Income Pooling and Women's Expenditure Power: Evidence from Papua New Guinea

John Gibson[1], Dimitri Margaritis[1] and Ann Sullivan[2]

November 2001

[1]Department of Economics and [2]Department of Political Science and Public Policy, University of Waikato, Hamilton, New Zealand

(corresponding author email: dimitri@waikato.ac.nz)

Introduction

A growing number of studies are examining the evidence for the "unitary" household model where households act as if they have a single set of preferences. Amongst the evidence assembled for these tests are studies of whether households pool men's and women's incomes (see Haddad, 1999 for a recent summary) and studies of whether parents reduce their consumption of 'adult' goods by as much for girls as they do for boys, as an indirect indicator of gender bias (Deaton, 1997, Gibson, 1997). A common policy recommendation from these studies is the need to raise women's incomes and to institute other changes that raise their status in household decision-making (Alderman, et al. 1995).

In this study we examine these questions, using data from the 1985-1987 Papua New Guinea Urban Household Survey (UHS). In particular, we are interested in determining whether the data support neoclassical models of common household member preferences or alternatives that allow preference heterogeneity as in models where intra-household allocations represent a Pareto efficient outcome. One advantage of our data is that we have knowledge of not only women's and men's incomes but also their expenditure. Such information provides further insights in understanding the nature of intra-household expenditure patterns. For example, it enables us to assess the robustness of testing (1) Pareto efficiency in the allocation of household income accrued by its female and male adult members and (2) the hypothesis that households pool their income, in situations where female and male expenditures replace income in the allocation decision process. Furthermore, we would like to know whether women are more likely to purchase goods that are generally regarded as enhancing the household's overall welfare or are socially desirable. The policy implications of this result are interesting as they suggest that government interventions should target specifically women's expenditure rather than income. An attempt is also made to use household and community wide socio-economic indicators to determine what are some of the factors that contribute to the differences in intra-household allocations and ways by which policy interventions will become more effective.

Economists are increasingly aware of the effect that assumptions about economies of household size have on poverty and inequality comparisons (Coulter, Cowell and Jenkins, 1992). Thus, the second issue that we propose to address is that of equivalence scales, which are measures of how much children cost relative to an adult. Such measures are important for the analysis of poverty and inequality and the design of policy interventions to improve living standards. For example, whether people in Indian households headed by women are more likely to be poor than are those in male-headed households depends on the adjustment made for size economies (Dreze and Srinivasan, 1997). In the transition economies, the rising relative cost of housing has made size economies more important, shifting the incidence of poverty toward small households and affecting conclusions about whether public interventions should be aimed at children or at the elderly (Lanjouw, Milanovic and Paternostro, 1998).

In this section we study these issues using data from recent household surveys in the developing country of Papua New Guinea. This particular country has the highest degree of inequality in the Asia-Pacific region, with a Gini coefficient of 0.51 for per capita expenditures (World Bank, 2001). The poverty situation is also serious, with over one-third of the population classified as poor. Moreover, the evidence points to a worsening in the poverty situation for this country, with increasing inequalities amongst the poor and a fall in their standard of living relative to the poverty line (Gibson, 2000 and 2001). Estimates of adult equivalence scales using either of the two data sets employed in this study suggest that children in the 0-6 year age group are equivalent to 0.5 adults, other children are equivalent to adults, and economies of household size appear absent. However, there is controversy about the identifying assumptions of these scales (Deaton, 1997).

In spite of their importance, there is no generally accepted method of measuring household size economies. The Engel method, which simply uses the food budget share as a welfare indicator, is popular (Lancaster, Ray and Valuezuela, 1999) but it lacks a theoretical basis (Deaton 1997). The development of other methods, based on the effect of public goods within the household, is currently blocked by an empirical puzzle about the relationship between household size and measured food demand (Deaton and Paxson, 1998).

This paper reports evidence on the empirical fragility of Engel estimates of household size economies. The evidence may also help to resolve the puzzle about food demand raised by Deaton and Paxson (1998), namely that food budget shares fall as household size rises, holding outlay per head constant. A unit increase in (log) household size decreases the food share by up to 10 percentage points in a group of poor countries (Thailand, Pakistan, and African households in South Africa), by 1-2 percentage points in Taiwan and the U.S., and by less in France and in Britain (Deaton and Paxson, 1998).[1] The results use a survey where a sample of households had each adult reporting daily purchases in diaries, while a matched sample had a single respondent give a verbal recall of the household's expenditure for the past fortnight. Both of these methods are widely used by household expenditure surveys in developed and developing countries (ILO, 1994). Regression results from the matched samples, when compared with the results from Monte Carlo experiments, suggest that recalled food expenditures have measurement errors correlated with household size. These correlated errors cause a negative bias in the coefficient on household size in regression models of food budget shares and, consequently, cause Engel estimates of size economies to be overstated. This error is shown to affect the cross-sectional pattern of poverty in the sampled population.

Deaton and Paxson (1998) reject a measurement error explanation for this puzzle because they cannot see why the respondent in a recall survey should be worse informed about others' food consumption than about non-food consumption. But there are grounds for this asymmetry in reporting errors, and for expecting the asymmetry to rise with household size. For example, in the survey used here, a household with two people makes an average of 50 food purchases per fortnight, while a household with 10 people makes 140 food purchases per fortnight. Thus, the respondent from the larger household is the one more likely to forget food purchases when giving her verbal report on expenditures in the previous fortnight. But whether the household has two people or 10 people, it still needs only one gas stove, so the reporting task for non-foods is easier and less proportional to household size. While purchase frequency of some non-food items rises with household size (e.g., gas for the stove), the rate of increase is less than it is for food purchases,[2] and these items are only a small part of total non-food spending. Furthermore, household surveys often impute the value of consumption of certain non-food items (e.g., durables and rent), using observations and measurements

[1] It is perhaps no coincidence that the two countries in Deaton and Paxson's sample with the least puzzling results (Britain and France) collect household expenditures using the diary method, while the other countries gather data by recall.

[2] A household with two people makes an average of 25 non-food purchases per fortnight while a household with 10 people averages 50 non-food purchases per fortnight.

made by the interviewers. The measurement error in these imputations is likely to be orthogonal to household size, in contrast to the errors in food expenditures.

Perhaps because of these differential reporting requirements, a food Engel curve estimated on the sample who received the recall method look most similar to Deaton and Paxson's African households in South Africa – the food budget share falls sharply with increasing household size (at constant outlay per person). But when the sample who received the diary method are used, the results look most similar to their results for Britain – there is a statistically insignificant relationship between the food share and household size.

This paper is organised as follows. Section 1 gives a brief review of the literature on models of household behaviour and outlines the models that we use to determine the patterns of male and female household expenditure and obtain estimates of household size economies. Section 2 discusses the UHS data and reports on the results of the efficiency and pooling tests as well as the socio-economic factors that contribute to the differences in intra-household allocations. Section 3 describes how the data on household expenditures were gathered using the diary method and the recall method on matched samples. It also presents empirical results on size elasticities, which are then used to examine the pattern of poverty across household size groups. Concluding remarks are given in section 4.

1. Models of Household Behaviour

Theoretical models of household behaviour start with the neoclassical model of common preferences (Samuelson 1956, Becker, 1974) and extend to a broad collective framework that includes models of conflict (Folbre, 1986a), cooperative bargaining (Manser and Brown 1980, McElroy and Horney 1981, and Lundberg and Pollak 1993), non-cooperative bargaining (Ulph 1988, Kanbur 1991, Lundberg and Pollak 1993) and Pareto efficiency (Chiappori 1988a, 1988b). In all these models household members may use distinct utility functions. Yet, in the neoclassical models households maximize a 'consensus' utility function (Samuelson 1956) or a 'benevolent' dictator utility function (Becker 1974) subject to a single budget constraint. Income is allocated in a way that the marginal rate of substitution between any two goods is the same as for any other pair of goods (Thomas 1990). In the neoclassical paradigm the well-being of the individual takes center stage, yet individual welfare is intrinsically linked to a multitude of interactions among household members. Thus it comes as no surprise that an increasing number of studies find both little theoretical appeal and empirical support for unitary models of household behaviour based on a common set of preferences. Indeed, Folbre (1986b, p. 251) notes that "it is the juxtaposition of women's lack of economic power with the unequal allocation of household resources that lends the bargaining power approach much of its persuasive appeal."

Collective models of household behaviour provide a far richer structure in comparison to the neoclassical model. Yet, Chiappori (1988a, 1988b, 1992) convincingly argues that such structures often impose a fairly restrictive framework on the household allocation problem that is in turn hard to test. He proposes a cooperative model that assumes household decisions are always Pareto efficient in the sense that no household member can be better off without making another member worse off. The efficient cooperative framework of Chiappori (1988a, 1988b, 1992), Bourguignon and Chiappori (1992), and Browning et al. (1994) is appealing because of its simplicity. Its key feature is that the rules for resource allocation within the household come from the data and are not assumed. This is particularly convenient for assessing the relevance of collective (or non-unitary) models (Alderman et al 1995, Deaton 1997). This approach is also useful for discussing equivalence scales (Deaton 1997).

A simplifies version of the Chiappori model is set up as follows. Following Deaton (1997) we assume that each household consists of two adult members F (female) and M (male), with or without children, whose decisions regarding expenditure are represented by vectors of purchases q^F and q^M, respectively. The process by which the two parties engage to resolve a possible resource allocation conflict is not critical to the model, what is important is that the allocations are efficient in a Pareto sense. Given efficiency, the optimal choice for each member is to maximize an individual utility function subject to an effective budget constraint. For example, the constrained utility maximization problem for member F is to take q^M as given and solve:

$$\text{Max } u^F (q^F , q^M) \text{ s.t. } p^F . q^F = g^F (p,y) \qquad (1)$$

where u is the utility function and g is the sharing rule, the function that determines the level of purchases of member F in terms of goods and services prices (p) and household income (y). The solution to (1) will be a set of demand functions:

$$q^F = R [g^F (p,y), p^F] \qquad (2)$$

There exists a similar set of demand functions for member M given as:

$$q^M = R [g^M (p,y), p^M] \qquad (3)$$

where the sharing rule for M satisfies

$$g^M (p,y) = y - p^F . q^F \qquad (4)$$

An interesting feature of this model is that it nests within the income pooling model where M and F act as if they maximize a single utility function either reached by consensus or as a result of the benevolence of the party that controls the majority of the household’s income. Different kinds of behaviour, e.g bargaining types, can give rise to different sharing rules, but once the sharing rule is set individual demands will be characterized by (2) and (3) (Deaton 1997). Thus this framework is flexible enough to give rise to testable restrictions that can help one identify the household’s allocation mechanism. In particular, it is possible to test whether efficiency holds, and to examine specific forms of efficiency, such as dictatorial behaviour (Deaton 1997). The derivation of testable restrictions requires information on individual incomes (y^F and y^M). Combining (2) and (3) and assuming total household income is the sum of individual incomes plus any joint income, a set of household demand functions can be derived as:

$$q = R [g^F (p, y^F, y^M ,y), p^F] + R [y- g^F (p, y^F, y^M ,y), p^M] \qquad (5)$$

where individual incomes can vary while holding total income constant. Upon differentiating (5) with repect to y^F and y^M , we obtain for each good (i) a ratio of income derivatives given by:

$$\frac{\partial q(i)/\partial y^{M}}{\partial q(i)/\partial y^{F}} = \frac{\partial g^{M}/\partial y^{M}}{\partial g^{M}/\partial y^{F}} \qquad (6)$$

where it is worth noting that the right hand side of (6) is independent of (i). This result enables us to test for Pareto efficiency in the household allocation mechanism by computing the ratio of derivatives in the left hand side of (6) for all goods in the sample, and testing that they are equal to each other. The unitary model of household behaviour is a stronger version of the efficiency test. It requires that the left hand side ratio is equal to unity for each good (i) which is to say that an extra dollar of male income is expected to be spent the same way as an extra dollar of female income. Thus we have two tests nested within the collective household behaviour model, a test of efficiency and the more restrictive test of income pooling. To implement these tests we choose an Engel curve specification of (5) which sets expenditures (c) as a function of individual incomes and household demographic characteristics.

Most of the empirical evidence is not in support of the dictatorial or income pooling model. In particular, Bourguignon et al. (1993) using French data report a rejection of the pooling model but not the less restrictive efficiency model. As Deaton (1997, p. 227) points out "it would be a productive exercise to replicate these tests using data from developing countries, since they provide an obvious first step in an enquiry into the structure of household allocation." This is an important objective of this study. In addition, a novel feature of this research is that the predictions underlying the efficiency and pooling tests relate to both the income and (total) expenditure derivatives of (6).

To gain further understanding on the relationship between households of different composition and living standards, poverty and inequality, we use Engel curve analysis to estimate household size economies. We adopt a two-stage procedure. First, we carry out a Monte Carlo experiment to assess the effects of measurement errors on food budget shares and the Engel estimates of the size elasticity. In the second step, we use the diary and recall survey data to compute Engel estimates of household size economies.

The Engel curve food share model used by Deaton and Paxson (1998) provides the starting point:

$$\frac{p_f q_f}{x} = w_f = \alpha + \beta \ln\left(\frac{x}{n}\right) + \gamma \ln n + \sum_{j=1}^{J-1} \vartheta r_j + \delta \cdot z + u \qquad (7)$$

where $p_f q_f$ is expenditure on food, x is household total expenditure, n is household total size, $r_j = n_j / n$ is the proportion of persons in the household in demographic group j, z is a vector of other household characteristics, u is a disturbance term, and α, β, γ, ϑ, and δ are parameters to be estimated. According to the 'public goods' perspective on household size economies, $\hat{\gamma}$ should be positive, reflecting the income effect on food demand that results when resources are released by extra persons sharing in the consumption of public goods.

This same equation, reparameterised, provides Engel estimates of size economies. For example, Lanjouw and Ravallion (1995) use data from Pakistan to estimate:

$$w_f = \alpha + \beta \ln\left(\frac{x}{n^{1-\sigma}}\right) + \sum_{j=1}^{J-1} \vartheta_j r_j + \delta \cdot z + u\,, \qquad (8)$$

which is identical to equation (7) because $\gamma=\beta\sigma$.[3] According to equation (8), if x^0 is the outlay of a 1-person household, an n-person household of the same composition needs total outlay of $x^0 n^{1-\sigma}$ to have the same food share (and the same welfare level, by assumption). Lanjouw and Ravallion estimate σ to be 0.4, suggesting that ten individuals, each spending, say, \$1 per day in separate single-dweller households, can achieve the same welfare level living as a 10-person single household with total expenditures of just \$4 per day ($10^{0.6}$=3.98). These size economies imply surprisingly large falls in food spending per head for consumers in a poor country. With an average food share of 0.5 in Pakistan, the per-head spending on food in the 10-person group declines by 60 percent, from, say, 50 cents per day to 20 cents per day.

The Monte Carlo experiments may show whether measurement errors could cause such large estimates of size economies, since it is possible to retrieve an estimate of σ from the ratio of γ to β. Even errors in expenditures that are uncorrelated with any explanatory variables affect both dependent and independent variables because ln (x/n) and w_f are constructed from the same information ($x = p_f q_f + p_{nf} q_{nf}$ and $w_f = p_f q_f / x$). Because ln (x/n) and ln (n) are negatively correlated, by construction, errors in ln (x/n) are likely to bias $\hat{\gamma}$, but in an unpredictable direction (Deaton and Paxson, 1998). Bias in $\hat{\gamma}$ is even more likely if the errors are correlated with household size or with the true value of expenditures.[4]

2. Testing for efficiency and pooling in intra-household allocations

2.1 Data

The empirical work in this section of the paper uses data from the Papua New Guinea Urban Household Survey, which was carried out in six provinces in 1985-1987. This sample of 1091 households has several useful features, including (i) the sampling frame was relatively recent (the 1980 Census), (ii) the sampling of households was staggered over the year in each town to remove seasonal effects, and (iii) the data were collected with personal income and expenditure diaries, which were completed by all adults (and included questions on expenditures made by children) for a 14 day period (the customary pay cycle).

It is these income and expenditure diaries that provide the opportunity to study intra-household allocations in a way that is not typically available in other studies. Rather than just obtaining the household total expenditures recorded in the diaries, we also obtained disaggregated information on the total expenditures made by males and the total made by females. It is important to stress that these gender-specific data do not refer to consumption because there is no way to know whether the spending recorded in the diary was to be consumed just by the individual person or shared with others in the household. Nevertheless, the control over expenditures may be different than the control over income and the data available from this survey allow this question to be examined.

---

[3] By rewriting $\beta \ln(x/n^{1-\sigma})$ as $\beta \ln x - (1-\sigma)\beta \ln n$ it is clear that $\beta \ln(x/n) + \beta\sigma \ln n = \beta \ln(x/n^{1-\sigma})$.

[4] Rodgers, Brown, and Duncan (1993) report evidence that errors in survey responses are not independent of either the true values of the variable being measured or of other explanatory variables in the causal model.

Some of the descriptive statistics for the main variables available from the survey are reported in Table 1. It is apparent that, on average, men control over four-fifths of household income (which includes wages, business, informal sector and 'other' activities). It is notable that men's control over expenditures – at least those recorded in the personal diaries – is less marked than is their control over income, with their share of the total being only 60 percent. Indeed, the higher value of female diary expenditures than female cash income suggests that women are spending some of the income earned by men.

It is also apparent that incomes exceed expenditure, in part because of the need for saving but also because of the transfers of income from urban households to rural relatives (Morauta, 1984). The expenditures recorded in diaries account for approximately three-quarters of the value of households' total consumption expenditure, with the other expenditures (mainly infrequent purchases and recurring rental charges and utilities) captured using a household recall form rather than the personal diaries.

Table 1: Descriptive Statistics for Key Variables

	Mean	Standard Deviation
Male Cash Income	254.13	294.28
Female Cash Income	60.61	128.11
Male Diary Expenditures	100.05	163.41
Female Diary Expenditures	66.75	74.81
Value of household consumption of:		
Food	77.59	49.83
Beverages, tobacco, betelnut	28.09	32.72
Clothing	8.20	17.33
Rent and utilities	46.92	71.55
Household goods and furnishings	10.26	33.31
Medical and health expenses	3.68	12.26
Transport and communication	26.56	61.23
Entertainment and education	16.77	43.32
Miscellaneous	11.91	20.97
Total Goods and Services	229.37	189.80

Note: Values are fortnightly household totals, in Kina, and at the time of the survey K1.00=US$1.00.

2.2. Results

Engel curve expenditure equations were estimated for the nine categories of goods and services listed in Table 1, controlling for a number of household demographic characteristics. Those characteristics included household size and composition and province dummies. The province dummies are used to capture the staggered price effects as different provinces were surveyed at different time periods between 1995 and 1997. We report results for the full sample, which includes households with zero female or zero male income, in Table 2, and for the sub-sample, which only includes the non-zero income and diary expenditures, in Table 3. The income and expenditure derivatives reported in Tables 2 and 3 come from sets of Seemingly Unrelated Regression Equations (SURE), which enable the comparisons of coefficients (or their ratios) across equations, as is required in the test of the Pareto efficiency hypothesis. In these tables, male and female incomes (or alternatively, expenditures) are treated as exogenous because Hausman tests did not indicate any inconsistency problems in these SURE results, as compared with Three-Stage Least Squares results which allow endogenous incomes. Further confirmation for this modelling choice comes from the Appendix Tables, which present the results of single equation estimates using both OLS and IV, and which also report Hausman tests and tests of the over-identifying restrictions on the IV model.

Both male and female incomes or total expenditures are significantly, and with the exception of the female income effect on household goods and furnishings, positively associated with expenditures on the various categories of goods and services in our sample. The negative effect of female income on the expenditure for household goods and furnishings is somewhat surprising. It may be explained as reflecting a main income-earner male responsibility towards a predominantly ‘big ticket’ one-off item purchases. Using the full sample, we cannot reject the pooling or efficiency hypotheses when using income as the control variable. However, theses hypotheses are rejected when using total expenditures as the control variable.

Table 2: Tests for Efficient Allocations and Pooling of Household Resources, Using Alternatively Cash Incomes and Diary Expenditures as Income Measures

	Using Male and Female Cash Incomes			Using Male and Female Diary Expenditures		
	$\frac{\partial c}{\partial y^M}$	$\frac{\partial c}{\partial y^F}$	$\frac{\partial c/\partial y^M}{\partial c/\partial y^F}=1$	$\frac{\partial c}{\partial x^M}$	$\frac{\partial c}{\partial x^F}$	$\frac{\partial c/\partial x^M}{\partial c/\partial x^F}=1$
Food	0.056 (0.004)	0.059 (0.010)	$\chi^2_{(1)}=0.10$ [0.75]	0.079 (0.006)	0.363 (0.015)	$\chi^2_{(1)}=1434$ [0.00]
Beverages, tobacco, betelnut	0.027 (0.003)	0.020 (0.007)	$\chi^2_{(1)}=0.43$ [0.51]	0.038 (0.005)	0.139 (0.013)	$\chi^2_{(1)}=226$ [0.00]
Clothing	0.009 (0.002)	0.009 (0.004)	$\chi^2_{(1)}=0.01$ [0.92]	0.051 (0.003)	0.040 (0.007)	$\chi^2_{(1)}=1.37$ [0.24]
Rent and utilities	0.052 (0.007)	0.023 (0.017)	$\chi^2_{(1)}=0.50$ [0.48]	0.105 (0.013)	0.185 (0.031)	$\chi^2_{(1)}=12.19$ [0.00]
Household goods and furnishings	0.030 (0.003)	-0.003 (0.008)	$\chi^2_{(1)}=0.12$ [0.73]	0.052 (0.006)	0.065 (0.015)	$\chi^2_{(1)}=0.82$ [0.37]
Medical and health expenses	0.010 (0.001)	0.007 (0.003)	$\chi^2_{(1)}=0.43$ [0.51]	0.007 (0.002)	0.051 (0.005)	$\chi^2_{(1)}=348$ [0.00]
Transport and communication	0.068 (0.006)	0.081 (0.014)	$\chi^2_{(1)}=0.95$ [0.33]	0.115 (0.010)	0.278 (0.025)	$\chi^2_{(1)}=111$ [0.00]
Entertainment and education	0.039 (0.004)	0.046 (0.010)	$\chi^2_{(1)}=0.49$ [0.48]	0.076 (0.008)	0.154 (0.018)	$\chi^2_{(1)}=39.3$ [0.00]
Miscellaneous	0.015 (0.002)	0.020 (0.005)	$\chi^2_{(1)}=0.96$ [0.33]	0.028 (0.004)	0.092 (0.009)	$\chi^2_{(1)}=184$ [0.00]
Test for efficient allocations			$\chi^2_{(8)}=2.74$ [0.95]			$\chi^2_{(8)}=64.22$ [0.00]

Notes: Results are from Seemingly Unrelated Regressions, with Hausman tests indicating no inconsistency compared with Three-Stage Least Squares where male and female incomes (expenditures) are treated as endogenous ($\chi^2_{(41)}=55.7$ $p<0.07$ for the income model and 48.5 $p<0.23$ for the expenditures model). Standard errors in ( ) and probability of a Type I error in [ ].

Table 3: Tests for Efficient Allocations and Pooling of Household Resources, Using Alternatively Cash Incomes and Diary Expenditures as Income Measures:
Sub-Sample With Non-Zero Incomes and Diary Expenditures (n=419)

	Using Male and Female Cash Incomes			Using Male and Female Diary Expenditures		
	$\frac{\partial c}{\partial y^M}$	$\frac{\partial c}{\partial y^F}$	$\frac{\partial c/\partial y^M}{\partial c/\partial y^F}=1$	$\frac{\partial c}{\partial x^M}$	$\frac{\partial c}{\partial x^F}$	$\frac{\partial c/\partial x^M}{\partial c/\partial x^F}=1$
Food	0.042 (0.008)	0.076 (0.014)	$\chi^2_{(1)}=7.20$ [0.01]	0.105 (0.015)	0.365 (0.023)	$\chi^2_{(1)}=224$ [0.00]
Beverages, tobacco, betelnut	0.023 (0.005)	0.018 (0.010)	$\chi^2_{(1)}=0.12$ [0.73]	0.067 (0.011)	0.148 (0.017)	$\chi^2_{(1)}=29.96$ [0.00]
Clothing	0.007 (0.003)	0.017 (0.006)	$\chi^2_{(1)}=4.57$ [0.03]	0.049 (0.007)	0.032 (0.011)	$\chi^2_{(1)}=0.71$ [0.40]
Rent and utilities	0.014 (0.012)	0.047 (0.022)	$\chi^2_{(1)}=5.01$ [0.03]	0.054 (0.027)	0.131 (0.043)	$\chi^2_{(1)}=4.82$ [0.03]
Household goods and furnishings	0.061 (0.008)	-0.039 (0.015)	$\chi^2_{(1)}=21.11$ [0.00]	0.061 (0.020)	0.045 (0.031)	$\chi^2_{(1)}=0.10$ [0.75]
Medical and health expenses	0.007 (0.002)	0.012 (0.003)	$\chi^2_{(1)}=2.75$ [0.10]	0.009 (0.004)	0.033 (0.007)	$\chi^2_{(1)}=22.31$ [0.00]
Transport and communication	0.039 (0.011)	0.102 (0.020)	$\chi^2_{(1)}=16.39$ [0.00]	0.178 (0.024)	0.280 (0.037)	$\chi^2_{(1)}=7.75$ [0.01]
Entertainment and education	0.025 (0.007)	0.049 (0.012)	$\chi^2_{(1)}=5.37$ [0.02]	0.073 (0.015)	0.148 (0.024)	$\chi^2_{(1)}=12.89$ [0.00]
Miscellaneous	0.012 (0.004)	0.027 (0.007)	$\chi^2_{(1)}=7.50$ [0.00]	0.046 (0.008)	0.100 (0.013)	$\chi^2_{(1)}=24.78$ [0.00]
Test for efficient allocations			$\chi^2_{(8)}=15.02$ [0.06]			$\chi^2_{(8)}=15.00$ [0.06]

Notes: Results are from Seemingly Unrelated Regressions, with Hausman tests indicating no inconsistency compared with Three-Stage Least Squares where male and female incomes (expenditures) are treated as endogenous ($\chi^2_{(41)}=28.94$ p<0.93 for the income model and 17.71 p<0.99 for the expenditures model). Standard errors in ( ) and probability of a Type I error in [ ].

The two categories of goods, that support the pooling hypothesis in the latter case above, are clothing and household goods and furnishings. Using the sub-sample, we do not reject efficiency but do reject pooling (using either income or expenditures). Using expenditure as the control variable, we find again that the two categories of goods that provide support to the pooling hypothesis are clothing and household goods and furnishings. Using income, we find that beverages, tobacco, betelnut is the only category where males and females pool resources. This result is not entirely surprising, as for example, we know from our data that betelnut represents about a 5 percent budget share for both men and women and also soft drinks are part of this expenditure category.

The results reported above suggest that women's control over household expenditures may affect the budget allocations of the household in ways that control over incomes do not. To further examine this point, men's share of total diary expenditures was regressed on a variety of household and community characteristics, in order to find the factors that might affect the control over expenditure by each gender. The results are reported in Table 4. The following groups of explanatory variables were used:

- Men's share of total household income, of total adult school years, and of total household composition. For all three of these variables, it is expected that a higher share accruing to men would raise men's share of expenditures.

- Household expenditures, size and demographic composition (seven age and gender groups), where these variables are included as controls for the standard of living of the household, in case gender expenditure shares differ between rich and poor households.

- An indicator for whether the household head was born overseas, where it is expected that there will be greater autonomy for women in such households. This autonomy may result from the less traditional beliefs of such households, and also from the practical fact that such households are richer, with women who are more educated and more likely to be able to drive, which is likely to increase their ability to shop and to manage household finances.

- The burglary rate in the community, the distance from the dwelling to the nearest police station and the density of police stations. It is expected that women will be less likely to do shopping on behalf of the household in areas where crime is more prevalent and police presence is less apparent. These factors may be especially important in urban Papua New Guinea, which has a severe crime problem.[5]

- An indicator for whether the household lives in a tradition village or squatter settlement within the urban area, and an indicator for whether the household is headed by a migrant (born in another province) and lives in an ethnically mixed community.[6] It is expected that there will be less security for women when shopping if they are relatively recent arrivals into more ethnically diverse communities because they will not have the 'protection' that membership of a clan system brings.

Table 4 reports results derived from two specifications of the men’s share of total diary expenditures along with the descriptive statistics of the variables used in the OLS regressions. The results of Table 4 are very much in line with what we expected to find. Both the men’s share of total income and the male percentage of adults in residence are positively associated with the men’s share of total diary expenditures. The coefficient on the men’s share of adult school years is negative but statistically insignificant. The coefficients on the various age and gender groups are also negative with the notable exception of the coefficient on the share of males in the 0-6 age group which is positive and significant. This finding corroborates

[5] According to Levantis (1997) almost 15 percent of the urban workforce in PNG depend on crime as their main source of income. According to international comparisons, Papua New Guinea cities have the highest rates of victimization for violent crime in the world (exceeding well-known spots such as Rio de Janeiro and Johannesburg), and the rates for property crime are amongst the highest in the world (Zvekic and Alvazzi, 1995).

[6] Ethnic mix is indicated by having a sum of squared population shares less than 0.25, where the population shares relate to the province of birth of the household head.

evidence reported in earlier studies (e.g. see Deaton 1997 and Gibson 1997) of discrimination against girls relative to boys for the youngest age groups. The presence of an overseas born head of household is found to have a negative and significant coefficient in both regressions. Similarly, the proximity to the dwelling and the density of policy stations exert negative and significant effects on male shares of diary expenditures while a migrant head of household in a mixed community has a positive and significant effect.

Table 4: Household and Community Determinants of Male Shares of Diary-Expenditures

	Mean (std. deviation)	Regression coefficients (i)	(ii)
Men's share of total income	0.769	0.155	0.155
	(0.334)	(6.86)	(6.84)**
Men's share of adult school years	0.588	-0.011	-0.012
	(0.317)	(0.43)	(0.47)
Male share of adult residents	0.567	0.434	0.434
	(0.226)	(5.94)	(5.96)**
Share of residents male:: 15-50 yrs	0.343	-0.040	-0.041
	(0.268)	(0.75)	(0.77)
Share of residents male: 7-14 yrs	0.101	-0.206	-0.199
	(0.134)	(2.56)	(2.46)*
Share of residents male: 0-6 yrs	0.102	-0.274	-0.262
	(0.134)	(3.30)**	(3.12)**
Share of residents female: 50+ yrs	0.022	-0.483	-0.480
	(0.085)	(3.62)**	(3.60)**
Share of residents female: 15-50 yrs	0.230	-0.478	-0.477
	(0.158)	(4.79)**	(4.79)**
Share of residents female: 7-14 yrs	0.082	-0.235	-0.229
	(0.122)	(2.71)	(2.63)**
Share of residents female: 0-6 yrs	0.080	-0.153	-0.145
	(0.117)	(1.76)+	(1.67)+
In (per capita expenditures)	8.224	0.013	0.013
	(0.804)	(1.14)	(1.12)
In (household size)	1.545	-0.013	-0.015
	(0.701)	(0.93)	(1.12)
Overseas-born household head	0.070	-0.085	-0.089
	(0.255)	(2.84)**	(2.97)**
Burglary rate in Census Division	0.256	0.047	0.002
	(0.104)	(0.59)	(0.03)
Distance to nearest police station	0.454	0.022	0.019
	(0.877)	(2.70)**	(2.27)*
Police stations per 100,000 people	6.055	-0.022	-0.039
	(0.876)	(2.43)*	(3.46)**
Village or settlement Census Unit	0.434		0.007
	(0.496)		(0.46)
Migrant in mixed community	0.574		0.053
	(0.495)		(2.56)*
Constant		0.466	0.546
		(3.51)**	(3.85)**
		$F_{(16,1072)}$=60.7**	$F_{(18,1070)}$=54.6**
Adjusted R^2		0.468	0.470

Note:
Absolute value of t-statistics in parentheses; *significant at 5%, ** significant at 1%; +significant at 10%; N=1089

3. Household economies of size

3.1 Data: The household survey experiment

This section of the paper uses data from the household survey that was carried out between April and December 1996 in Port Moresby, the capital of Papua New Guinea (PNG).[7] To avoid telescoping errors, there were two interviews, two weeks apart, so that the start of the recall period was signaled by the first interview. Expenditure data were collected on all food (36 categories) and other frequent expenses (20 categories) during the recall period. The expenditure estimates include the imputed value of own-production, net gifts received, and food stock changes (measured by the interviewer). An unbounded recall of the previous 12 months covered 31 categories of infrequent expenses. An inventory of 16 durable assets, plus questions about and measurements of the dwelling, were used to collect the data needed to estimate the value of the flow of services from durable goods and dwellings.

For the questionnaire where each adult recorded daily expenditures, the diaries were organized by means of acquisition (purchases, gifts, own-production) and lists of easily forgotten items were included. Interviewers visited every three to four days to check that respondents had made written daily reports. Interviewers also measured food stocks at the beginning and end of the 14-day period. Data on the infrequent expenditures of these households were collected using the same unbounded 12-month recall as used by the other questionnaire. The data on durable goods and dwellings also were obtained in the same manner as the other questionnaire.

Households within finely defined area units were randomly allocated into two groups: one receiving the diary method and the other receiving the recall method. This is a break from previous studies that apply the two methods sequentially to the same household. This sequential design was not used because it may be subject to conditioning bias – persons who learn to report their expenditures in diaries may, subsequently, be atypically accurate recall respondents, while persons initially given recall questionnaires may find daily recording to be onerous and be atypically bad diary-keepers.

Two-stage sampling was used, with 38 clusters initially selected with probability proportional to estimated size from a frame made up of the 1990 Census plus areas of recent settlement (500 clusters in total). The frame was divided into nine strata, corresponding to major districts of the city, with equal sampling rates used for each strata. At the second stage, circular systematic sampling was used to select six households in each cluster, and a further four households who were "reserves" that were surveyed only if some of the original six selections were absent or refused at the time of the initial interviews. Households were divided into the diary and recall samples at this stage, using a coin toss to decide whether the 1st, 3rd, and 5th households received diaries and the 2nd, 4th and 6th received the recall questionnaire, or *vice versa*. The alternating pattern of diary and recall samples continued into the "reserve" households if they were needed.

---

[7] Data for this research are 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. All views in this paper are those of the authors and should not be attributed to the World Bank.

There was a potential sample of 228 households (38×6) but eight households had missing data due to absence at the time of the second visit for the recall interview or the departure of a household during the diary-keeping period. The reserve households were not used as replacements when this problem arose because a bounded recall would not be possible for these households. To keep the matched nature of the two samples for this analysis, a non-response by a recall household causes the closest diary household within the same cluster to be dropped from the sample (and *vice versa*). This leaves a sample of 212 households. The sample is weighted because selection based on the 1990 Census under-represents clusters found to be larger during the household listing.

Summary statistics for the two samples are reported in Table 5. The average food budget share was 51 percent for the sample where the diary method was used and 45 percent when the recall method was used. This significant difference in food shares is unlikely to be due to differing characteristics of the households across the two samples. All of the demographic variables, except the share of seven to 14 year old males, had no significant differences in means between the two samples. This is to be expected because the demographic and employment variables were collected at the initial interview and are thus independent of the choice of questionnaire. The other variable affected by the choice of questionnaire - per capita expenditure - had an average that was 13 percent lower when the recall method was used, but this difference was not statistically significant.

Table 5: Description of the Data (*N* = 212)

	Diary Sample		Recall Sample		
Variable	Mean	Std. Deviation	Mean	Std. Deviation	t-test for equal means
Food budget share	0.5085	0.1677	0.4502	0.1610	2.83**
ln (per capita expenditure)	7.6998	0.8969	7.5958	0.8389	1.25
ln (household size)	1.6843	0.6612	1.7394	0.6954	0.83
rm06	0.0884	0.1226	0.0776	0.1088	0.60
rf06	0.0918	0.1211	0.0985	0.1237	0.39
rm714	0.1122	0.1365	0.0787	0.1029	2.55**
rf714	0.0721	0.0997	0.0933	0.1165	0.93
rm1550	0.3417	0.2573	0.3735	0.2351	0.75
rf1550	0.2529	0.1543	0.2450	0.1567	0.29
rm51+	0.0352	0.0799	0.0122	0.0364	1.68
rf51+	0.0058	0.0262	0.0211	0.0845	1.16
Adult employment rate	0.6159	0.2550	0.6037	0.2702	0.43

Note: Variables beginning with *r* are demographic ratios, so that e.g., rf714 is the ratio of the number of females aged 7-14 to total household numbers.

Means and standard deviations are calculated using household sampling weights.

The *t*-test uses standard errors corrected for clustering, sampling weights and stratification.

**=significant at p<0.05 (2 sided).

3.2 Monte Carlo Experiments and Empirical Results

The Monte Carlo experiments use a simplified form of equation (7):

$$w_f = \alpha + \beta \ln\left(\frac{x}{n}\right) + \gamma \ln n + u, \qquad (9)$$

where α=1.6, β=-0.14, and γ=-0.007. The value chosen for γ is similar to the estimates of γ made by Deaton and Paxson using a household survey from France, with expenditures reported in diaries. The experiments aim to answer the following question: what sort of measurement errors could cause equation (9) to give estimates of γ like those found when using household surveys from poor countries with expenditure data collected by recall, specifically, $-0.09 \leq \hat{\gamma} \leq -0.05$. This experimental design is not meant to imply that reports of expenditures made in diaries have no errors and that estimates of γ coming from such data are the true values. Rather, this design is used to reveal *differential bias* – the extent to which estimates of γ coming from recall surveys could diverge from the estimates coming from diary surveys due to measurement error in recall surveys that is over and above the measurement error in diary surveys.

The experiments allow uncorrelated errors in expenditures, errors that are correlated with the true value of expenditures, and errors that are correlated with household size. Each experiment moves from a situation of no error to cases of increasingly severe measurement error. Initially, just food expenditures are measured with error, with the experiments carried out as follows: Samples of 1000 observations on (log) total expenditure, x and household size, n were generated from a bivariate normal distribution: $N_2[\mu_x, \mu_n, \sigma_x^2, \sigma_n^2, \rho]$, with μ_x=9.5, μ_n=6.7, σ_x=0.7, σ_n=3.5, and ρ=0.2.[8] The regression errors, u were normal, with mean zero and standard deviation 0.1. Any draws with food budget shares outside the range 0.05-0.95 were dropped.[9] Total expenditure, x was partitioned into food expenditures, $x_f = x \cdot w_f$ and non-food expenditures, $x_{nf} = x - x_f$. A proportionate error was added to true food expenditures, so that the observed variable was $\ln \tilde{x}_f = \ln x_f + v$. In the first experiment the measurement error is independent: $v \sim N(0, \sigma_v^2)$, with three values of σ_v used; 0.1, 0.2, and 0.3. In the second experiment $v = \varphi \ln x_f + \varepsilon$, where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and $E(\varepsilon, x_f)$=0. In the third experiment $v = \lambda \ln n + \varepsilon$, where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and $E(\varepsilon, n)$=0. In the second and third experiments, the values used for φ and λ were -0.3, -0.2, -0.1, 0.1, 0.2, and 0.3. The error-ridden total expenditure and food share variables were reconstructed as $\tilde{x} = \tilde{x}_f + x_{nf}$ and $\tilde{w}_f = \tilde{x}_f / \tilde{x}$, and equation (3) was estimated.

The results of the Monte Carlo experiments can be summarized by the following three observations: First, errors in measuring food expenditures that are negatively correlated with either household size or with the true value of food expenditures are the only type of errors that could cause estimates of γ to be like those found when using surveys from poor countries

[8] These values were drawn afresh at each experiment. Values for household size were rounded to the nearest integer, with the minimum constrained to n=1.

[9] The mean food budget share generated by these assumptions was 0.50 (standard deviation=0.16). The means and standard deviations of the generated x, n, and w_f variables match the data collected in the field.

with expenditure data collected by recall, specifically, $-0.09 \le \hat{\gamma} \le -0.05$ (see Table 6, row 2b. and 3b.). Second, if measurement errors are correlated with the true value of expenditures, the coefficient on ln (*x*/*n*), $\hat{\beta}$ will suffer attenuation bias (i.e., towards zero) but if errors are correlated with household size, there will be no effect on $\hat{\beta}$ (see row 2a. and 3a.). Third, if the true level of size economies (according to the Engel method) is σ=0.05, errors in measuring food expenditures that are negatively correlated with either true values (row 2c.) or with household size (row 3c.) will cause $\hat{\sigma}$ to be biased upwards, with a range of values that includes the estimate of σ=0.4 found by Lanjouw and Ravallion (1995).

Table 6: Monte Carlo Results for Food Share Model

		Experiment 1: Independent measurement errors $v \sim N(0,\sigma_v^2)$			
		No error	$\sigma_v = 0.1$	$\sigma_v = 0.2$	$\sigma_v = 0.3$
1a.	$E(\hat{\beta})$	-0.1379	-0.1344	-0.1241	-0.1082
1b.	$E(\hat{\gamma})$	-0.0073	-0.0047	0.0030	0.0146
1c.	$E(\hat{\sigma})$	0.0518	0.0339	-0.0254	-0.1377
		Experiment 2: Errors correlated with true values $v = \varphi \ln x_f + \varepsilon$, $\varepsilon \sim N(0,0.4)$			
		no error	φ = -0.1	φ = -0.2	φ = -0.3
2a.	$E(\hat{\beta})$	-0.1379	-0.1282	-0.0940	-0.0560
2b.	$E(\hat{\gamma})$	-0.0073	-0.0383	-0.0448	-0.0331
2c.	$E(\hat{\sigma})$	0.0518	0.2986	0.4763	0.5904
			φ = 0.1	φ = 0.2	φ = 0.3
2d.	$E(\hat{\beta})$		-0.0728	-0.0221	0.0001
2e.	$E(\hat{\gamma})$		0.0547	0.0706	0.0543
2f.	$E(\hat{\sigma})$		-0.7594	-3.3308	1636.7
		Experiment 3: Errors correlated with household size $v = \lambda \ln n + \varepsilon$, $\varepsilon \sim N(0,0.4)$			
		no error	λ = -0.1	λ = -0.2	λ = -0.3
3a.	$E(\hat{\beta})$	-0.1379	-0.1263	-0.1262	-0.1242
3b.	$E(\hat{\gamma})$	-0.0073	-0.0289	-0.0582	-0.0844
3c.	$E(\hat{\sigma})$	0.0518	0.2282	0.4603	0.6792
			λ=0.1	λ=0.2	λ=0.3
3d.	$E(\hat{\beta})$		-0.1200	-0.1142	-0.1072
3e.	$E(\hat{\gamma})$		0.0357	0.0686	0.1003
3f.	$E(\hat{\sigma})$		-0.2999	-0.6039	-0.9393

Note:
Results based on 10,000 replications of the model: $w_f = \alpha + \beta \ln(x/n) + \gamma \ln n + u$.
The true values are α=1.6, β=-0.14, γ=-0.007, and $\sigma = \gamma/\beta$, so the implied true value is σ=0.05. Each series is 1000 observations.

When measurement errors in non-food expenditures that are uncorrelated with the errors in food expenditures are introduced, the pattern of results is largely unchanged. If the errors in non-food expenditures are independent, i.e., $\ln \tilde{x}_{nf} = \ln x_{nf} + g$ where $g \sim N(0,0.4)$, the effect of food expenditure errors is amplified slightly. If the errors in non-food expenditures vary negatively with household size, $g = -0.2 \ln n + \zeta$ where $\zeta \sim N(0,0.4)$ and there is only a weak correlation between food expenditure errors and household size (λ=-0.1), the estimate of γ tends to be positive. But as food expenditure errors become more strongly correlated with household size (λ≤-0.2) the expected value of $\hat{\gamma}$ moves into the range $-0.09 \leq \hat{\gamma} \leq -0.05$. Errors that are correlated with household size have no effect on $\hat{\beta}$, which is the same as was found when there were no errors in non-food expenditures.

The model is based on the specification used by Deaton and Paxson (Equation 7), excluding variables not available in an urban areas. In addition to per capita expenditure and household size, there are eight demographic ratios, r_j; and the proportion of adults who are employed. The employment rate may affect the food share because of higher caloric requirements for workers or because of the higher cost of meals eaten out of the home. The model also includes dummy variables for the calendar quarter in which the household was surveyed. The model is separately estimated on the sample of households whose expenditures were reported in diaries and on the sample whose expenditures were recalled in an interview. Slope and intercept dummy variables from a model estimated on the pooled sample are used to test which, if any, of the coefficients differ between the two samples. The model is estimated by both Ordinary Least Squares (OLS) and Instrumental Variables (IV), and the estimation methods account for the clustered, stratified, and weighted nature of the sample.

OLS Results

The method used to collect expenditure data affects the relationship between household size and food share but does not affect any of the other coefficient estimates of the food Engel curve (Table 7). When the Engel curve is estimated on the sample of households whose expenditures were recalled in an interview, household size appears to exert a negative and statistically significant effect on the food budget share, holding per capita outlay constant. The effect is even larger than that found for the poorest households in Deaton and Paxson's sample; a unit increase in the logarithm of household size decreases the budget share of food by 12 percentage points (10 percentage points if using the unweighted data). But when the Engel curve is estimated on the sample of households whose expenditures were reported in diaries, household size has no statistically significant effect on the food budget share, and the point estimate is almost zero – a result that is similar to what Deaton and Paxson find for Britain. Hence, the variation in data collection methods across the group of countries studied by Deaton and Paxson may account for some of their results.

The Monte Carlo experiments suggest that measurement errors in expenditures that are negatively correlated with either household size or with the true value of food expenditures could cause negative bias in the coefficient on household size. According to the Monte Carlo results, we can distinguish which of these two types of measurement error are present by what happens to the coefficient on ln (*x*/*n*). If errors are correlated with true values, $\hat{\beta}$ will be biased toward zero but if errors are correlated with household size, there will be no effect on $\hat{\beta}$. The coefficient on ln (*x*/*n*) when the diary sample is used is identical to the coefficient when the recall sample is used. This suggests that the measurement errors in the recall sample are negatively correlated with household size.

Table 7: OLS Estimates of the Food Engel Curve

	Diary Sample		Recall Sample		*t*-test for equal
Explanatory variable	coefficient	$\|t\|$	coefficient	$\|t\|$	coefficients
ln (per capita expenditure)	-0.1329	9.26	-0.1328	4.66	0.01
ln (household size)	-0.0026	0.17	-0.1262	2.73	2.64
rm06	0.3998	1.08	0.1991	1.16	0.50
rf06	0.2304	0.63	0.1657	1.14	0.17
rm714	0.0766	0.22	0.1301	0.64	0.14
rf714	0.4232	1.00	0.2813	1.82	0.36
rm1550	0.3228	0.96	0.3034	3.68	0.06
rf1550	0.3532	0.99	0.1952	1.25	0.45
rm51+	0.3406	0.88	0.8271	2.15	0.98
Adult employment rate	0.1207	2.41	0.0123	0.19	1.24
Constant	1.1788	3.32	1.4919	5.34	0.67
R^2	0.5457		0.4008		
$F_{(12,18)}$	30.46		7.86		

Note: Variables beginning with *r* are demographic ratios, so that e.g., rf714 is the ratio of the number of females aged 7-14 to total household numbers. The omitted group is elderly females. Models also contain two quarterly dummies.

Reported absolute *t*-values are corrected for clustering, sampling weights and sample stratification.

$F_{(12,18)}$ is an adjusted Wald (W) test for zero slopes: $\frac{d-k+1}{kd}W \sim F(k, d-k+1)$, where d is the number of clusters minus the number of strata (29), and k is the number of slope variables.

Instrumental Variables Results

The IV specification of Deaton and Paxson, where per capita cash income instrumented for ln (x/n), was not exactly replicated because income data were not gathered by the PNG survey.[10] Measures of wealth were available (dwelling quality and value of durables) but were closely related to imputed non-food expenditures and hence directly influenced the food budget share. Therefore, the instruments chosen were the average number of school years of each adult in the household and the age of the household head, which are both good predictors of per capita expenditures – raising the R^2 in the first stage regression of ln (x/n) on the exogenous variables from 0.41 to 0.71. Over-identification tests suggested that these two variables did not play a direct role in the explanation of food budget shares, so they should be valid instruments, even though they have sometimes appeared in food demand studies elsewhere. One concern with using schooling as an instrument is that illiteracy (a close correlate) could cause measurement error in expenditure data that are collected by the diary

[10] Deaton and Paxson use IV to deal with random measurement errors in ln (x/n) that might bias the γ coefficient because of the correlation between ln (x/n) and ln n.

method. However, this should not be a problem in the current study because adult literacy rates are high by developing country standards (approximately 90 percent).[11]

The IV estimates of the Engel curve, reported in Table 8, show the same pattern as the OLS estimates. Just as before, the only variable with a statistically significant difference in coefficients between the diary and recall samples is household size. The use of an instrument for ln (x/n) causes only small changes in the value of the coefficients on household size, compared with the OLS results.[12] There also are small changes in the coefficients on per capita expenditures and a widening of the standard errors surrounding all coefficients. However, the Durbin-Wu-Hausman tests suggest that there is no need to use the IV estimates.

Table 8: IV Estimates of the Food Engel Curve

Explanatory variable	Diary Sample coefficient	Diary Sample $\|t\|$	Recall Sample coefficient	Recall Sample $\|t\|$	*t*-test for equal coefficients
ln (per capita expenditure)	-0.1536	6.70	-0.1039	2.73	1.12
ln (household size)	-0.0147	0.81	-0.1049	2.42	1.77
rm06	0.3513	0.97	0.2335	1.23	0.29
rf06	0.1685	0.48	0.2058	1.28	0.10
rm714	0.0411	0.12	0.1853	0.83	0.39
rf714	0.3559	0.81	0.2928	1.84	0.16
rm1550	0.2895	0.88	0.3354	3.61	0.15
rf1550	0.3067	0.89	0.2048	1.27	0.30
rm51+	0.2876	0.77	0.8862	2.29	1.28
Adult employment rate	0.1265	2.43	0.0154	0.22	1.23
Constant	1.3958	3.47	1.2046	3.34	0.34
R^2	0.5385		0.3894		
$F_{(12,18)}$	22.42		7.07		
Durbin-Wu-Hausman test	t=1.06		t=1.35		
Over-identification test	$\chi^2_{(2)}$=0.25		$\chi^2_{(2)}$=0.02		

Notes: See Table 3.
Instruments for ln (per capita expenditure) are the average number of school years of each adult in the household and the age of the household head.

[11] Illiterates were typically the elderly, so sometimes other family members recorded in the diaries on their behalf or interviewers visited more frequently than usual (roughly every second day) and did the recording.

[12] This is also true when the unweighted data are used. For the recall sample, the OLS coefficient on household size is –0.1000 and the IV coefficient is –0.0928. For the diary sample, the OLS coefficient is -0.0061 and the IV coefficient is –0.0237.

Engel Estimates of Size Economies

The results reported in Tables 7 and 8 suggest that an appropriate model for the pooled sample needs only one slope dummy variable, for household size, and that it is safe to use OLS to estimate such a model. The results of this model are:

$$w_f = 1.480 - 0.140 \ln (x/n) - 0.083 \ln n + 0.049 [\ln n * \text{Diary Dummy}]$$
$$(9.85) \qquad (3.41) \qquad (4.36)$$
$$+ \text{ demographic ratios} + \text{adult employment rate} + \text{quarterly dummies}$$
$$R^2=0.44; \qquad F_{(13,17)} = 15.84$$

t-statistics in ( ) corrected for clustering, sampling weights and stratification

These results suggest that when an Engel curve is estimated with expenditure data collected by recall, a unit increase in the logarithm of household size will cause the observed food share to fall by five percentage points more than it would if the data were collected by respondents reporting expenditures in diaries. The elasticity of per capita food expenditure with respect to household size (which equals γ/w_f given per capita expenditure) is estimated as –0.184 when using expenditure data collected by recall but only –0.067 when using expenditure data reported in diaries (at the average food shares in Table 5).

According to the Engel method, the size economies parameter σ can be estimated from the ratio of the coefficients on ln n and ln (x/n). The coefficients from the pooled model reported above give estimates of σ for the diary and recall samples of:

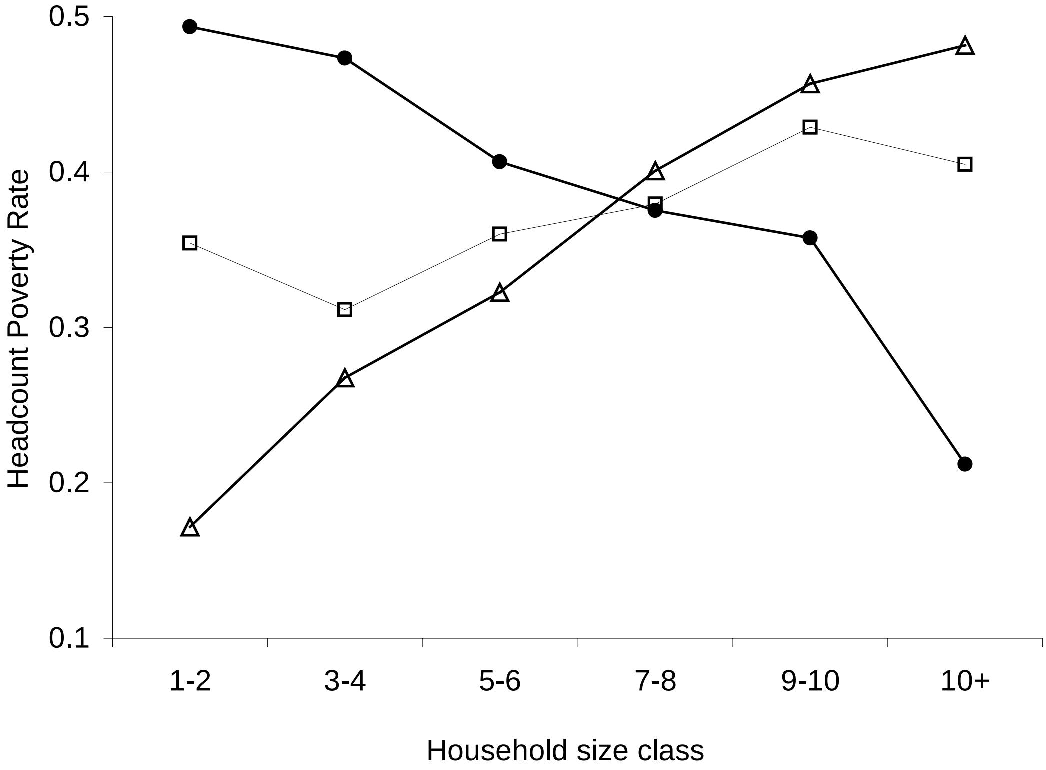
Diary Sample	Recall Sample
$\sigma = 0.24$	$\sigma = 0.59$
s.e.(σ) = 0.15	s.e.(σ) = 0.18
$H_0: \sigma = 0$ $\chi^2_{(1)} = 2.54$ (p<0.12)	$H_0: \sigma = 0$ $\chi^2_{(1)} = 10.77$ (p<0.01)

Engel estimates of size economies appear to be sensitive to the method used to collect household expenditure data. When expenditure data are collected from recall interviews, size economies appear large. But when expenditure data are collected by having a similar sample of households record expenditures in diaries, estimated size economies are much smaller and a per capita normalization of total expenditures (i.e., σ=0) would not be rejected. Thus it is possible that the estimate of σ=0.4 made by Lanjouw and Ravallion (1995) for households in Pakistan is biased upwards by the use of recall data, which may have measurement errors in expenditures that are correlated with household size.

Figure 1 shows the pattern of headcount poverty by household size class for the various estimates of the economies of household size parameter, σ.[13] These poverty rates are based on normalised estimates of effective household size, $n^{1-\sigma}$ so that the poverty rate is always the same (37 percent) for the average sized household. It is clear that when no allowance is made for economies of size, the headcount poverty rate increases rapidly with household size, approaching almost 50 percent for households with more than ten members. In contrast, when the estimate of size economies calculated from the recall sample (σ=0.59) is used, the largest

[13] The size classes are used to give smoother estimates because of the small sample size. For the same reason, the poverty rates are calculated from all areas of Papua New Guinea rather than just from the capital city.

households have a poverty rate of around 20 percent but the smallest have a poverty rate of 50 percent. The robustness of this pattern must be questioned, however, because the estimate of σ from the diary sample implies a slight rise in poverty with increasing household size.



4. Conclusion

In this study we provide further quantitative evidence on the role of gender in household resource allocation decisions. Information on intra-household expenditure patterns is of considerable interest to policymakers seeking effective means by which intervention programs will lead to socially desirable outcomes.

The results provide statistical evidence to support the hypotheses of (1) Pareto efficiency in the allocation of household income accrued by its female and male adult members and (2) that the household income is pooled so that ownership of income does not matter in the allocation mechanism. The latter finding is in contrast to evidence available from other developing countries. Further statistical results indicate that the pooling model is rejected when female and male expenditures replace income in the allocation decision process. In particular, women are found to be more likely to purchase goods that are generally regarded as enhancing the household's overall welfare or are socially desirable. The policy implications of this result are interesting as they suggest that government interventions should target specifically women's expenditure rather than income. An attempt is also made to use household and community wide socio-economic indicators to determine what are some of the factors that contribute to the differences in intra-household allocations and ways by which policy interventions will become more effective.

Economists who want to measure economies of household scale face unpalatable choices. They can use the atheoretical Engel method, which works but makes no sense (Deaton, 1997) or they can try a method – based on public goods within the household – that makes sense but does not work.[14] This public goods method does not work because its prediction of higher food expenditure per head in larger households (of equal per capita outlay) is rejected by the data (Deaton and Paxson, 1998). It is no coincidence that the prediction fails most when household budget surveys have a single respondent give a verbal recall of the household's expenditures from the previous week, fortnight, or month. In contrast, the two countries in Deaton and Paxson's sample with results closest to predictions collect household expenditures by having each adult report their daily purchases in diaries.

To test the effect that household survey methods have on the relationship between household size and food demand we used the diary and recall methods to collect data on the expenditures of two random samples of households in the same location and then estimated a food Engel curve on both samples. The elasticity of per capita food expenditure with respect to household size, given per capita total expenditure, was estimated to be -0.18 when using data collected by recall, but just -0.07 when using data reported in diaries. Similarly, if the expenditure data collected by recall are used, Engel estimates of scale economies appear much larger than if the expenditure data reported in diaries are used.

The most plausible interpretation of these results is that food expenditure data collected with the recall method have measurement errors that are correlated with household size. As household size increases, it becomes increasingly harder for a survey respondent to accurately recall expenditures on food because of the rise in the number of transactions. It is easier to recall expenditures on non-food items because these may be purchased only sporadically. These measurement errors in expenditure data cause a negative bias in the coefficient on household size in regression models of food demand. Thus, one reason why the Engel method of measuring scale economies works is because it mistakes these errors in the food expenditures of large households for genuine scale economies.

[14] Other methods based on estimation of a system of demand equations (for example, see Lancaster *et. al.*, 1999) may be too involved for many applied economists who just want to estimate an equivalence scale as an input into some other modeling task.

References

Alderman, H., P. Chiappori, L. Haddad, J. Hoddinott, and R. Kanbur. (1995) "Unitary versus collective models of the household: is it time to shift the burden of proof?" *World Bank Research Observer* 10(1): 1-19.

Becker, G.S. (1974) "A theory of social interactions" *Journal of Political Economy*, 82(6): 1063-1093.

Bourguignon, F., Browning, M., Chiappori, P.-A., and V. Lechene (1993) "Intrahousehold allocation of consumption: a model and some preliminary evidence" *Annals d'Economie et de Statistique*, 29: 137-156.

Bourguignon, F., and P.-A. Chiappori (1992) "Collective models of household behaviour: an introduction" *European Economic Review*, 36: 355-364.

Browning, M., Bourguignon, F., Chiappori, P.-A., and V. Lechene (1994) "Income and outcomes: a structural model of intrahousehold allocation" *Journal of Political Economy*, 102(6): 1067-1096.

Chiappori, P.-A. (1988a) "Rational household labour supply" *Econometrica*, 56(1): 63-89.

Chiappori, P.-A. (1988b) "Nash bargained household decisions" *International Economic Review*, 29(4): 791-796.

Chiappori, P.-A. (1992) "Collective labour supply and welfare" *Journal of Political Economy*, 100(3): 437-467.

Coulter, F.A.E., Cowell, F.A. and S.P. Jenkins (1992) "Equivalence scale relativities and the extent of inequality and poverty" *Economic Journal*, 102: 1067-1082.

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

Deaton, A., and C. Paxson (1998) "Economies of scale, household size, and the demand for food" *Journal of Political Economy*, 106(5): 897-930.

Dreze, J., and P.V. Srinivasan. (1997) "Widowhood and poverty in rural India: some inferences from household survey data" *Journal of Development Economics* 54: 217-234.

Folbre, N. (1986a) "Cleaning house" *Journal of Development Economics*, 22: 5-40.

Folbre, N. (1986b) "Hearts and spades: paradigms of household economics" *World Development*, 14: 245-255.

Gibson, J. (1997) "Testing for boy-girl discrimination with household expenditure data: results for Papua New Guinea" *Applied Economics Letters* 4(9): 643-646.

Gibson, J. (2000) “The impact of growth and distribution on poverty in Papua New Guinea” *Applied Economics Letters* 7(8): 541-544.

Gibson, J. (2001) “Measuring chronic poverty without a panel” *Journal of Development Economics* 65: 243-266.

Haddad, L. (1999) “The income earned by women: impacts on welfare outcomes” *Agricultural Economics* 20(1): 135-141.

ILO (1994) "Household income and expenditure surveys" *Year Book of Labour Statistics, Sources and Methods,* Volume 6, International Labour Office, Geneva.

Kanbur, R. (1991) “Linear expenditure systems, children as public goods and intra-household inequality” mimeo, University of Warwick and the World Bank.

Lancaster, G., Ray, R., and M. Valenzuela (1999) “A cross-country study of equivalence scales and expenditure inequality on unit record household budget data” *Review of Income and Wealth* 45(4): 455-482.

Lanjouw, P., Milanovic, B., and S. Paternostro (1998) “Poverty in the transition countries: a case of children pitted against the elderly?” mimeo, Development Economics Research Group, The World Bank.

Lanjouw, P. and M. Ravallion (1995) “Poverty and household size” *The Economic Journal,* 105(Nov): 1415-1434.

Levantis, T., (1997) “Urban unemployment in Papua New Guinea - its criminal. Pacific” *Economic Bulletin*, 12(2): 73-84.

Lundberg, S. and R.A. Pollak (1983) “Separate spheres bargaining and the marriage market” *Journal of Political Economy*, 10: 987-1010.

Manser, M. and M. Brown (1979) “Bargaining analysis of household decisions” in *Women in the Labour Force*, C.B. Lloyd, E.S. Andrews, and C.L. Gilroy (eds.), New York: Columbia University Press.

McElroy, M.B. and M.J. Horney (1981) “Nash bargained household decisions: toward a generalization of the theory of demand” *International Economic Review*, 22(2): 333-349.

Morauta, L. (1984) "Income, unemployment and welfare in low-income urban areas" mimeo, Institute of Applied Social and Economic Research, Boroko, Papua New Guinea.

Rodgers, W., Brown, C. and G. Duncan (1993) “Errors in survey reports of earnings, hours worked, and hourly wages” *Journal of the American Statistical Association*, 88(424): 1208-1218.

Samuelson, P.A. (1956) “Social indifference curves” *Quarterly Journal of Economics*, 70: 1-21.

Thomas, D. (1990) "Intra-household resource allocation: an inferential approach" *Journal of Human Resources*, 25(4): 635-664.

Ulph, D. (1988) "General noncooperative Nash model of household consumption behaviour" Discussion Paper 88-205, University of Bristol.

Zvekic, U. and Alvazzi del Frate, A., (1995) *Criminal Victimization in the Developing World*, United Nations Interregional Crime and Justice Research Institute: Rome.

World Bank (2001) *World Development Indicators*, New York: Oxford University Press.

Appendix

Single Equation Results

Table A1: Equation-by-Equation Results, Full Sample, Income

	Instrumental Variable Results[a]		Ordinary Least Squares Results		Hausman test[b]	Over-identication[c]
	$\partial c/\partial y^M$	$\partial c/\partial y^F$	$\partial c/\partial y^M$	$\partial c/\partial y^F$	(p-value)	(p-value)
Food	0.072	0.129	0.056	0.059	0.20	0.83
	(0.013)	(0.033)	(0.008)	(0.014)		
Beverages, tobacco,	0.046	-0.004	0.027	0.020	0.98	0.06
betelnut	(0.012)	(0.024)	(0.005)	(0.009)		
Clothing	0.018	0.014	0.009	0.009	0.92	0.98
	(0.005)	(0.012)	(0.003)	(0.005)		
Rent and utilities	0.052	0.123	0.052	0.023	0.96	0.52
	(0.023)	(0.053)	(0.018)	(0.020)		
Household goods and	0.016	0.039	0.030	-0.003	0.99	0.91
furnishings	(0.011)	(0.021)	(0.018)	(0.012)		
Medical and health	0.010	0.025	0.010	0.007	0.92	0.75
expenses	(0.003)	(0.007)	(0.002)	(0.004)		
Transport and	0.083	0.098	0.068	0.081	0.99	0.83
communication	(0.017)	(0.032)	(0.013)	(0.037)		
Entertainment and	0.056	0.089	0.039	0.046	0.79	0.67
education	(0.013)	(0.022)	(0.007)	(0.014)		
Miscellaneous	0.032	0.068	0.015	0.020	0.00	0.34
	(0.007)	(0.014)	(0.003)	(0.007)		

Note: Reported derivatives are coefficients from regressions that also include household size and demographic composition variables and province dummy variables. Heteroscedastically robust standard errors in ( ).

[a] The instruments include the number of adult males and females in the household, the number of school years of those males and females, and dummy variables indicating when there are no adult males or females. The test for excluding these instruments from the first stage regression is $F_{(6, 1089)}=28$, which is significant at $p<0.01$.

[b] The Hausman test is based on the vector of contrasts between the IV and OLS results, and if $p<0.05$ a lack of consistency in the OLS results, due to the possible endogeneity of incomes, would be indicated.

[c] The test for over-identification is based on the relationship between the instruments and the disturbances of the IV regression. This test indicates one aspect of the validity of the instruments, which is that they are orthogonal to the regression disturbances.

Table A2: Equation-by-Equation Results, Full Sample, Diary Expenditures

	Instrumental Variable Results[a]		Ordinary Least Squares Results		Hausman test[b] (*p*-value)	Over-identication[c] (*p*-value)
	$\partial c/\partial x^M$	$\partial c/\partial x^F$	$\partial c/\partial x^M$	$\partial c/\partial x^F$		
Food	0.241	0.329	0.079	0.363	0.82	0.90
	(0.050)	(0.052)	(0.038)	(0.023)		
Beverages, tobacco,	0.204	0.008	0.038	0.139	0.79	0.37
betelnut	(0.058)	(0.055)	(0.022)	(0.017)		
Clothing	0.072	0.035	0.051	0.040	0.99	0.86
	(0.023)	(0.023)	(0.012)	(0.011)		
Rent and utilities	0.187	0.286	0.105	0.185	0.99	0.70
	(0.108)	(0.106)	(0.023)	(0.069)		
Household goods and	0.043	0.105	0.052	0.065	0.99	0.98
furnishings	(0.049)	(0.035)	(0.009)	(0.013)		
Medical and health	0.030	0.064	0.007	0.051	0.97	0.93
expenses	(0.013)	(0.014)	(0.003)	(0.014)		
Transport and	0.314	0.253	0.115	0.278	0.95	0.82
communication	(0.090)	(0.083)	(0.048)	(0.055)		
Entertainment and	0.190	0.231	0.076	0.154	0.63	0.77
education	(0.058)	(0.052)	(0.009)	(0.029)		
Miscellaneous	0.107	0.163	0.028	0.092	0.03	0.49
	(0.030)	(0.028)	(0.011)	(0.012)		

Notes: See Table A1.

Table A3: Equation-by-Equation Results, Sub-Sample with Positive Male and Female Incomes and Diary Expenditures (n=419)

	Instrumental Variable Results[a]		Ordinary Least Squares Results		Hausman test[b] (*p*-value)	Over-identication[c] (*p*-value)
	$\partial c/\partial y^M$	$\partial c/\partial y^F$	$\partial c/\partial y^M$	$\partial c/\partial y^F$		
Food	0.007	0.199	0.042	0.076	0.91	0.70
	(0.046)	(0.076)	(0.011)	(0.021)		
Beverages, tobacco,	0.053	0.004	0.023	0.018	0.99	0.55
betelnut	(0.031)	(0.045)	(0.010)	(0.015)		
Clothing	0.009	0.042	0.007	0.017	0.97	0.99
	(0.016)	(0.024)	(0.004)	(0.007)		
Rent and utilities	0.022	0.116	0.014	0.047	0.99	0.64
	(0.071)	(0.112)	(0.020)	(0.028)		
Household goods and	0.030	0.012	0.061	-0.039	0.99	0.93
furnishings	(0.043)	(0.049)	(0.044)	(0.039)		
Medical and health	0.015	0.011	0.007	0.012	0.99	0.75
expenses	(0.007)	(0.009)	(0.002)	(0.005)		
Transport and	0.042	0.120	0.039	0.102	0.99	0.93
communication	(0.048)	(0.059)	(0.016)	(0.050)		
Entertainment and	0.015	0.126	0.025	0.049	0.94	0.79
education	(0.026)	(0.043)	(0.008)	(0.015)		
Miscellaneous	0.027	0.079	0.012	0.027	0.04	0.92
	(0.019)	(0.025)	(0.005)	(0.010)		

Notes: See Table A1.

Table A4: : Equation-by-Equation Results, Sub-Sample with Positive Male and Female Incomes and Diary Expenditures (n=419)

	Instrumental Variable Results[a]		Ordinary Least Squares Results		Hausman test[b] (p-value)	Over-identication[c] (p-value)
	$\partial c/\partial x^M$	$\partial c/\partial x^F$	$\partial c/\partial x^M$	$\partial c/\partial x^F$		
Food	-0.117	0.575	0.105	0.365	0.99	0.99
	(0.247)	(0.210)	(0.044)	(0.035)		
Beverages, tobacco, betelnut	0.240	-0.036	0.067	0.148	0.99	0.39
	(0.257)	(0.217)	(0.023)	(0.025)		
Clothing	0.113	0.024	0.049	0.032	0.99	0.99
	(0.100)	(0.080)	(0.018)	(0.015)		
Rent and utilities	0.649	-0.237	0.054	0.131	0.99	0.97
	(0.550)	(0.445)	(0.030)	(0.058)		
Household goods and furnishings	-0.024	0.145	0.061	0.045	0.99	0.94
	(0.159)	(0.137)	(0.008)	(0.021)		
Medical and health expenses	0.079	0.005	0.009	0.033	0.99	0.83
	(0.069)	(0.054)	(0.004)	(0.009)		
Transport and communication	0.165	0.265	0.178	0.280	0.99	0.96
	(0.221)	(0.198)	(0.067)	(0.085)		
Entertainment and education	0.148	0.204	0.073	0.148	0.98	0.85
	(0.203)	(0.189)	(0.025)	(0.040)		
Miscellaneous	0.270	0.028	0.046	0.100	0.96	0.99
	(0.190)	(0.141)	(0.020)	(0.017)		

Notes: See Table A1.