ravallion, M., and Dearden, L. (1988) Social security in a "moral economy": An analysis for Java. *Review of Economics and Statistics* 70(1): 36-44.

Saunders, C, and Grootaert, C. (1980) Reflections on the LSMS group meeting. *Living Standards Measurement Study Working Paper* No. 10, The World Bank.

Wood, D., and Knight, J. (1985) The collection of price data for the measurement of living standards. *Living Standards Measurement Study Working Paper* No. 21, The World Bank.

World Bank. (1999) *Papua New Guinea: Poverty and Access to Public Services* Washington, DC: World Bank.

Table 1: Descriptive statistics for cluster-level market prices, unit values and price opinions

	Mean market price[a]	Mean unit value[a]	Mean price opinion[a]	No. of clusters with data on[b]		Correlation with market prices	
				Unit values	Price opinions	Unit values	Price opinions
Sweet potato	43.9	59.0	42.5	93	118	0.74	0.74
Banana	54.2	75.9	51.3	92	118	0.65	0.71
Rice	114.7	107.3	115.5	114	118	0.75	0.93
Flour	143.6	114.9	158.3	95	116	0.43	0.72
Biscuits	444.4	450.0	452.4	112	118	0.50	0.83
Canned fish	432.7	437.0	422.7	115	118	0.42	0.56
Betelnut	510.8	566.0	419.9	107	117	0.63	0.64
Soft drink	272.8	263.3	287.9	100	118	0.73	0.91
Beer	558.3	507.0	586.8	63	116	0.86	0.93

[a] Toea per kilogram, as calculated from cluster-level averages. 130 toea=US$1 in 1996.
[b] Out of a possible *n*=120.

Table 2: Aggregate food poverty measures for Papua New Guinea, 1996

Cost of poverty line food basket calculated from:	Headcount index	Poverty gap index	Poverty severity index
Market prices	22.0	5.9	2.4
Unit values	30.0	8.9	3.8
Price opinions	23.8	6.8	2.8

Note: Based on the food poverty lines in Figure 3. The poverty estimates are in terms of adult-equivalents.

Table 3: Summary Comparisons of Estimates Using Market Prices, Picture Prices and Unit Values

Data source and estimation method	AB1	AB2	Corr	Cost-benefit ratio (l_i) of tax rise for: Sweet potato	Banana	Rice
Market prices				1.47 (3)	1.39 (1)	1.44 (2)
PP1 (missing=reg/qtr mean)	0.089	0.904	0.958	1.46 (3)	1.40 (1)	1.41 (2)
PP2 (cluster medians)	0.048	1.448	0.938	1.45 (2)	1.40 (1)	1.47 (3)
UV1 (missing=reg/qtr mean)	0.369	3.323	0.804	1.49 (3)	1.40 (2)	1.35 (1)
UV2 (cluster medians)	0.653	4.844	0.669	1.48 (3)	1.42 (2)	1.34 (1)
Unit Values (Deaton method)	1.415	7.775	0.737	1.53 (3)	1.34 (1)	1.43 (2)

Note: AB1 is the aggregate bias on the own-price elasticities, AB2 is the aggregate bias on own- and cross-price elasticities, "Corr" is the correlation between the elements of the elasticity matrix and the market price elasticities. The calculations exclude the elasticities for "other goods" derived from the adding-up and homogeneity restrictions.

PP refers to "picture prices" and UV to "unit values". The cost-benefit ratio, l_I is calculated from equation (5), using an inequality aversion parameter, ε=0.5. The values in ( ) are the good's rank in terms of l_i, where "1" denotes the good with the lowest cost-benefit ratio from a marginal tax increase.

Table 4: Results for the sub-sample with each cluster having a unit value available[a]

	Price Elasticities of Demand Calculated From: Market Prices	Cluster Medians of Picture Prices	Cluster Medians of Unit Values	Deaton Procedure
Own-Price Elasticity for Sweet potato	-1.19	-1.30	-0.90	-2.05
Banana	-1.12	-0.70	-1.34	-2.16
Rice	-1.59	-1.77	-1.95	-3.00
Aggregate Bias (own-price elasticities only)[b]		0.22	0.26	3.53
Aggregate Bias (own- and cross-price elasticities)[b]		1.23	2.07	6.88
Correlation with elasticities from market prices[b]		0.89	0.88	0.95

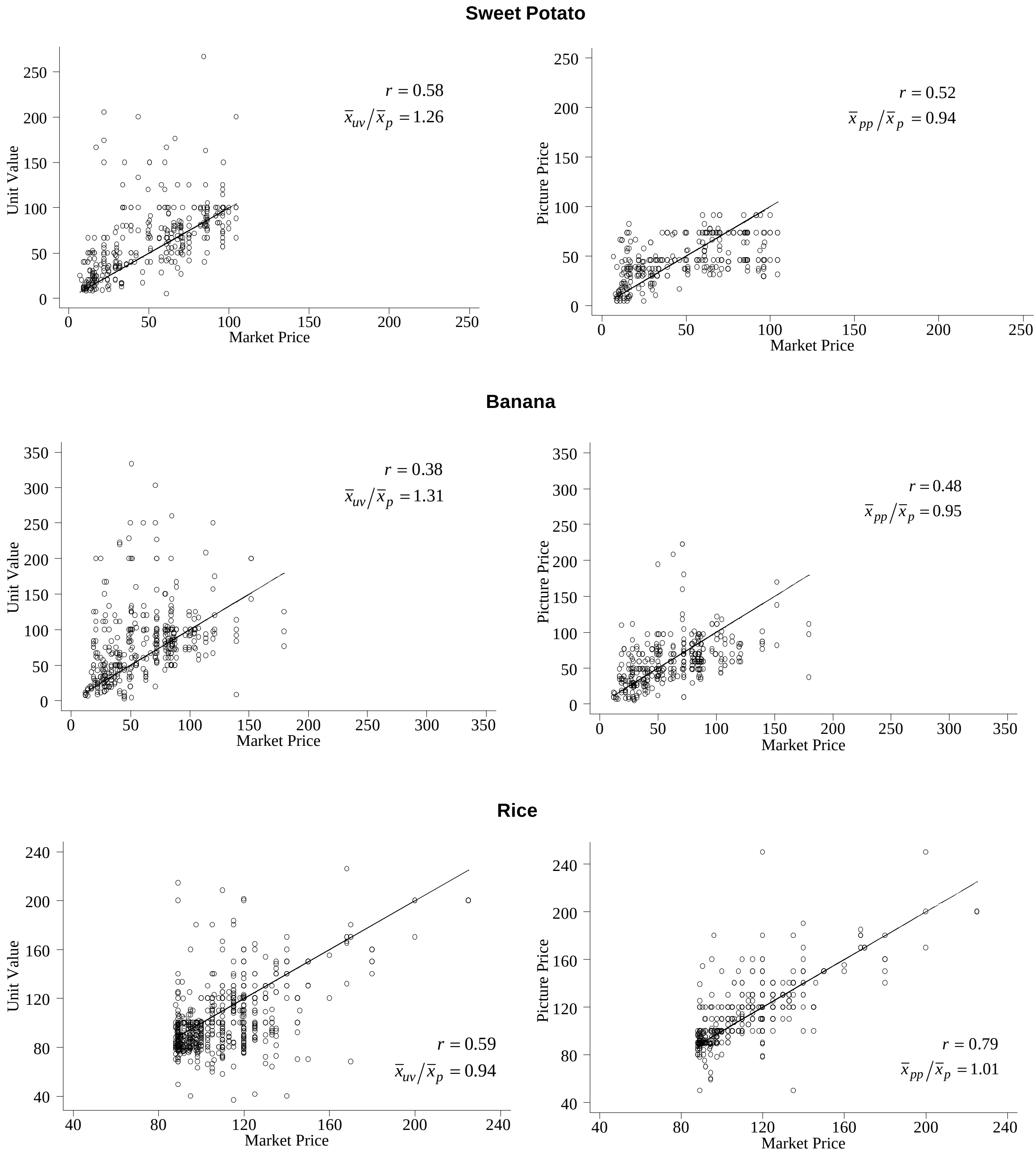
Notes

[a] 86 clusters, containing 755 households.

[b] Calculations exclude the elasticities for "other goods" derived from the adding-up and homogeneity restrictions.

Figure 1: Examples of Photographs Used When Eliciting Price Opinions

Figure 2: Comparisons of Market Prices and Household-Specific Unit Values and Picture Prices



Note: Prices are in toea per kilogram (130 toea=US$1 in 1996). The 45° line shows the points where market prices equal unit values (or picture prices).

Figure 3: Regional Food Poverty Lines

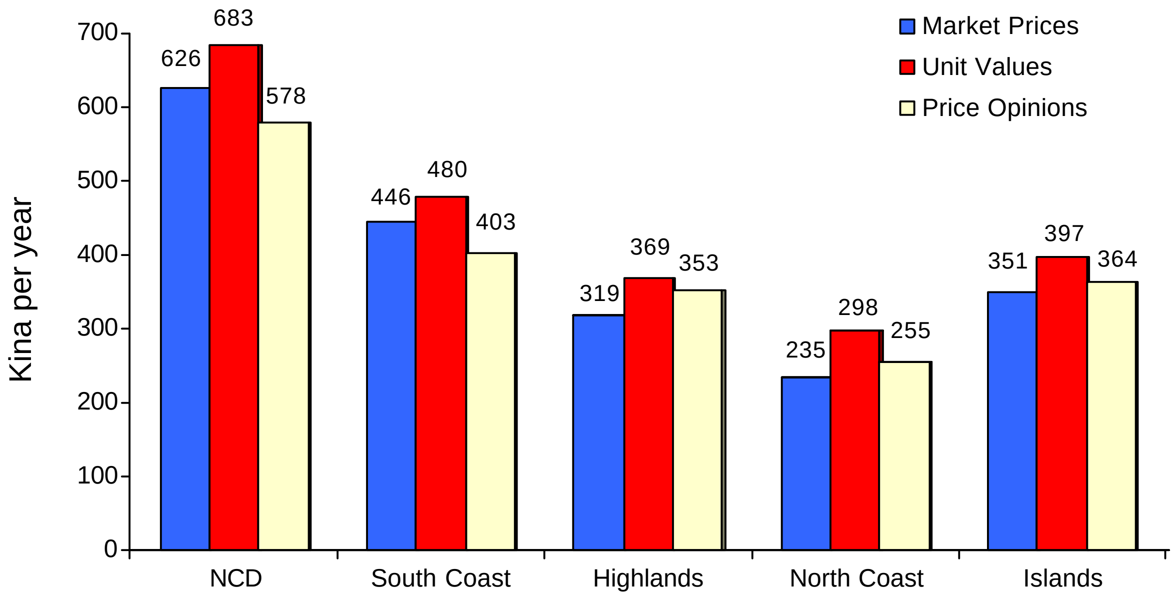
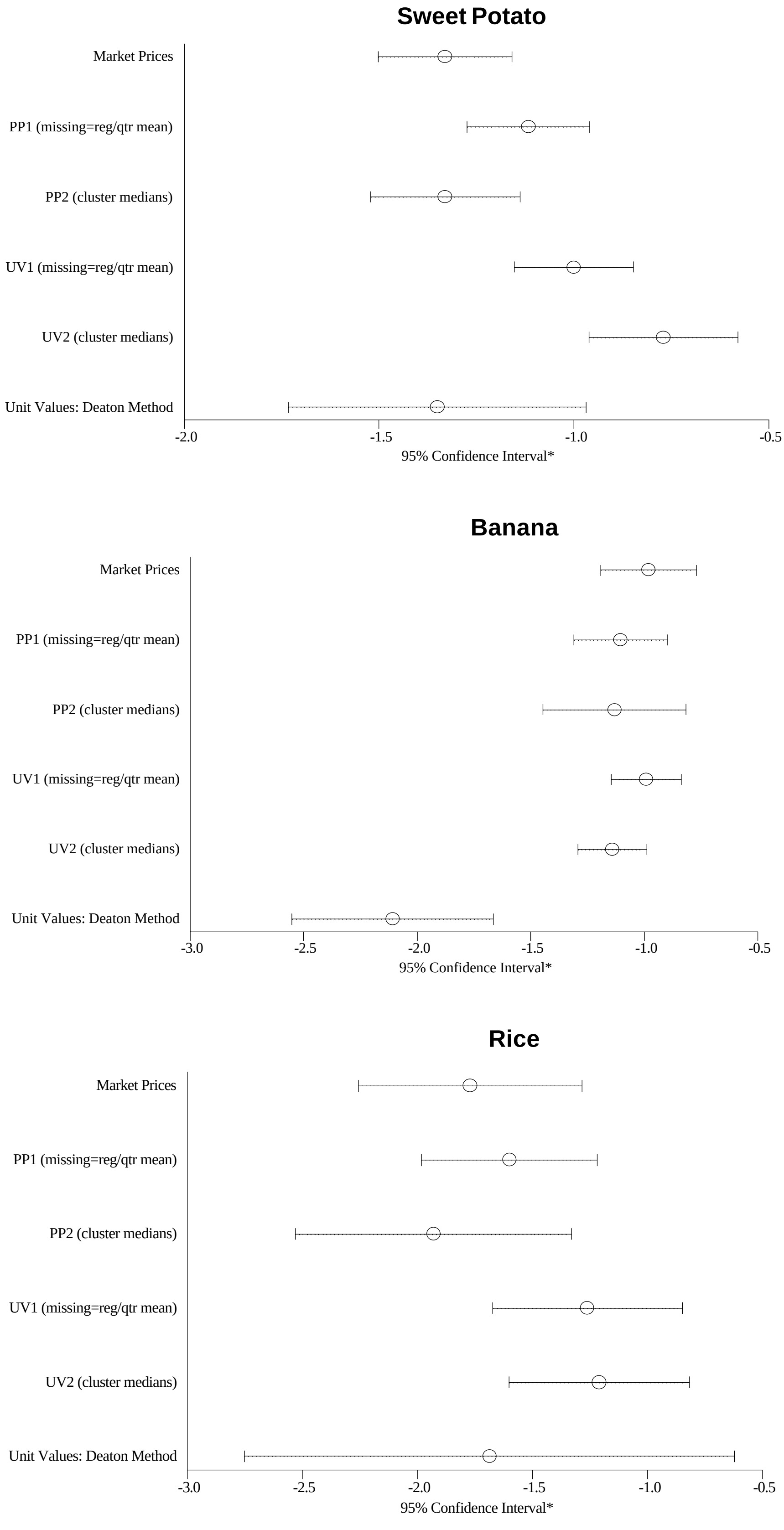


Figure 4: Own-Price Elasticity Comparisons for Market Prices, Picture Prices and Unit Values



Note: *68% Confidence Interval (± 1 std error) for the elasticities from the Deaton method

Notes

[1] In one example, the price of canned tomato paste had to be used as a substitute for all non-food prices (which were poorly measured) in Côte d'Ivoire (Glewwe, 1991).

[2] In some applications it is also possible to substitute assumptions for data. For example, researchers often use additivity assumptions, such as in the linear expenditure system, to get price elasticities from household budget data, without any prices required. But additive preferences imply that expenditure and own-price elasticities are roughly proportional, forcing a tradeoff between equity and efficiency, and leading to recommendations of uniform rates of commodity taxes regardless of the patterns in the data (Deaton, 1997).

[3] One reason why this was never pursued in the field may have been that subsequent LSMS papers were critical of the idea, calling it 'novel but risky' and suggesting that it would be subject to numerous potential sources of bias (Wood and Knight, 1985). We believe that our development of the idea, based on a representative sample of households each shown a defined specification (in the form of a photograph), overcomes several of these biases. We are also aware of prices being collected from 'key informants' (the Ibu PKK) in the IFLS, although comparisons of those prices with the prices collected from market surveys do not seem to be available.

[4] In addition to these short period measures of consumption, the estimate of household's total expenditure used an annual recall of 31 categories of infrequent expenses and an inventory of durable assets, which provides estimates of the flow of annual services from durables and dwellings.

[5] To provide some standardization, all of the households had been given empty sacks with marked graduations for recording their production from food gardens. Volumetric conversion factors were then established for each food.

[6] These correlations should not be seen as either atypically low or reflective of the unusual conditions in PNG. A comparison of market prices and unit values for 33 items in the 1997-98 Vietnam Living Standards Survey (VLSS) yields an average correlation of only 0.25 (Le, Gibson and Rozelle, 2002). Using a more restricted set of foods, and data from the 1992-93 VLSS, Deaton and Grosh (2000) report a median correlation of 0.34. A caveat to both comparisons is that the unit values in the VLSS are meant to refer to the previous 12 months while the market prices are from the month when the household was actually surveyed.

[7] The correlations with market prices are even lower for the unit values applied to self-produced foods $(\bar{r} = 0.35)$ and for the unit values for gifts received $(\bar{r} = 0.36)$. There is also little agreement amongst the different types of unit values: for those households who both purchased and produced either sweet potato, banana or betelnut, the average correlation between the two types of unit values is only 0.26. For those who both purchased and received gifts, the average correlation is 0.43.

[8] We can test this final argument, at least in the case of rice, because the brand used for the market price survey ("Trukai") accounted for 86 percent of rice sales in PNG in 1996 and most of those sales were for the specified 1 kilogram pack size. (We are grateful to Neville Whitecross of Trukai Industries, Port Moresby, for these details.) The correlation between unit values and market prices is almost the same for the households who report purchasing only one kilogram of rice during the recall period (r=0.61) as it is for other households (r=0.57). Thus, even when the pack size for the unit value report corresponds to that used for the market price survey, there is a low correlation between unit values and market prices, suggesting that reporting errors are important. The intracluster correlation in the rice prices collected from the market survey is 0.82 so variation in the prices charged by different trade stores within each cluster is unlikely to account for all of the discrepancy with unit values. Moreover, this variation in market prices within a cluster would also affect the calculated reliability of the picture prices, so it cannot account for the relatively poor performance of the unit values.

[9] The average correlation is no higher (r=0.63) if a more broadly defined unit value is formed, based on the ratio of the combined value of purchases, net gifts received and own-production to the combined quantity.

[10] This lack of unit values particularly affects rural areas. For example, a unit value for beer is available for 35 of the

40 clusters in the capital city but in only 28 of the 80 clusters elsewhere. Hence, the spatial distribution of prices may not be measured in a reliable way when unit values are relied on as the proxy for market prices.

[11] However, even without this sample selection issue, there is still bias in the unit values. For example, in the 93 clusters where a unit value for sweet potato is available, the average market price is 46.8 toea per kilogram (slightly above the average across all clusters), which is still 20 percent below the mean unit value for those same clusters.

[12] An analysis of covariance also showed that urban/rural price differentials within regions were less important than inter-regional price variations (World Bank, 1999).

[13] The NCD is an exception, with the average price formed directly from the raw prices rather than from the cluster-level prices. This reflected the assumption that there is less need for the average to reflect the spatial distribution of prices within a city than there is in larger geographical regions (World Bank, 1999).

[14] This is equivalent to US$250 per year, and refers to adult-equivalents rather than per capita.

[15] Capéau and Dercon (1998) find that in rural Ethiopia, unit values cause poverty measures to be more than one-fifth higher than when using other price data.

[16] The elasticities are not needed for evaluating the welfare effects of *marginal* tax and subsidy reforms. The existing demand structure, and some social weights for aggregating the effects across households, provides sufficient information when price changes are small (Ahmad and Stern, 1984).

[17] All of the other foods and non-foods are aggregated into a composite fourth commodity in the demand system and we assume that leisure is separable from goods demand (an assumption forced by the fact that the survey did not gather data on wage rates).

[18] For both AB1 and AB2 the calculation excludes the results for “other goods” which are simply derived from the other elasticities.

[19] To check that there was not some flaw in the programming, the market prices were passed through the STATA code for the Deaton procedure. The correlation between these elasticities and the market price elasticities reported in Figure 4 and Table 3 was 0.999.

[20] This expression for the cost-benefit ratio of a marginal tax increase is adapted by Deaton (1997) from the more usual one (see, for example, Ahmad and Stern (1984), equation (38)) and allows for both quantity and quality responses to tax-induced price changes.

[21] This finding is sensitive to the value of the inequality aversion parameter used. As ε increases, the equity effects of not taxing sweet potato and banana, which tend to be consumed by the poor, dominate the tax derivative effects and the rankings are not sensitive to differences in the price elasticities. However, attempts to econometrically estimate ε, using the approach of Ravallion and Dearden (1988), suggests that ε is likely to be close to zero in PNG.