An Analysis of Household Behaviour to Price Shocks in Vietnam: Can Unit Values Substitute Prices?[1]

Yoko Niimi[2]
University of Sussex

March 2004

Abstract:
This paper estimates a demand system of food commodities based on the Vietnamese household survey data. When estimating consumer behaviour for developing economies, one of the problems that researchers often come across is the unavailability of price data. A common practice has been to use unit values, ratios of expenditures to quantities purchased, as proxies for market prices. However, Deaton (1988) criticises the direct substitution of unit values for prices in the demand analysis because bias is likely to result as a consequence of quality effects and measurement error. In order to overcome this problem, Deaton (1988, 1990 and 1997) proposes a procedure that corrects the bias and hence enables us to obtain price elasticities in the absence of price information. Given the availability of price data in the case of Vietnam, this paper investigates the validity of Deaton's method, which has hardly been examined despite its popular application. The results demonstrate that Deaton's method fails to reproduce the price elasticities estimated with market prices and pose a question on its usefulness.

---

[1] I am grateful to my supervisor, Professor L. Alan Winters, for his valuable guidance and comments. I would also like to thank the Poverty Research Unit at Sussex for permission to access to the Vietnamese Living Standard Measurement Surveys.
[2] School of Social Sciences and Cultural Studies, University of Sussex, Falmer, Brighton, BN1 9SN, UK.
E-mail: Y.Niimi@sussex.ac.uk

1. Introduction

Despite the ongoing discussion about the link between trade liberalisation and poverty, there has been limited empirical research on how trade shocks are transmitted to poverty impacts at the household level (Winters *et al*., 2002). One possible effect of trade reforms on poverty is through price changes following, for instance, a devaluation of local currency and a reduction or elimination of tariff and non-tariff barriers (Winters, 2002). The poverty implication of such price effects can be significant, particularly when price changes take place in the agricultural sector. In the case of Vietnam, for example, the domestic price of rice, one of Vietnam's leading export commodities, increased during the 1990s in response to the *doi moi* (economic renovation) reforms including agricultural and trade reforms (Niimi *et al*., 2004). The welfare implication of the rice price increase is likely to have been considerable given that rice is the most important consumption good of Vietnamese households – on average it accounted for a 28% share in total household expenditure in 1992-93. The figure is even more striking for poor households[3] whose expenditure share of rice was 36% in the same year.

For the evaluation of the price effects, it is important to have a good understanding of how households respond towards price changes. Such knowledge is also required for policy makers for designing effective indirect tax or subsidy reforms. In order to estimate consumer behaviour, however, accurate price information is generally required. In the case of developed countries, this requirement is usually met by the availability of time-series data. In contrast, the absence of price data in developing countries is often compensated by extensive household survey data, from which unit values can be derived as a ratio of expenditure to quantity and used as a proxy for market prices (Deaton, 1988). Yet Deaton (1988) argues that unit values cannot be treated as prices because they are contaminated by quality effects and measurement error.

Nevertheless, Deaton (1988, 1990 and 1997) recognises the usefulness of unit values that contain vital information on the spatial distribution of prices and derives a procedure for recovering such information from unit values so that price elasticities can still be estimated in the absence of market price data. Although his methodology has widely been applied to household survey data of various countries (e.g. Friedman and Levinsohn, 2002; Gracia and Albisu, 1998; and Stavrev and Kambourov, 1999), there has been very little work carried out that actually tests the validity of Deaton's method (e.g. Brubakk, 1997; and Gibson and Rozelle, 2002). This is rather alarming because if Deaton's procedure does not generate accurate estimation of a demand system, any subsequent welfare analysis of price changes would simply be invalid. The 1992-93 and 1997-98 Vietnam Living Standards Measurement Surveys (VLSS)[4] data are not only rich in microeconomic issues and constitute a panel data set of over 4,300 households, but they also contain price data that were collected along with the surveys. Based on these data sets, it is possible to formally investigate Deaton's approach.

The main aim of this paper is therefore to examine the usefulness of Deaton's method for estimating demand functions for developing countries where adequate price data are often missing through its application to the case of Vietnam. The next section briefly discusses alternative ways of evaluating the welfare effect of price changes at the household level. Section 3 raises some methodological issues concerning the use of unit values in the estimation of a demand system and

[3] Poor households are defined as those whose consumption expenditure is below that is required to obtain a balanced diet of 2,100 calories per day.

[4] These surveys were carried out by the General Statistical Office (GSO) and the Ministry of Planning and Investment, with financial assistance from the United Nations Development Programme (UNDP) and the Swedish International Development Agency (SIDA) and technical assistance from the World Bank. 4,800 and 6,000 households were surveyed in 1992-93 and 1997-98 respectively and both surveys are nationally representative.

describes Deaton's procedure. Section 4 specifies demand models and variables employed. Section 5 discusses the validity of Deaton's methodology. The final section ends with concluding remarks.

2. Measuring Welfare Effects of Price Changes

The welfare implication of price changes can be examined through an analytical framework developed by Winters (2002), which explores the link between trade liberalisation and poverty. Economic growth is the key causal link between openness and poverty alleviation in the long run. In the medium and short term, liberalisation-induced shocks trickle down to households who are consumers as well as producers in the economy via their direct effects on product and factor markets and indirectly through changes in government revenues and social spending. The price channels are the main focus of this paper, although the other channels are equally important for the welfare analysis of trade reforms.

One way of analysing the impact of trade reforms through the price channels is to use an indirect utility function, in which the household's utility is written as a function of its income and prices (Deaton, 1997):

$$(1) \qquad u_h = \Psi_h(x_h, p)$$

where Ψ_h is the indirect utility function of household h, x_h is household income (or consumption expenditure) and p is a price vector of consumption goods. If, for simplicity, intertemporal aspects are ignored and saving is assumed to be zero, the welfare effect of a price change of good i can be expressed as follows:

$$(2) \qquad \partial u_h / \partial \ln p_i = \left[\partial \Psi_h / \partial l \ln x_h\right] \cdot p_i \left(q_{hi} - c_{hi}\right) / x_h$$

where q_{hi} and c_{hi} are the production and consumption of good i by household h respectively (Deaton, 1997). Note that this is based on the so-called farm household model whereby households are regarded as a production-consumption unit. Deaton (1997) refers to the last term on the right-hand side of (2) as the net benefit ratio (NBR). It is defined as the value of net sales as a proportion of income and can be interpreted as the short-term elasticity of real income with respect to the price of the commodity. If the term $[\delta\Psi_h / \delta \ln x_h]$, the marginal utility of income, is assumed to be positive, the welfare effect of a price change on households depends on their net supply position. In other words, whether trade liberalisation contributes to poverty alleviation crucially depends on how the poor earn and spend their income (Deaton, 1997).

Although equation (2) provides an explicit picture of how price changes affect household welfare, the problem with this method is that the use of the differential calculus limits the analysis to the effect of infinitesimal price changes and it assumes no household response. Hence it cannot explain how finite price changes influence households. In the case of finite changes, the magnitude of the welfare effect does not just depend on the amount of production and consumption, but also depends on how individual households accommodate price changes (Deaton, 1997). Given the observed changes in rice prices in Vietnam, it is important to undertake a behavioural analysis of household response. One way of doing so is to place a specific form upon equation (1), namely the Almost Ideal Demand System (AIDS) developed by Deaton and Muellbauer (1980b). The AIDS demand function in its budget share form can be derived as follows:[5]

[5] See Deaton and Muellbauer (1980b) for the derivation.

(3) $$w_{ih} = \alpha_i + \sum_{j=1} \theta_{ij} \ln p_j + \beta_i \ln\left(\frac{x_h}{P}\right)$$

where w_{ih} is household h's budget share of good i, p_j is the price of good j, x_h is h's total consumption expenditure, and P is a price index defined by:

(4) $$\ln P = \alpha_0 + \sum_i \alpha_i \ln p_i + \frac{1}{2}\sum_i \sum_j \theta_{ij} \ln p_i \ln p_j$$

One of the AIDS' desirable properties is that it is a flexible representation of an arbitrary demand system, which, in turn, generates an approximation to any arbitrary indirect utility function. It is also free from the restrictive assumption of homotheticity. This essentially allows the model to capture the differences in the bundle of consumption among different income groups. Another reason for the popularity is its tractability. Although the AIDS equation (3) must meet the usual properties of a demand function, namely adding-up, homogeneity, symmetry and negativity conditions, they can be ensured by satisfying the restrictions on its parameters alone except the last condition. Moreover, apart from P, the AIDS expressed as (3) is linear in the parameters, which would make its estimation much simpler. In many practical situations, by exploiting the collinearity of the prices, P is often approximated as proportional to some known price index. Most commonly used price index P^* is a Stone price index where $\ln P^* = \Sigma\, w_i \ln p_i$, which allows us to estimate (3) linearly. Given these advantages, the AIDS is employed in the present analysis of household demand behaviour. The next section draws some attention to a number of issues arising from the use of unit values in the estimation of a demand system.

3. Methodological Issues: Price vs. Unit Value

One of the problems that researchers often come across is the unavailability of price data in developing countries, which is vital for estimating a demand system. Although extensive household surveys are regularly conducted in many developing countries, these surveys rarely collect data on market prices (Deaton and Grosh, 2000). The absence of price data at the household level is usually compensated by the fact that in these surveys households are asked to report both expenditures and quantities consumed on a wide range of commodities. A common practice has therefore been to derive unit values as a ratio of expenditure to quantity and use them as a direct substitute for market prices (Deaton, 1988).

Deaton (1988) argues, however, that there are a number of reasons as to why unit values cannot be treated as if they were market prices. Firstly, since commodities are usually subject to some degree of aggregation, unit values reflect the quality of different goods within the commodity group, which is chosen by households. One obvious consequence is that there is a risk of simultaneity bias in any attempt to use unit values to "explain" demand patterns. Moreover, this choice of quality may itself be affected by price and income as consumers alter both quantity and quality of goods they buy in response to price or income changes. Each unit value is thus likely to vary less proportionately with its own price. Secondly, unit values are calculated by dividing expenditures by quantities, both of which are almost always measured with errors. As a consequence, unit values are also bound to be contaminated by measurement error and spuriously negatively correlated with quantities (Deaton, 1988).

Nevertheless, Deaton (1988) recognises the usefulness of unit values that contain vital information on the spatial distribution of prices and argues for the possibility of relying on spatial variation in prices to identify demand behaviour. He exploits the fact that households in household surveys are in general geographically clustered and assumes that all households within the same cluster face the same market prices. The basic idea of his model is that, based on the assumption of no price variation in each cluster, within-cluster variation in purchases and unit values is used to estimate the effects of incomes and household characteristics on quantities and qualities. Within-cluster variation in unit values is also used to identify the influence of measurement error and to separate it out from genuine price variation. The demand system can then be estimated on the basis of inter-cluster variation in corrected quantities and unit values (Deaton, 1988). In this way, it is possible to treat market prices as unobservable variables in Deaton's model.

Deaton's procedure is essentially based on the following two equations:

$$w_{ihc} = \alpha_i^0 + \beta_i^0 \ln x_{hc} + \sum_{m=1} \gamma_{im}^0 \cdot z_{mhc} + \sum_{j=1} \theta_{ij} \ln p_{jc} + f_{ic} + u_{ihc}^0, \tag{5}$$

$$\ln v_{ihc} = \alpha_i^1 + \beta_i^1 \ln x_{hc} + \sum_{m=1} \gamma_{im}^1 \cdot z_{mhc} + \sum_{j=1} \psi_{ij} \ln p_{jc} + u_{ihc}^1 \tag{6}$$

where w_{ihc} is the budget share of good i of household h in cluster c, x_{hc} is the household's total expenditure, v_{ihc} is the calculated unit value, and z_{mhc} is a vector of household characteristics, all of which are observed. However, we cannot observe price p_{jc}, the cluster fixed effects f_{ic}, or the two error terms u^0_{ihc} and u^1_{ihc}. It should be noted that there is no subscript h for prices as they are assumed to be common within the same cluster. The cluster fixed effects f_{ic} represent unobservable taste variation among different clusters, but they are shared by all households in each cluster and hence no subscript h for this variable either. In addition, f_{ic} must be uncorrelated with the unobservable prices so that price elasticities can be measured even when both variables are unobservable. The second error term u^0_{ihc} incorporates the usual unobservables as well as any measurement error in the share. Consequently, it will be correlated with the error term u^1_{ihc} in the unit value equation (6), which includes measurement error in unit values. However, it is assumed to be uncorrelated with all the other regressors including the fixed effects (Deaton, 1990).

As far as the unit value equation (6) is concerned, unlike the unobservable prices, unit values vary among households as each household chooses different qualities within the same commodity group. These unit values are thus not identical to the prices, but are related to them by equation (6) whereby β^1_i is the elasticity of quality with respect to total expenditure. The matrix Θ represents the response of unit values to prices, which would be an identity matrix if unit values were equal to market prices up to measurement error (Deaton, 1990). Note that the unit value equation does not contain the cluster fixed effects. Apart from quality effects and measurement error, unit value is a direct indication of price. The introduction of an additional fixed effect in the unit value equation would break this link between prices and unit values. The inclusion of the fixed effects would therefore prevent unit values giving any useful information about prices and remove any possibility of identification. (Deaton, 1997).

Based on these assumptions, the methodology proposed by Deaton (1990) basically has two following stages to the estimation of price elasticities.

First Stage:

By postulating no price variation in the same cluster, within-cluster variation in purchase patterns and unit values is used to estimate total expenditure elasticities and quality effects. All the variables in both equations are demeaned by their cluster means and ordinary least squares (OLS) are then applied to the demeaned equations. This allows the consistent estimation of all the parameters except θ_{ij} and ψ_{ij} since it removes all the prices, the fixed effects and the cluster means of the idiosyncratic errors. In addition, the residuals from each equation e^0_{ihc} and e^1_{ihc} can be used to estimate the variances and covariances of the measurement errors u^0_{ihc} and u^1_{ihc} in equations (5) and (6). They can be constructed from:

(7a) $$\tilde{\sigma}_{ij} = (n - C - k)^{-1} \sum_c \sum_{h \in c} e^0_{ihc} e^0_{jhc}$$

(7b) $$\tilde{\omega}_{ij} = (n_i^+ - C - k)^{-1} \sum_c \sum_{h \in c} e^1_{ihc} e^1_{jhc}$$

(7c) $$\tilde{\chi}_{ij} = (n_i^+ - C - k)^{-1} \sum_c \sum_{h \in c} e^1_{ihc} e^0_{jhc}$$

where n is the total number of households in the survey, n^+_i is the number of households who report purchases of good i, C is the number of clusters, and k is the number of explanatory variables. Note that the budget share equation (5) includes all the households in the survey, while the unit value equation (6) includes only those households who purchase good i (and thus unit values can be derived).

Second Stage:

The main task of the second stage is to use between-cluster information in the data to estimate the price responses. Here the effects of total expenditure and household characteristics obtained from the first-stage estimation are netted out, and cluster averages of the "corrected" budget shares and unit values are computed:

(8) $$\tilde{y}^0_{ic} = \frac{1}{n_c} \sum_{i \in c} (w_{ihc} - \tilde{\beta}^0_i \ln x_{hc} - \sum_{m=1} \tilde{\gamma}^0_{im} Z_{mhc})$$

(9) $$\tilde{y}^1_{ic} = \frac{1}{n^+_{ic}} \sum_{i \in c} (\ln v_{ihc} - \tilde{\beta}^1_i \ln x_{hc} - \sum_{m=1} \tilde{\gamma}^1_{im} Z_{mhc})$$

Superimposed tildes indicate the estimates from the first-stage regressions. As the number of observations in the first-stage regressions increases, the estimates tend to their true values, and $\tilde{y}^0_{ic}$ and $\tilde{y}^1_{ic}$ also tend to the true cluster means which are:

(10) $$y^0_{ic} = \alpha^0_i + \sum_{j=1} \theta_{ij} \ln p_{jc} + f_{ic} + u^o_{ic}$$

(11) $$y^1_{ic} = \alpha^1_i + \sum_{j=1} \psi_{ij} \ln p_{jc} + u^1_{ic}$$

where u^0_{ic} and u^1_{ic} are the cluster means of the errors in equations (5) and (6). The cluster size in household surveys is typically the same and not large, even when the total number of households in surveys increases. As a result, although averaging over clusters will reduce the effects of the measurement error, it will not eliminate the bias. Both the covariance and the variance must thus be corrected appropriately. This can be done by the use of a standard errors-in-variable estimator, whereby the error covariance and variance estimated at the first stage are used to capture both the measurement error in unit values and any possible correlation between the measurement error in the budget share and unit value equations.

By the definition of OLS, the between-cluster estimation of equations (10) and (11) would be $B_{OLS}=S^{-1}R$ where S is the variance-covariance matrix of y^1_{ic} and R is the covariance matrix of y^1_{ic} and y^0_{ic}. They are respectively:

(12a) $$\{S\}_{ij} = \operatorname{cov}\left(y_{ic}^{1}, y_{jc}^{1}\right)$$

(12b) $$\{R\}_{ij} = \operatorname{cov}\left(y_{ic}^{1}, y_{jc}^{0}\right)$$

Yet S is likely to overestimate the variance-covariance matrix of true prices because it includes the effects of the measurement error in (6). R is similarly contaminated by any covariance in the measurement error between the two equations (5) and (6). The situations can be illustrated as:

(13a) $$S = \Psi \mathrm{M} \Psi' + \Omega N_{+}^{-1}$$

(13b) $$R = \Psi \mathrm{M} \Theta' + \mathrm{X} N^{-1}$$

where M is the variance-covariance matrix of the unobservable true prices, $N_{+}^{-1} = \operatorname{plim} C^{-1} \sum_{c} D(n_c^{+})^{-1}$, $D(n_c^{+})^{-1}$ is a diagonal matrix formed from the elements of n^{+}_{ic}, and N^{-1} is the corresponding matrix formed from the n_c's. The matrices Ω and X are the variance-covariance matrices for u's:

(14a) $$\{\Omega\}_{ij} = \operatorname{cov}\left(u_{ihc}^{1}, u_{jhc}^{1}\right) = \omega_{ij}^{11}$$

(14b) $$\{\mathrm{X}\}_{ij} = \operatorname{cov}\left(u_{ihc}^{1}, u_{jhc}^{0}\right) = \chi_{ij}^{10}$$

The elements of these matrices are those defined in equation (7) above. In order to eliminate the effects of the measurement error, the OLS estimator must be corrected by removing the second terms on the right hand side of (13). Hence the correct estimator would be:

(15) $$\tilde{B} = \left(\tilde{S} - \tilde{\Omega} \tilde{N}_{+}^{-1}\right)^{-1} \left(\tilde{R} - \tilde{\mathrm{X}} \tilde{N}^{-1}\right)$$

where the tildes for S and R indicate that the matrices are evaluated using the estimates from (8) and (9) rather than from (10) and (11), while for Ω and X they are the estimates from (7). Finally, the probability limit of the matrix B in the presence of quality effects is:

(16) $$\operatorname{plim} \tilde{B} = B = (\Psi')^{-1} \Theta'$$

It should be noted that what is estimated is still not the matrix of price responses, Θ, but the matrix that is still contaminated by the quality effects through Ψ. Hence the effect of price on budget share has to be extracted from the ratio (16). This can be done with the theory linking the price elasticity of quality to the usual price and total expenditure elasticities proposed by Deaton (1988). Deaton (1988) defines quality as a property of commodity aggregates. Based on this definition, if q_{ihc} denotes a vector of quantities of each item within the group i consumed by household h in cluster c, a group quantity index Q_{ihc} can be defined as:

(17) $$Q_{ihc} = k_i \cdot q_{ihc}$$

where k_i is a vector used to add the quantity of individual items in the group. Hence it would be a vector of ones if quantities are reported as weights and Q_{ihc} would be the total physical quantity of the commodity group i measured in kilos. Similarly, the price vector can be expressed as:

(18) $$p_{ic} = \pi_{ic} \cdot p_i^0$$

where π_{ic} is a measure of the level of prices in the group and p^0_i is a reference price vector. If group expenditure is denoted by x_{ihc}, which is $p_{ic} \cdot q_{ihc}$, the following identity can be drawn:

$$(19) \qquad x_{ihc} = p_{ic} \cdot q_{ihc} = Q_{ihc}\left(p_{ic} \cdot q_{ihc} / k_i \cdot q_{ihc}\right) = Q_{ihc}\pi_{ic}\left(p_i^0 \cdot q_{ihc} / k_i \cdot q_{ihc}\right)$$

where the term in brackets is the measure of quality denoted by ξ_{ihc}. In other words, quality can be defined as the value of a bundle of goods at fixed reference prices relative to its physical volume (Deaton, 1997). Equation (19) also tells us that expenditure is the product of quantity, price and quality. Furthermore, from equation (19), unit value can be defined as:

$$(20) \qquad v_{ihc} = x_{ihc} / Q_{ihc} = \pi_{ic}\left(p_i^0 \cdot q_{ihc} / k_i \cdot q_{ihc}\right) = \pi_{ic}\xi_{ihc}$$

Deaton (1997) then uses the assumption of weak separability to incorporate these quality effects into the utility maximisation problem. When separable groups are formed for each of M commodity groups, the utility function is written as:

$$(21) \qquad u = V\left[v_1(q_1),.....,v_M(q_M)\right]$$

If any subset of commodities appears only in these separable subutility functions, quantities purchased within each group can be expressed as a function of group expenditure and prices within the group alone (Deaton and Muellbauer, 1980a). For a utility-maximising consumer, the cost function for group i is defined as $c_i(u_i, p_i)$ (omitting the subscripts for household and cluster for now) so that:

$$(22) \qquad x_i = c_i\left(u_i, p_i\right)$$

From equations (18) and (20) above, this cost function at a reference price vector can be expressed as:

$$(23) \qquad c_i\left(u_i, p_i^0\right) = \xi_i Q_i$$

According to Deaton (1997), in this formulation group utility u_i is a monotone increasing function of the product of quality and quantity so that overall utility is:

$$(24) \qquad u = V^*\left(\xi_1 Q_1,.....,\xi_M Q_M\right)$$

which is maximised subject to the budget constraint:

$$(25) \qquad \sum_{i=1}^{M} \pi_i \xi_i Q_i = x$$

This is a standard utility maximisation problem and has standard demand functions as solutions written as:

$$(26) \qquad \xi_i Q_i = g_i\left(x, \pi_1,.....,\pi_M\right)$$

As long as the product of quality and quantity is treated as the object of preference, the standard apparatus of demand analysis applies as usual (Deaton, 1997).

Moreover, if the goods in i form a separable group in household preferences, there exists a subgroup demand function of total group expenditure and within-group prices:

(27) $$q_i = f_i(x_i, p_i) = f_i\left(x_i / \pi_i, p_i^0\right)$$

where the last expression comes from using (18) and the fact that demand functions are homogeneous of degree zero. According to the definition of quality (the expression in brackets of equation (19)) and equation (27), quality depends only on the composition of demand within the group (since p^0_i and k_i are constant) and hence on the ratio of group expenditure to group price. In consequence:

(28) $$\frac{\partial \ln \xi_i}{\partial \ln \pi_j} = \frac{\partial \ln \xi_i}{\partial \ln x_i}\left(\frac{\partial \ln x_i}{\partial \ln \pi_j} - \delta_{ij}\right)$$

where δ_{ij} is the Kronecker delta that is unity if $i = j$ and zero otherwise. The term in brackets is the price elasticity of Q_i with respect to π_j, ε_{ij}. The first term is closely related to quality elasticity $\beta^1{}_i$ in equation (6), which can be expressed as follows by the chain rule:

(29) $$\beta_i^1 = \frac{\partial \ln \xi_i}{\partial \ln x} = \frac{\partial \ln \xi_i}{\partial \ln x_i} \cdot \frac{\partial \ln x_i}{\partial \ln x}$$

Given that the last term on the right-hand side of (29) is the total expenditure elasticity of the group, ε_i, (28) can be re-expressed as:

(30) $$\frac{\partial \ln \xi_i}{\partial \ln \pi_j} = \beta_i^1 \varepsilon_{ij} / \varepsilon_i$$

Equations (28) and (30) show that the effects of price on quality operate as income effects. In other words, an increase in the group price depresses group demand through the group price elasticity and it is this fall in demand that generates the change in quality (Deaton, 1988). Further, given (20) and (30), price elasticity of unit value, or the parameter ψ_{ij} in (6) is defined as:

(31) $$\psi_{ij} = \delta_{ij} + \beta_i^1 \varepsilon_{ij} / \varepsilon_i$$

Finally, in the case of the budget share form, the price and total expenditure elasticities of quantities ε_{ij} and ε_i can be defined through the following steps. Firstly, since the logarithm of the share is the sum of the logarithms of quantity and quality less the logarithm of expenditure, the expenditure elasticity of (quantity) demand is:

(32) $$\frac{\delta \ln w_i}{\delta \ln x} = \frac{\beta_i^0}{w_i} = \varepsilon_i + \beta_i^1 - 1$$

As for the price elasticities,

(33) $$\frac{\delta \ln w_i}{\delta \ln \pi_j} = \frac{\theta_{ij}}{w_i} = \varepsilon_{ij} + \psi_{ij}$$

If we rearrange (32) and (33) and express them in matrix notation, we have:

(34) $$E = -\Psi + D(\overline{w})^{-1}\Theta \qquad \text{(price elasticities)}$$

(35) $$e = \iota - \beta^{1} + \beta^{0} D(\overline{w})^{-1} \qquad \text{(expenditure elasticities)}$$

where E is the matrix of price elasticities, e is the vector of total expenditure elasticities, and $D(.)$ is an operator that converts the vector argument into a diagonal matrix.

In sum, according to (30), separability implies that quality shading in response to price is determined by the price, expenditure and quality elasticities of the commodity group. Equation (31) shows that if the expenditure elasticity of quality is zero, the unit value of the commodity i moves proportionately with the price and is independent of prices outside the group. Yet if the quality elasticity is positive, unit values will respond less than proportionately with its own price because the quality is adjusted downward through the income effect. The shortfall will depend on the size of the quality elasticity as well as cross-price elasticities (Deaton, 1997).

As outlined in this section, Deaton (1988, 1990 and 1997) provides a useful solution to the potential problems of unit values, namely measurement error and quality effects, so that unit values can be used as a substitute for market prices. His work has clearly given a breakthrough to the demand analysis of developing economies, for which appropriate price information is often missing. However, to my best knowledge, there has been very limited work so far, which tests Deaton's approach. Brubakk (1997), for instance, compares price elasticities estimated by Deaton's method with those calculated using market price data on the basis of the Norwegian data set. He first estimates a demand system of five commodities using Deaton's procedure (i.e. estimates equations (5) and (6) above) and re-estimates it by applying the observed price data to Deaton's method. Given that an identification problem arises from the presence of both the cluster fixed effect and cluster-specific price in the budget share equation, he defines the cluster fixed effect as a function of some variables that are likely to explain inter-cluster differences, which is expressed as:

(36) $$f_{ic} = \mu_i + \sum_{s=2}^{4} \eta_{si} D_{sc} + \sum_{r=2}^{6} \upsilon_{ri} L_{rc} + \sum_{d=2}^{3} \pi_{di} PO_{dc}$$

where D_{sc}=1 if households in cluster c are observed in season (quarter) s and zero otherwise, L_{rc}=1 if cluster c is in region r and zero otherwise, and PO_{dc}=1 if cluster c falls into population density category d and zero otherwise. He thus adds (36) to the share equation (5) as the cluster fixed effect. Brubakk (1997) concludes that these two approaches produce significantly different results. He points out that some of the assumptions made in Deaton's method may not be valid and thus responsible for the differences in the estimation. For example, he argues that the assumption of no correlation between prices and cluster fixed effects might not be realistic, particularly if the fixed effects include some seasonality. In addition, the weak separability assumption as well as the assumption of the constant relative price structure across clusters may require some investigation for their justification (Brubakk, 1997).

On the other hand, Gibson and Rozelle (2002) carried out an experiment while the Papua New Guinea Household Surveys were being conducted in 1995 and 1996 to test alternative ways of collecting price data. They used three ways to obtain information on prices: (1) from the household survey data as unit values, (2) from a market price survey conducted in each cluster along with the household surveys, and (3) from the "opinions" of household respondents who were shown pictures of various items during the surveys and asked to report the local price for the product in the picture. They consider the price elasticities estimated with the prices from the market price survey as the standard and compare them with those calculated on the basis of the two alternative price proxies.

Their demand system consists of sweet potato, banana, rice and other goods. They first estimate it using a share-log functional form using market prices and unit values respectively. They also estimate the demand system through Deaton's procedure. A comparison shows that even with the application of Deaton's procedure the use of unit values leads to biased estimates of price elasticities, while the picture price series generate the estimates with less bias. They argue that part of the poor performance of Deaton's method may reflect the sample selection problem of several clusters having no unit value available. This seems to be supported by the fact that Deaton's method does better on the sub-sample of those clusters for which unit values are available for all the commodities (Gibson and Rozelle, 2002).

Both works illustrate the failure of Deaton's procedure to produce the "true" price elasticities. Nonetheless, it cannot be said that their analyses are satisfactory. For instance, the price elasticities are poorly determined with both prices and unit values in Brubakk (1997), which implies that they are too noisy to make a legitimate comparison. The way he defines the cluster fixed effects (equation (36)) is also questionable as there are so many other factors that can explain inter-cluster differences. Moreover, Gibson and Rozelle (2002) estimate the demand system of relatively homogeneous commodities. Hence this demand system might not be suitable to test Deaton's method that is designed to deal with the quality issues arising from the heterogeneity of commodity groups. It seems that what we are facing is the trade-offs between estimation bias and extra noise Deaton's procedure can create. Even if Deaton's method eliminates the bias caused by the quality effects and measurement error, it is likely to produce greater variance. If we take into account of the extra costs associated with Deaton's method, his method is perhaps not required for homogeneous goods given that the bias is likely to be relatively small.

Given the problems with the existent studies on the validity of Deaton's method, the main aim of this paper is therefore to improve our current understanding of its usefulness on the basis of Vietnamese data.

4. Model Specification

This section specifies estimation models and describes variables used in the models.

4.1 Model Specification

In order to examine Deaton's procedure, a demand system for Vietnamese households is estimated. To keep the demand function as flexible as possible, the AIDS model described in Section 2 is employed here. Table 1 summarises the models.

Table 1: List of Models

Model	Equations	Prices/Unit values	Estimation methods
A	(37a) (37b)	Prices	OLS
B1	(37a) (37b)	Household-specific unit values	OLS
B2	(37a) (37b)	Commune mean unit values	OLS
C	(38a) (38b) (38c)	Unit values	Stage 1: OLS Stage 2: Errors-in-variable estimation

Model A is our basic model whereby observable market prices are used for the estimation, while Models B1 and B2 are based on a direct substitution of market prices by unit values. Model C follows Deaton's method.

Model A (Basic Model)

(37a)	$w_{ihc} = \alpha_i + \beta_i \ln\left(\frac{x_{hc}}{P_c^*}\right) + \sum_{m=1}\gamma_{im} Z_{mhc} + \sum_{j=1}\theta_{ij} \ln p_{jc} + u_{ihc}$
(37b)	$\ln P_c^* = \sum_{i=1}\overline{w}_{ic} \ln p_{ic}$

The basic model, Model A, is estimated on the basis of market price data. Equation (37a) is the linear AIDS, in which w_{ihc} is the budget share of good i of household h in commune c, x_{hc} is the household's per adult equivalence food consumption expenditure, Z_{mhc} is a vector of household characteristics, p_{jc} is the price of good j observed in commune c, α_i, β_i, γ_{im} and θ_{ij} are parameters to be estimated, and u_{ihc} is an idiosyncratic error term. By assuming weak separability between food and non-food goods, the demand system is estimated for food only, which consists of 10 categories as listed in Table 2 below. Note that the price variables do not have subscript h as they are assumed to be invariant within the same commune. Similarly, the Stone price index (37b) is constructed using the commune average of budget shares, and thus does not vary within communes either. Although the Stone price index only approximates the true value of the price index of the non-linear AIDS (Deaton and Muellbauer, 1980b), given the great complication of Deaton's model, it is desirable to keep the estimation as simple as possible. As a result, all the models employ the Stone price index in the present analysis.

For the purpose of comparison with Deaton's method, Model A is estimated without imposing any restrictions for the demand functions because of the complication associated with Deaton's procedure explained below. Hence equations (37a) and (37b) are estimated for each commodity by OLS. This also applies to Model B whose estimation is based on unit values. There will be two versions of Model B; one with household-specific unit values (B1) and the other with commune mean unit values (B2). Given that household-specific unit values are more likely to be contaminated by errors, which can cause some attenuation bias, Model B2 with commune mean unit values may provide less biased elasticity estimation.

Model C follows Deaton's procedure described in the previous section.[6] The only difference from equations (5) and (6) is that Model C adopts the AIDS demand system for the budget share equations to be consistent with Models A and B.

Model C (Deaton's Method)

(38a)	$w_{ihc} = \alpha_i^0 + \beta_i^0 \ln\left(\frac{x_{hc}}{P_c}\right) + \sum_{m=1}\gamma_{im}^0 Z_{mhc} + \sum_{j=1}\theta_{ij}^* \ln p_{jc} + f_{ic} + u_{ihc}^0$
(38b)	$\ln P_c = \sum_{i=1}\overline{w}_{ic} \ln p_{ic}$
(38c)	$\ln v_{ihc} = \alpha_i^1 + \beta_i^1 \ln x_{hc} + \sum_{m=1}\gamma_{im}^1 Z_{mhc} + \sum_{j=1}\psi_{ij} \ln p_{jc} + u_{ihc}^1$

If we substitute the Stone price index (38b) into the budget share equation (38a), we obtain:

(38d) $$w_{ihc} = \alpha_i^0 + \beta_i^0 \ln x_{hc} + \sum_{m=1}\gamma_{im}^0 Z_{mhc} + \sum_{j=1}\left(\theta_{ij}^* - \beta_i^0 \overline{w}_{jc}\right)\ln p_{jc} + f_{ic} + u_{ihc}^0$$

[6] For the detailed description of the estimation procedure, see Section 3 above.

Given that the price variables, $\ln p_{ic}$, are unobservable, the Stone price index (38b) is also unobservable. Hence if we were going to impose the restrictions, they would have to be imposed on the matrix Θ instead of Θ^*, whose elements are:

$$(39) \qquad \theta_{ij} = \theta^*_{ij} - \beta^0_i \overline{w}_{jc} \qquad \rightarrow \qquad \text{i.e. } \theta^*_{ij} = \theta_{ij} + \beta^0_i \overline{w}_{jc}$$

This implies that the Slutsky symmetry restriction would be imposed by:

$$(40) \qquad \theta^*_{ij} = \theta^*_{ji}, \text{ or equivalently, } \theta_{ij} + \beta^0_i \overline{w}_{jc} = \theta_{ji} + \beta^0_j \overline{w}_{ic}$$

This would require the imposition of a non-linear restriction on the matrix *B* expressed by equation (16) above. Furthermore, in order to estimate the restricted complete demand system for Model C, it is necessary to assume some plausible quality elasticity for the 10th category (other food), which cannot be deduced by the restrictions alone. As a consequence, the unrestricted version of Model C is estimated and its results are compared with those generated by the unrestricted Models A, B1 and B2 to examine the validity of Deaton's method. The comparison based on the estimation of the restricted complete demand system is left for future work.

4.2 Variables

The demand system consists of 10 categories as listed in Table 2. The VLSS contains data on consumption expenditures on 45 food items. For this exercise, they are aggregated into 10 different commodity groups. These expenditures include consumption of purchased goods as well as of home production. The dependent variables are the budget shares of each commodity group in total food expenditure. It should be pointed out that households with zero budget shares are included because we want to find out the total demand response, of both purchasers and non-purchasers, to price changes.

Table 3 exhibits a list of explanatory variables, which are urban and regional dummies, the logarithm of per adult equivalence food consumption expenditure and of household size, the dummy variables for female headed households, ethnic minorities (non-Kinh), farm households, and seasonality based on the date of the interview, and price variables. Note that the prices used in the regression come from the commune price questionnaires carried out along with the VLSS in 1992-93 and 1997-98. These data were collected over a number of goods[7] at the commune level both in the rural and urban sectors[8] where 150 and 194 communes were surveyed in 1992-93 and 1997-98 respectively. Three separate observations were made in each commune, although they did not necessarily involve actual purchases. In some communes, fewer than three observations were made either because of a lack of three distinct markets within the commune or for some other reasons (World Bank, 2000 and 2001).

It should be noted that *communes* are not same as *clusters* and in the case of the VLSS a commune consists of 2 clusters, each of which has 16 households. Since Deaton's method is examined in this analysis, it is necessary to assume no price variation within the same commune as Deaton does for the same cluster. However, because the commune size is twice as large as that of the cluster, it may be questionable to assume prices are invariant within the same commune. On the other hand, it

[7] The data for 1992-93 contain price information for 36 food items, 31 non-food items, 9 pharmaceutical products, 7 agricultural inputs, and 5 types of services from local markets, while 36, 33, 10, 7 and 6 items were covered for each category in 1997-98.

[8] In 1992-93, price data were collected only for rural areas as part of the VLSS. However, a separate set of prices is available for urban areas, which was collected by the GSO and their values are claimed to be comparable with those of the rural prices (World Bank, 2000).

could still be argued that the price data were collected at the relatively geographically disaggregated level. To limit the complexity, this assumption is thus maintained in this exercise. Further, that households belong to the same commune implies that households were located in the same commune *and* interviewed in the same quarter of the year throughout this analysis. Hence if some households were interviewed in a different quarter from the rest of households in the commune, they are treated as if they belonged to a different commune.

Table 2: Dependent Variables (Budget Shares)

Group	Component	1992-93		1997-98	
		Mean	S.D.	Mean	S.D.
Rice	Ordinary rice Sticky rice	0.468	0.188	0.409	0.161
Staple	Corn/Maize Barley, malt, millet and kaoliang Bread, wheat or wheat flour Wheat/egg (dry) noodles Rice noodles Arrow root noodles Cassava Sweet and ordinary potatoes	0.051	0.062	0.039	0.044
Meat	Pork Beef and buffalo Chicken Duck and other poultry Other meat Processed meat	0.117	0.095	0.152	0.089
Fish	Fresh fish, shrimp Dried and processed fish and shrimp Other seafood	0.098	0.087	0.094	0.067
Vegetables	Water morning glory Kohlrabi Cabbage Tomatoes Peanuts, sesame seeds Beans Other vegetables	0.058	0.038	0.058	0.035
Fruits	Oranges Bananas Mangoes Other fruit	0.023	0.028	0.029	0.028
Sugar	Sugar, molasses	0.011	0.014	0.014	0.013
Spice	Salt Fish sauce and dipping sauce MSG	0.050	0.036	0.041	0.021
Dairy	Chicken or duck eggs Milk and other milk products	0.012	0.022	0.017	0.025
Other Food	Lard, cooking oil Cake, candy, candied fruit Alcohol and beer Coffee Tea Beverages Food and drink away from home Other	0.111	0.115	0.148	0.124

Source: Calculations based on the VLSS.

Table 3: Explanatory Variables

	1992-93		1997-98	
	Mean	**S.D.**	**Mean**	**S.D.**
Per adult equivalence total food expenditure ('000 dong)	852	575.0	1642	961.5
Geographical Variables				
Dummy variable for urban sector	0.188	0.391	0.201	0.401
Dummy variable for Northern Uplands	0.159	0.366	0.172	0.378
Dummy variable for Red River Delta	0.241	0.428	0.228	0.419
(North Central)				
Dummy variable for Central Coast	0.116	0.321	0.116	0.321
Dummy variable for Central Highlands	0.027	0.161	0.027	0.161
Dummy variable for South East	0.111	0.314	0.111	0.314
Dummy variable for Mekong River Delta	0.204	0.403	0.204	0.403
Demographic Variables				
log of household size	1.517	0.466	1.461	0.481
Dummy variable for female headed households	0.260	0.439	0.271	0.445
Dummy variable for ethnic minority (non-Kinh)	0.143	0.351	0.143	0.350
Dummy variable for farm households	0.712	0.453	0.634	0.482
Seasonality				
(Interviewed 1st quarter)				
Interviewed 2nd quarter	0.156	0.363	0.305	0.461
Interviewed 3rd quarter	0.304	0.460	0.283	0.451
Interviewed 4th quarter	0.271	0.444	0.261	0.439
Price variables ('000 dong per kilogram)				
Rice (weighted average of ordinary and sticky rice)	1.814	0.252	3.616	0.571
Staple (weighted average of maize, dry noodles, arrow root noodles, cassava and sweet potatoes)	6.798	2.552	11.006	2.295
Meat (weighted average of pork, chicken and duck)	13.027	2.581	20.243	3.311
Fish	9.050	3.764	9.012	2.144
(average of sea fish and fresh water fish in 1992-93)				
(price of scad (fish) in 1997-98)				
Vegetables (weighted average of water morning glory, cabbage, tomato, peanuts and dry beans)	2.315	0.951	4.050	1.327
Fruits (weighted average of orange, banana and mango)	2.781	1.135	6.037	2.142
Sugar	4.675	0.850	7.099	0.395
Spice (weighted average of fish sauce, salt and MSG)	11.446	2.779	17.057	2.455
Dairy (weighted average of eggs and milk)	12.049	1.954	16.774	1.477
Other food	1.000	n.a.	1.000	n.a.
Unit values ('000 dong per kilogram)				
Rice (weighted average of ordinary and sticky rice)	1.887	0.309	3.517	0.690
Staple (weighted average of maize, dry noodles, arrow root noodles, cassava and sweet potatoes)	6.104	3.713	8.781	3.861
Meat (weighted average of pork, chicken and duck)	11.999	2.961	17.829	3.972
Fish (unit value of fresh fish and shrimps)	6.435	2.608	9.974	4.075
Vegetables (weighted average of water morning glory, cabbage, tomato, peanuts and beans)	1.899	1.453	3.490	1.896
Fruits (weighted average of orange, banana and mango)	2.145	1.357	4.637	2.283
Sugar	4.107	0.987	6.824	0.792
Spice (weighted average of fish sauce, salt and MSG)	12.431	4.849	16.079	4.999
Dairy (weighted average of eggs and milk)	11.530	3.214	17.533	5.046
Other food	1.000	n.a.	1.000	n.a.

Source: Calculations based on the VLSS.
Note: Expenditures and prices are expressed in nominal terms.

Based on the assumption that every household within the same commune faces the same prices, the commune price of each commodity is assigned to all the households when calculating price indices for each commodity group.[9] The indices are weighted averages of individual products within the group where commune average expenditures (on purchased goods only) are used for weights. These weighted price indices are therefore invariant within communes. As for unit values, the usual way of deriving unit values is to calculate the ratio of total expenditure for the whole commodity group to the sum of the quantities of individual products within the group. However, in order to be comparable with the price variables, a weighted average of the unit values of the products for which price data are available is calculated for each group. There are three versions of unit values. The first two sets of unit values are constructed using household-specific unit values and expenditures and thus the resultant indices vary within communes (UV1 and UV2). In the case of UV1, missing unit values are left as they are. UV1 is used for Model C and no replacement for missing unit values implies that the unit value equation (38c) is estimated on the sub-sample of those households who made purchases of a particular commodity. The second set, UV2, is the same as UV1 except that missings are replaced by the average of the nearest aggregated level.[10] It is used for Model B1. Finally, the last set, UV3, is constructed using commune mean unit values and expenditures and missing indices are replaced by the average of the nearest aggregated level. Hence UV3 does not vary within communes and it is employed for Model B2.

Price and expenditure elasticities are calculated based on the estimated parameters from each model. For Models A, B1 and B2, the following formula are used:

(41) $$\varepsilon_{ij} = -\delta_{ij} + \left(\theta_{ij} - \beta_i^0 \overline{w}_j\right) / \overline{w}_i$$ (price elasticities)

(42) $$\varepsilon_i = 1 + \beta_i^0 / \overline{w}_i$$ (expenditure elasticities)

As for Model C, Deaton (1990) defines price and expenditure elasticities as equations (34) and (35) respectively in matrix notation.

5. Estimation Results

This section compares price and expenditure elasticities estimated on the basis of Models A, B1, B2 and C and discusses the usefulness of Deaton's procedure. The estimation of Model A is treated as the standard, against which the estimation of the alternative models is compared. They are computed at the sample mean budget shares. The estimated elasticities are summarised in Tables 4 and 5.

Both expenditure and price elasticities computed from Model A are well determined and seem plausible. Meat, fish, fruits and dairy products are found to be luxury goods in both years. Staple food, fish and vegetables are relatively elastic towards price changes.

[9] When commune prices are missing, the mean price for the urban/rural sector of each region interviewed in the same quarter is used for computing the price indices as long as at least one household in the commune purchased that particular product.

[10] For instance, if the commune price index is missing, then the mean price index for the urban/rural sector of the region interviewed in the same quarter is employed.

Table 4: Expenditure Elasticities (Unconstrained)

	1992-93				1997-98			
	Model A	**Model B1**	**Model B2**	**Model C**	**Model A**	**Model B1**	**Model B2**	**Model C**
Rice	0.626 (0.010)	0.830 (0.011)	0.631 (0.010)	0.519 (0.011)	0.501 (0.011)	0.704 (0.010)	0.503 (0.011)	0.358 (0.015)
Staple	0.914 (0.042)	0.656 (0.041)	0.887 (0.043)	0.781 (0.058)	0.993 (0.046)	0.690 (0.041)	1.080 (0.047)	0.943 (0.057)
Meat	1.460 (0.026)	0.753 (0.025)	1.446 (0.026)	1.592 (0.028)	1.430 (0.023)	0.804 (0.020)	1.410 (0.024)	1.521 (0.026)
Fish	1.234 (0.028)	0.938 (0.028)	1.271 (0.028)	1.092 (0.033)	1.065 (0.028)	0.851 (0.025)	1.033 (0.029)	0.881 (0.033)
Vegetables	0.949 (0.023)	0.989 (0.023)	0.916 (0.024)	0.496 (0.037)	0.820 (0.025)	0.855 (0.022)	0.787 (0.025)	0.512 (0.037)
Fruits	1.499 (0.042)	1.219 (0.041)	1.502 (0.043)	1.117 (0.056)	1.367 (0.040)	1.102 (0.035)	1.362 (0.041)	0.994 (0.052)
Sugar	1.231 (0.046)	0.952 (0.044)	1.263 (0.046)	1.283 (0.052)	0.988 (0.037)	0.876 (0.032)	0.989 (0.037)	1.044 (0.041)
Spice	0.568 (0.024)	0.503 (0.023)	0.585 (0.024)	0.476 (0.033)	0.551 (0.019)	0.630 (0.016)	0.582 (0.019)	0.485 (0.028)
Dairy	1.569 (0.062)	0.853 (0.060)	1.524 (0.063)	1.567 (0.070)	1.627 (0.060)	0.870 (0.052)	1.596 (0.061)	1.673 (0.084)

Source: Calculations based on the VLSS.
Note: Figures in the parentheses are standard deviations.

Table 5: Price Elasticities (Unconstrained)

	1992-93				1997-98			
	Model A	**Model B1**	**Model B2**	**Model C**	**Model A**	**Model B1**	**Model B2**	**Model C**
Rice	-0.854 (0.039)	-0.964 (0.033)	-0.912 (0.050)	-0.790 (0.126)	-0.998 (0.030)	-0.896 (0.025)	-0.668 (0.033)	-0.616 (0.092)
Staple	-1.348 (0.045)	-0.969 (0.023)	-1.211 (0.046)	-0.649 (0.089)	-1.910 (0.066)	-1.025 (0.030)	-1.710 (0.062)	-1.014 (0.285)
Meat	-0.735 (0.099)	-0.364 (0.061)	-0.597 (0.110)	-0.604 (0.322)	-0.531 (0.094)	-0.292 (0.049)	-0.620 (0.103)	-0.787 (0.338)
Fish	-1.050 (0.052)	-0.881 (0.037)	-1.362 (0.064)	-1.289 (0.177)	-1.222 (0.053)	-0.866 (0.033)	-1.034 (0.057)	-0.941 (0.132)
Vegetables	-1.047 (0.026)	-0.902 (0.015)	-0.895 (0.027)	-0.539 (0.042)	-1.123 (0.030)	-0.939 (0.017)	-1.072 (0.038)	-0.726 (0.060)
Fruits	-0.913 (0.055)	-0.804 (0.035)	-0.891 (0.060)	-0.678 (0.196)	-0.924 (0.045)	-0.816 (0.030)	-1.094 (0.048)	-0.853 (0.117)
Sugar	-0.571 (0.161)	-0.406 (0.078)	-0.334 (0.137)	-0.650 (0.235)	-0.830 (0.277)	-0.831 (0.107)	-1.118 (0.238)	-0.955 (0.862)
Spice	-0.921 (0.044)	-0.767 (0.019)	-1.091 (0.085)	-0.562 (0.143)	-0.921 (0.054)	-0.830 (0.017)	-0.896 (0.062)	-0.593 (0.100)
Dairy	-0.433 (0.176)	0.060 (0.134)	-0.656 (0.133)	-0.756 (0.481)	-0.510 (0.282)	-0.192 (0.119)	-0.848 (0.132)	-1.353 (0.795)

Source: Calculations based on the VLSS.
Note: Figures in the parentheses are standard deviations. They are obtained from 1,000 replications of the bootstrap using the commune-level data and are defined as half the length of the interval around the bootstrap median that contains 0.638 of the bootstrap replications. This follows Deaton (1997: 319), but unlike Deaton who uses the interval around the bootstrap mean, the bootstrap median is employed here as the distribution of the bootstrap replications appear to be rather skewed for some commodities.

Model A's estimation can be compared with that of Models B1 and B2 whose estimation is based on a direct substitution of unit values for price data. Although both models determine the elasticities relatively precisely, there is a noticeable difference in the estimation of Models B1 and B2. Model B1 generates rather inelastic household behaviour towards expenditure and price changes compared with Model B2 whose estimation is closer to Model A's. This appears to indicate that working with the commune-mean unit values reduces the errors of observations by averaging them and as a result the bias in the estimation seems to become smaller.

On the other hand, from Section 4, we know that unit values are still not the same as prices because of the quality effects and measurement error. Hence Model C, which corrects for these effects, is supposed to generate even closer estimation to the "true" elasticities. However, although the expenditure elasticity estimation of Model C is relatively similar to Model A's estimation except for vegetables and fruits in both years, it is rather surprising to see the poor performance of Model C in the price elasticity estimation. Like Model B1, Model C also estimates inelastic household response to price changes for many commodities compared with Model A. In order to make a more explicit comparison, Table 6 summarises the difference between the estimated elasticities of Model A and those of Models B1, B2 and C respectively. It is calculated as follows assuming that the elasticities estimated by Model A are true:

$$(43) \qquad difference = \frac{\left|E^A - E^S\right|}{\left|E^A\right|} \qquad \text{for } S = B1, B2 \text{ and } C$$

where E is a vector of elasticities (expenditure or own-price). The last rows of the tables show the unweighted average difference of each model.

Table 6: Estimation Difference

Expenditure Elasticities

	1992-93			1997-98		
	B1	**B2**	**C**	**B1**	**B2**	**C**
Rice	0.326	0.008	0.171	0.405	0.004	0.285
Staple	0.282	0.030	0.146	0.305	0.088	0.050
Meat	0.484	0.010	0.090	0.438	0.014	0.064
Fish	0.240	0.030	0.115	0.201	0.030	0.173
Vegetables	0.042	0.035	0.477	0.043	0.040	0.376
Fruits	0.187	0.002	0.255	0.194	0.004	0.273
Sugar	0.227	0.026	0.042	0.113	0.001	0.057
Spice	0.114	0.030	0.162	0.143	0.056	0.120
Dairy	0.456	0.029	0.001	0.465	0.019	0.028
Average	0.262	0.022	0.162	0.256	0.028	0.158

Own-Price Elasticities

	1992-93			1997-98		
	B1	**B2**	**C**	**B1**	**B2**	**C**
Rice	0.129	0.068	0.075	0.102	0.331	0.383
Staple	0.281	0.102	0.519	0.463	0.105	0.469
Meat	0.505	0.188	0.178	0.450	0.168	0.482
Fish	0.161	0.297	0.228	0.291	0.154	0.230
Vegetables	0.138	0.145	0.485	0.164	0.045	0.354
Fruits	0.119	0.024	0.257	0.117	0.184	0.077
Sugar	0.289	0.415	0.138	0.001	0.347	0.151
Spice	0.167	0.185	0.390	0.099	0.027	0.356
Dairy	1.139	0.515	0.746	0.624	0.663	1.653
Average	0.325	0.215	0.335	0.257	0.225	0.462

Source: Calculations based on the VLSS.

As far as expenditure elasticities are concerned, Model B2 produces the estimation with least difference. Although there seems greater divergence for the case of price elasticities for all models, Model B2's estimation is again closest to the estimation of Model A. In contrast, Table 6 confirms the sizeable difference between Model C's estimation and Model A's. For the expenditure elasticity estimation Model C performs better than Model B1, which is based on the household-specific unit values. Yet in the case of price elasticities, its difference from Model A's estimation is slightly larger in 1992-93 and considerably greater in 1997-98 compared with Model B1. In order to find the possible explanations, it is worth looking at each stage of Deaton's procedure.

Table 7: Θ_{ii} and Own-Price Elasticities

1992-93

	Θ_{ii}			Price Elasticities		
	No correction	**Measurement error correction**	**Measurement and quality correction**	**No correction**	**Measurement error correction**	**Measurement and quality correction**
Rice	0.0728	0.0839	0.0697	-0.844	-0.821	-0.790
Staple	-0.0019	-0.0015	-0.0008	-1.037	-1.030	-0.649
Meat	0.0559	0.0439	0.0437	-0.522	-0.624	-0.604
Fish	-0.0443	-0.0603	-0.0478	-1.451	-1.613	-1.289
Vegetables	0.0064	0.0100	0.0054	-0.888	-0.825	-0.539
Fruits	0.0037	0.0042	0.0023	-0.838	-0.819	-0.678
Sugar	0.0043	0.0040	0.0036	-0.615	-0.645	-0.650
Spice	0.0147	0.0170	0.0152	-0.709	-0.662	-0.562
Dairy	0.0029	0.0028	0.0027	-0.751	-0.759	-0.756

1997-98

	Θ_{ii}			Price Elasticities		
	No correction	**Measurement error correction**	**Measurement and quality correction**	**No correction**	**Measurement error correction**	**Measurement and quality correction**
Rice	0.0855	0.0982	0.0829	-0.791	-0.759	-0.616
Staple	-0.0091	-0.0126	-0.0084	-1.231	-1.319	-1.014
Meat	0.0410	0.0297	0.0255	-0.729	-0.804	-0.787
Fish	-0.0208	-0.0272	-0.0168	-1.222	-1.291	-0.941
Vegetables	-0.0019	-0.0035	-0.0026	-1.033	-1.060	-0.726
Fruits	-0.0011	-0.0058	-0.0043	-1.037	-1.201	-0.853
Sugar	0.0003	-0.0014	0.0004	-0.978	-1.104	-0.955
Spice	0.0124	0.0134	0.0131	-0.699	-0.676	-0.593
Dairy	-0.0000	-0.0058	-0.0061	-1.001	-1.349	-1.353

Source: Calculations based on the VLSS.

The second column of Table 7 shows the diagonal elements of the matrix Θ without any correction, i.e. $\Theta = B_{\mathrm{OLS}} = S^{-1}R$ where S is the variance-covariance matrix of unit values and R is the covariance matrix of budget shares and unit values at the commune level. This can be compared with the matrix Θ in the third column obtained after the measurement error correction, which is computed by equation (15). Given the small scale of the covariances between the residuals in the budget share and unit value equations (see σ_{01} in Table 8), the error correction does not change the estimation of Θ considerably apart from for dairy products in 1997-98, which alters proportionately in a more noticeable way. This explains as to way the corresponding price elasticities in the sixth column are relatively similar to those without any correction in the fifth column. Both price elasticities are computed by equation (34) using an identity matrix instead of the matrix Ψ in order to isolate the effects of measurement error by leaving the quality correction aside.

Once the quality effects are taken into account, however, some changes can be observed. This is mainly because the matrix Ψ is far from being an identity matrix. Recall that Ψ is the price elasticity of unit values computed by equation (31), which depends on the price, expenditure and quality elasticities of the commodity groups. According to equation (31), if the quality elasticities (β^1)

were negligible, Ψ would be an identity matrix. Yet in our case the quality elasticities are significant for almost all the commodities in both years as shown in Table 8. Tables 8 and 9 exhibit that the quality effects are particularly pronounced for staple food, vegetables and fruits in 1992-93 and for fish, vegetables and fruits in 1997-98. In addition, although the absolute values of many of the off-diagonals of the matrix Ψ are small, there are a number of cases where the figures are far from being zero. In other words, unit values respond less than proportionally to price changes, which implies that households do not just consume less, but also choose goods of a poorer quality when facing price increases.

Table 8: Quality Elasticities, and First-Stage Variances and Covariances.

	1992-93			1997-98		
	β^1	σ_{01}	σ_{11}	β^1	σ_{01}	σ_{11}
Rice	0.040***	0.00063	0.01412	0.105***	0.00122	0.02179
Staple	0.442***	0.00257	0.60154	0.186***	0.00149	0.26188
Meat	0.056***	0.00171	0.02421	0.087***	0.00174	0.02117
Fish	0.167***	0.00273	0.08834	0.223***	0.00133	0.06825
Vegetables	0.338***	0.00323	0.35736	0.224***	0.00125	0.21212
Fruits	0.368***	0.00188	0.30108	0.343***	0.00168	0.19783
Sugar	0.057***	0.00036	0.04934	0.017***	0.00004	0.01268
Spice	0.116***	0.00253	0.22226	0.073***	0.00101	0.13091
Dairy	0.031**	0.00054	0.04000	0.018	0.00061	0.03811

Source: Calculations based on the VLSS.

Table 9: Ψ Matrix

1992-93

	Rice	Staple	Meat	Fish	Vegetables	Fruits	Sugar	Spice	Dairy
Rice	0.939	-0.002	-0.040	-0.002	0.001	0.005	-0.018	0.018	-0.007
Staple	0.010	0.633	0.174	0.034	-0.044	-0.234	0.099	-0.596	0.027
Meat	0.003	0.002	0.979	0.009	-0.002	0.004	0.013	-0.013	0.012
Fish	-0.084	-0.006	0.066	0.802	0.013	-0.003	0.035	0.035	-0.007
Vegetables	0.066	0.006	0.154	0.030	0.633	-0.131	-0.101	-0.002	0.075
Fruits	-0.116	0.002	-0.174	0.149	0.042	0.777	-0.032	0.021	0.066
Sugar	-0.031	-0.000	0.028	-0.005	0.005	0.015	0.971	-0.004	0.003
Spice	-0.081	-0.003	0.001	0.017	0.001	0.035	0.046	0.863	0.035
Dairy	0.008	0.003	0.025	0.011	-0.000	-0.002	0.012	-0.012	0.985

1997-98

	Rice	Staple	Meat	Fish	Vegetables	Fruits	Sugar	Spice	Dairy
Rice	0.819	-0.007	0.009	-0.019	-0.021	0.001	-0.022	-0.034	0.037
Staple	-0.042	0.801	0.300	0.103	0.047	-0.004	0.331	0.086	-0.047
Meat	-0.007	0.004	0.955	0.006	0.003	0.006	0.008	0.014	-0.026
Fish	-0.018	0.046	0.095	0.762	0.012	-0.020	0.046	-0.052	-0.089
Vegetables	0.002	-0.020	-0.045	0.002	0.682	-0.009	0.199	-0.044	0.077
Fruits	-0.136	0.063	-0.051	0.016	0.041	0.706	-0.051	0.099	0.307
Sugar	-0.008	0.001	0.008	-0.002	0.004	-0.002	0.984	0.000	-0.006
Spice	-0.017	0.016	-0.011	-0.021	0.011	-0.003	-0.036	0.911	-0.030
Dairy	-0.002	0.000	-0.002	0.002	0.003	-0.001	0.004	0.002	0.985

Source: Calculations based on the VLSS.

Given the strong presence of the quality effects, after the quality effect correction in Deaton's procedure, the price elasticities in the last column of Table 7 (which are the same in Table 5 (Model C)) should be closer to those estimated by Model A. Nevertheless, as noted above, they are in fact far from them. One possible explanation is that Model C uses a different sample from the sample employed for the other models. For Models A, B1 and B2, the entire sample (4,302 households) is used for estimation by replacing missing prices/unit values by the mean values of the nearest aggregated level. This is because when we estimate the demand function (37), we would lose a great number of observations if we were to use only those households for which prices/unit values were available for all commodities. In contrast, Model C is run on the sub-sample of those households for whom unit values can be obtained (except for the first-stage budget share equations that are run on the full sample). Given that the second-stage estimation is at the commune-level, we do not face a problem of losing too many observations in the case of Model C. This sample difference might affect elasticity estimation and it is worth checking its effects.

Table 10 shows the sample difference for both periods. The difference becomes smaller in 1997-98, which is most probably due to the commercialisation of the Vietnamese economy (since unit values are computed from the purchase data only). However, the difference is still significant particularly for rice, fruits and dairy products. Hence Model C is re-estimated by replacing missing unit values with the mean values (i.e. UV2 used in Model B1). After correcting the sample difference, the average difference drops from 0.335 and 0.462 to 0.299 and 0.397 in 1992-93 and 1997-98 respectively. In other words, Model B2 in 1992-93 and Models B1 and B2 still outperform Deaton's model.

Table 10: No. of Households in the Samples

	1992-93		1997-98	
	Sample for A, B1 and B2	**Sample for C**	**Sample for A, B1 and B2**	**Sample for C**
Rice	4302	2946	4302	2692
Staple	4302	2851	4302	3450
Meat	4302	4005	4302	4217
Fish	4302	3221	4302	3603
Vegetables	4302	3857	4302	4007
Fruits	4302	2256	4302	3031
Sugar	4302	3228	4302	3748
Spice	4302	4289	4302	4273
Dairy	4302	1480	4302	2310

Source: Calculations based on the VLSS.

Table 11: Ψ Matrix (estimated with prices)

1992-93

	Rice	Staple	Meat	Fish	Vegetables	Fruits	Sugar	Spice	Dairy
Rice	0.411	0.008	0.031	-0.075	-0.000	0.020	-0.133	0.037	0.064
Staple	-0.069	0.755	0.160	-0.097	-0.109	0.010	0.259	0.068	-0.046
Meat	-0.024	-0.009	0.297	0.042	-0.011	0.033	0.052	0.052	0.024
Fish	-0.168	-0.006	0.184	0.173	-0.054	0.098	-0.167	0.003	0.102
Vegetables	0.130	-0.066	0.340	-0.164	0.675	-0.042	0.061	-0.218	-0.256
Fruits	-0.050	0.075	0.029	-0.101	0.219	0.316	0.464	-0.146	-0.081
Sugar	-0.234	-0.028	0.133	-0.094	0.052	-0.023	0.401	0.028	0.034
Spice	-0.098	0.007	-0.120	-0.001	-0.005	0.146	-0.288	0.211	0.056
Dairy	0.059	-0.003	0.095	0.091	-0.048	0.039	0.088	0.029	0.514

1997-98

	Rice	Staple	Meat	Fish	Vegetables	Fruits	Sugar	Spice	Dairy
Rice	0.550	0.054	0.034	-0.029	-0.037	0.002	0.157	0.048	-0.143
Staple	-0.009	0.706	-0.163	0.086	0.058	0.096	0.530	-0.037	0.303
Meat	0.079	0.035	0.484	-0.049	-0.024	0.024	-0.327	0.006	0.104
Fish	0.244	-0.195	0.357	0.184	-0.017	-0.012	-0.237	-0.111	0.324
Vegetables	0.112	0.009	0.080	-0.050	0.640	0.097	0.458	0.149	0.193
Fruits	-0.081	-0.001	0.553	0.164	-0.035	0.520	-1.146	-0.060	0.417
Sugar	-0.064	0.025	0.090	0.008	0.034	-0.006	0.039	-0.021	0.006
Spice	0.074	0.041	0.033	-0.034	0.079	-0.054	-0.233	1.081	0.266
Dairy	0.083	0.099	0.170	0.010	-0.050	0.056	-0.182	-0.029	0.218

Source: Calculations based on the VLSS.

Given the relatively small measurement errors identified above, other possible explanations may be found in the way Deaton's model corrects for the quality effects. Because we are in a better position of having price data, it is possible to estimate the unit value equation (38c) with *observable* prices. Following Deaton's procedure, OLS is applied to the demeaned equation (38c) and the purged unit values are regressed on the price variables. The Ψ matrix estimated in this way is exhibited in Table 11. The results clearly show that the estimation of the unit value response to prices in Deaton's procedure is significantly different from the case where price data are employed, which indicates a stronger presence of the quality effects. To summarise the overall quality effects in terms of the difference from the identity matrix, the following measure is calculated using both matrices in Tables 9 and 11.

$$\text{(44)} \qquad difference = \frac{1}{81}\sqrt{\sum_i \sum_j \left(\psi_{ij} - I_{ij}\right)^2}$$

With the Ψ matrix estimated in Deaton's procedure, the figures are 0.012 and 0.011 in 1992-93 and 1997-98. However, they increase to 0.026 and 0.031 respectively when using the matrix estimated with price data. Deaton's method clearly underestimates the quality effects, which may be one potential factor that prevents Deaton's procedure from generating the "true" price elasticities even after the corrections. Although Model C's price elasticities are re-computed using the Ψ matrix in Table 11 instead, the difference from Model A's estimation does not become smaller, if not larger. This indicates the need for closer examination of Deaton's procedure, particularly the appropriateness of the way the Ψ matrix and its relation to price elasticities are defined.

Other possible explanations include the issues concerning the validity of the assumptions made in Deaton's procedure. For instance, it is assumed that there is no correlation between the commune fixed effects and prices in the budget share equation. As Brubakk (1997) points out, however, the fixed effects could among other things represent some seasonality effects, which are likely to be correlated with prices. If that were indeed the case, the estimated price elasticities would be biased. Furthermore, the derivation of price elasticities in Deaton's model is based on his quality theory that links the price elasticity of quality to the usual price and total expenditure elasticities. This relies on the assumption of the fixed relative price structure across communes. If this assumption is invalid, the treatment of the quality issues described in Section 3 can no longer be applied. Given the time constraint, testing these assumptions is left for future work.

Another important issue to be raised is the implication of the relatively good performance of Model B2. If the substitution of commune mean unit values for prices can provide a relatively good estimation of the "true" price elasticities, Deaton's procedure would not be required as the latter involves greater complication in its computation. The whole purpose of Deaton's method is to solve the problems of the quality effects and measurement error. However, even if we assume Deaton's method eliminates the bias (which does not seem to be the case given the incorrect estimation of the Ψ matrix pointed out above), it is likely to produce greater variance. The matrix B defined as (15), for instance, relies on the estimation of each of its elements. Given the relatively small number of households in each commune, the estimates from equations (8) and (9) may well be far from the true commune means of equations (10) and (11). The problem of the limited sample size also questions the use of probability limit theorem for equation (16). Indeed Table 5 shows the imprecision of Model C's estimation in comparison with the other models. Given that Model B2 seems to reduce some of the estimation bias by working with the commune mean unit values, it might be a preferred model to Deaton's method, at least in this case, as it allows us to avoid the complication associated with Deaton's method as well as the resultant imprecision of the estimation. Furthermore, the fact that the results seem to suggest Deaton's procedure does not even correct the bias appropriately poses a serious question on the validity of his method.

6. Concluding Remarks

This paper examines the validity of Deaton's procedure based on the data from the Vietnamese household surveys. Although his method has widely been applied to many demand analyses, this work is one of the very few studies that empirically examine his model. The results demonstrate that Deaton's model does not generate satisfactory estimation of price elasticities, which is contrary to what has generally been assumed. If it generates the unbiased estimation of elasticities, Deaton's method may be still useful for analysing household behaviour in the absence of price information even thought it produces extra noise. However, the evidence does not even support that Deaton's model produces the unbiased estimation of price elasticities. In fact, as far as this case is concerned, the straightforward substitution of commune mean unit values for prices is found to provide a relatively good approximation to the "true" price elasticities.

These findings therefore pose a serious question on the usefulness of Deaton's method and certainly shed light upon the needs for further examination. In particular, the way the quality effects are estimated and how they are related to price elasticities should be investigated in the future. The justification for some of the assumptions, such as the fixed relative price structure across communes, also has to be explored. Moreover, one of the ultimate objectives of estimating demand functions is usually not just to obtain price and expenditure elasticities, but to evaluate the welfare effect of price changes, which was the initial motivation of this study. As mentioned at the beginning, if the demand system was incorrectly estimated, the subsequent welfare analysis would also be invalid. Hence the comparison between the welfare analysis based on the elasticities estimated with price data and those estimated with unit values through Deaton's procedure is also intended to be made in future work.

References:

Brubakk, L. (1997) "Estimation of Price Elasticities from Norwegian Household Survey Data", *Discussion Papers*, No.202, Oslo: Statistics Norway Research Department.

Deaton, A. (1988) "Quality, Quantity, and Spatial Variation of Price", *American Economic Review*, Vol.78, No.3, pp.418-430.

Deaton, A. (1990) "Price Elasticities from Survey Data: Extensions and Indonesian Results", *Journal of Econometrics*, Vol.44, pp.281-309.

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

Deaton, A. and M. Grosh (2000) "Consumption" in M. Grosh and P. Glewwe (eds) *Designing Household Survey Questionnaires for Developing Countries: Lessons from Ten Years of LSMS Experience*, Chapter 17, Washington, D.C.: World Bank.

Deaton, A. and J. Muellbauer (1980a) *Economic and Consumer Behavior,* New York: Cambridge University Press.

Deaton, A. and J. Muellbauer (1980b) "An Almost Ideal Demand System", *American Economic Review*, Vol.70, No.3, pp.312-26.

Friedman, J. and J. Levinsohn (2002) "The Distributional Impacts of Indonesia's Financial Crisis on Household: A "Rapid Response" Methodology", *World Bank Economic Review*, Vol.16, No.3, pp.397-423.

Gibson, J. and S. Rozelle (2002) "Is a Picture Worth a Thousand Unit Values?: Price Collection Methods, Poverty Lines and Price Elasticities in Papua New Guinea", paper presented at the Northeast Universities Development Consortium Conference (NEUDC), Williams College's Center for Development Economics, October 25-27, 2002.

Gracia, A. and L. M. Albisu (1998) "The Demand for Meat and Fish in Spain: Urban and Rural Areas", Agricultural Economies, Vol.19, No.3, pp.359-366.

Niimi, Y., P. Vasudeva-Dutta and L. A. Winters (2004) "The Storm in a Rice Bowl: Rice Reform and Poverty in Vietnam in the 1990s", *forthcoming Journal of the Asia Pacific Economy*, Vol.9, No.2.

Stavrev, E. and G. Kambourov (1999) “Estimation of Income, Own- and Cross-Price Elasticities: An Application for Bulgaria”, *Transition Economics Series*, No.6, Vienna: Institute for Advanced Studies.

World Bank (2000) *Viet Nam Living Standards Survey (VNLSS) 1992-93: Basic Information*, Washington, D.C.: World Bank.

World Bank (2001) *Viet Nam Living Standards Survey (VNLSS) 1997-98: Basic Information*, Washington, D.C.: World Bank.

Winters, L. A. (2002) “Trade, Trade Policy and Poverty: What are the Links?”, *The World Economy*, Vol.25, pp.1339-1367.

Winters, L. A., N. McCulloch and A. McKay (2002) “Trade Liberalisation and Poverty: The Empirical Evidence”, *Economics Discussion Papers*, No. 88, University of Sussex.