Household Budget Share Distribution and Welfare Implication: An Application of Multivariate Distributional Statistics

Manisha Chakrabarty[1] and Amita Majumder[2]

Abstract

In this paper the consequence of considering "food share" distribution as welfare measure, in isolation from the joint distribution of item budget shares, is examined through the unconditional and conditional distribution of "food share". The statistical properties of the joint distribution of household budget shares are studied through *Dirichlet distribution* in a three commodity set up. The consequence of considering "food share" as a welfare measure is then illustrated through the unconditional *Beta density*, which overlooks the household joint decision making across different consumption categories, and the conditional *Beta* density of "food share" derived from the joint *Dirichlet distribution.* The analysis based on household level rural data for West Bengal, India, for the year 2009-10 shows significant underrepresentation of households by the conventional unconditional "food share" distribution in the higher range of food budget shares that correspond to the lower end of the income profile.

Key words: *Dirichlet* distribution, *Beta* distribution, budget share.

JEL Classification Number: C46, D1, D3, D6.

[1] Associate Professor, Economics Group, Indian Institute of Management Calcutta, India. mchakrabarty@iimcal.ac.in.

[2] Professor, Economic Research Unit, Indian Statistical Institute, Kolkata, India. amita@isical.ac.in. Corresponding author. FAX: +913325778893.

Household Budget Share Distribution and Welfare Implication: An Application of Multivariate Distributional Statistics

1. Introduction

Household budget shares, defined as the share of total household resources spent for purchasing a specific class of goods, contain useful information on household level of living. In traditional Engel curve analysis, consumption expenditure on any particular commodity is analysed using utility-based theory-driven parametric or nonparametric empirical applications with respect to total consumption expenditure in a single equation framework. Of particular interest is the budget share on 'Food', which has led to a vast literature on welfare and poverty analysis and computation of equivalence scales (e.g., Deaton, 1997; Perali, 2002, 2003; Ravallion, 2002). However, for a given level of total expenditure, households make their choices simultaneously on the consumption vector (budget allocation across different commodity groups), although there is a notion of ranking goods according to a hierarchy of needs (Eswaran and Kotwal, 1993). The interdependence in budget allocation and heterogeneity in consumption pattern (Pasinetti, 1981) can indeed be captured through the complete demand systems approach (e.g., the Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980), the Quadratic Almost Ideal Demand System (QAIDS) of Banks, Blundell and Lewbel (1997) and the Exact Affine Stone Index (EASI) system) of Lewbel and Pendakur (2009)), which provide insight into the consumption pattern and can be applied in the context of True Cost of Living Indices (TCLI), optimal commodity taxation and welfare comparison. However, as noted by Barigozzi et al. (2012), the estimation of demand systems or Engel curves compresses household heterogeneity to the knowledge of the first few moments of household expenditure-level or budget-share distribution for the commodity under study. They also note that (i) to fully characterize such heterogeneity of household consumption pattern, one should perform distributional analyses that carefully investigate how the shape of household consumption expenditure and budget share distributions behave [see Chakrabarty et. al (2006)]; and (ii) a data-driven approach focused on distributional analysis may help to discover fresh stylized facts related to how households allocate their consumption expenditures.

It is with this background that we approach the problem of welfare analysis from a purely statistical viewpoint in terms of distribution of household budget shares. In particular, we explore the statistical properties of household budget share distributions and attempt to assess the impact on the analysis of welfare, when decision on food expenditure is isolated from the joint decision of budget allocation, by looking at the joint distribution of budget shares vis-a-vis the unconditional marginal distribution of budget share on food. The conditional distribution of food share obtained from this joint distribution provides the true expenditure distribution pattern of households under the ceteris paribus condition. A comparison of the welfare rankings based on the quantiles of conditional distribution of food expenditure derived from the joint distribution with the quantiles of the conventional food expenditure share distribution throws light on the severity of the consequences in measuring welfare by ignoring the interdependent nature of expenditure pattern.

The paper is organised as follows: Section 2 describes the joint, conditional and marginal distributions of “food” in a 3-commodity set up. Section 3 discusses the data and results. Finally, Section 4 concludes the paper.

2. The Distributions

Consider 3 commodities, viz., food and two others, with budget shares w_1, w_2 and w_3, respectively, such that $(w_1 + w_2 + w_3) = 1$. The joint probability density function (p.d.f.) can then be assumed to be a *Dirichlet Distribution* given by (Kotz et al., 2000)[3]:

$$f(w_1, w_2, w_3) = \frac{\Gamma(m+n+p)}{\Gamma(m)\Gamma(n)\Gamma(p)} w_1^{m-1} w_2^{n-1} {w_3}^{p-1}, \quad (2.1)$$

where *m, n, p > 0,* $w_1 \geq 0, w_2 \geq 0$ and $w_3 = (1 - w_1 - w_2) \geq 0.$

The conditional p.d.f. of w_1 given $w_2 = W$ is given by

$$g(w_1|W) = \frac{\Gamma(m+p)}{(1-W)\Gamma(m)\Gamma(p)} (\frac{w_1}{1-W})^m (1 - \frac{w_1}{1-W})^p, \quad (2.2)$$

with $0 < w_1 < 1 - W, \quad 0 < W < 1.$

The unconditional p.d.f. of food budget share can be assumed to follow a *Beta Distribution*[4], which is the marginal distribution of the *Dirichlet Distribution,* given by (Forbes et al., 2011)

$$h(w_1) = \frac{\Gamma(q+r)}{\Gamma(q)\Gamma(r)} w_1^{q-1} (1 - w_1)^{r-1} \quad (2.3)$$

[3] The Dirichlet distribution is a multivariate distribution whose components take values on (0,1) and which sum to one , which is satisfied for the budget shares. The automatic interdependence is captured by this multivariate Dirichlet distribution.

[4] Barigozzi et al. (2012) also chose Beta density as a benchmark to evaluate the performance of two household budget share distributions they proposed.

where $q, r > 0$, $0 < w_1 < 1$.

3. Data and Results

This study uses the household level consumption expenditure data contained in the unit records from the 66th (July, 2009- June, 2010) round of India's National Sample Surveys. For illustrative purpose data on the state of West Bengal (rural sector) have been chosen. The detailed expenditures on individual commodities have been aggregated into expenditure on three commodity groups, namely, 'food', "durables" and the rest (which we call 'others'). This group consists of clothing and footwear, pan, tobacco and intoxicants, personal care, fuel and light, rent and taxes, medical expenses, education expenditure and transport expenses. In other words, these three commodity groups comprise the total household budget.

First, a *Dirichlet Distribution* is fitted to the 3-commodity budget share data (food: w_1; others: w_2; durables: w_3). Table 1 presents the estimates of the parameters along with the corresponding standard errors. All parameters are significant at 1% level.

Next, the conditional distribution of the budget share of 'food' is obtained using (2.2), given a particular value of the budget share for 'others' (the budget share of the remaining commodity 'durables' is automatically determined by these two). This is repeated for a number of values of the budget share for 'others'. The parameters of all the conditional distributions, presented in Table 2, are significant at 1% level.

Finally, the unconditional *Beta Distribution,* given by (2.3) is fitted to the budget share for 'food'. The parameter estimates, along with the corresponding standard errors are presented in Table 3. Here again, the parameters are significant at 1% level. Figure 1 gives a comparison of the non parametric density and the fitted Beta density. It is quite clear that the Beta distribution is very close to the nonparametrically estimated density.

We now turn to the welfare implication under the conditional and unconditional distributions. Fixing a particular value of food share, we look at the proportion of persons at that value under the two alternative scenarios. Table 4 presents these results. Given that the difference of two Uniform variates on the interval [0,1] follows a Triangular distribution[5], we looked at

[5] This distribution of $X = |X_1 - X_2|$, where X_1, X_2 are two independent random variables with standard uniform distribution is characterized by $f(x) = 2 - 2x, 0 \leq x < 1;\ E(x) = \frac{1}{3}, \sigma^2 = V(x) = \frac{1}{18}$.

the 1σ interval of the absolute difference for each value of food share to see if it contains the value zero, in which case we can say that there is no difference. It is observed from the last column of Table 4 that the interval clearly rejects the null hypothesis of no difference $H_{0:}\ |conditional - unconditional| = 0$ for higher ranges of food share.

A graphical representation is provided by Figures 2(a) and 2(b), which give the cumulative distribution functions (c.d.f.'s) and the probability density functions (p.d.f.'s), respectively, under alternative scenarios. It is evident from Figure 2(a) that the distribution of food budget share moves leftwards as 'food' budget share is conditioned by higher and higher values of 'others' budget share[6]. It can also be seen that for a given value of food share in certain ranges, particularly in some higher ranges, the value of the unconditional c.d.f. is more than that of the conditional c.d.f.[7] This implies that the proportion of people above a particular level will be underestimated by the unconditional distribution. If food share, satisfying Engel's law, is taken as a measure of welfare with higher food share indicating lower rank in the welfare measurement, this phenomenon is of particular relevance in countries like India, where majority of the households have high food shares.

4. Concluding Remarks

In this paper the statistical properties of the 3-dimensional household budget share distribution is studied parametrically using *Dirichlet distribution*. This allows one to incorporate into the study the underlying automatic across-budget share correlation structure, which is embodied in the estimates of density parameters. The consequence of considering "food share" as a welfare measure in isolation from this joint distribution is illustrated through the conventional unconditional and conditional density of "food share". It may, however, be pointed out that the dependency of the budget share distribution on household income and other demographic characteristics is ignored here. We propose to extend this study incorporating this feature of conditioning of the distributions (both marginal and joint) using a non-parametric set-up in a future project.

[6] The main purpose here is to illustrate the distortion in the distribution of food budget share when considered in isolation. Here it is illustrated with respect to 'others'.

[7] Note that the range of food share for the conditional distribution is given by $0 < w_1 < 1 - W, \quad 0 < W < 1$, where W defines the $ceteris\ paribus$ condition.

References

Banks, J., R. Blundell and A. Lewbel (1997): ‘Quadratic Engel curves and consumer demand’, *The Review of Economics and Statistics,* 79, 527–539.

Barigozzi, Matteo, Alessi, Lucia, Capasso, Marco and Fagiolo, Giorgio (2012): ‘The distribution of household consumption-expenditure budget shares’, *Structural Change and Economic Dynamics,* 23, 69-91.

Chakrabarty, M , Anke Schmalenbach and Jeffrey Racine, 2006. "On the distributional effects of income in an aggregate consumption relation," *Canadian Journal of Economics*, Canadian Economics Association, vol. 39(4), pages 1221-1243, November.

Deaton, A. (1997): *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, World Bank Publications.

Deaton, A., and Muellbauer, J. (1980): ‘An almost ideal Demand System’, *American Economic Review*, 70, 312-26.

Eswaran, M. and Kotwal, A. (1993): ‘A theory of real wage growth in LDCs’, *Journal of Development Economics*, 42, 243-269.

Forbes, C., Evans, M., Hastings, N. and Peacock, B. (2011): *Statistical Distributions*. Hoboken, NJ: John Wiley.

Kotz, S., Balakrishnan, N., Johnson, N. L. (2000): *Continuous Multivariate Distributions: Volume 1*. New York: John Wiley.

Lewbel, A. and Pendakur, K (2009): ‘Tricks with Hicks: The EASI Demand System’, *American Economic Review*, 99, 827-863.

Pasinetti, L. L. (1981): *Structural Change and Economic Growth* (Cambridge: Cambridge University Press).

Perali, Federico, (2002): ‘Some curiosities about the Engel method to estimate equivalence scale’. *Economics Bulletin,* 4, 1–7.

Perali, Federico, (2003): *The Behavioural and Welfare Analysis of Consumption: The Cost of Children, Equity and Poverty in Colombia*, Kluwer Academic Publishers.

Ravallion, Martin (2002): *Poverty Comparisons*, Routeledge.

Table 1: Parameter estimates for the joint *Dirichlet* distribution of "food share", "others share" and "durables share" (Equation 2.1)

Parameters	
M	13.524*** (0.255)
N	8.566*** (0.161)
P	0.867*** (0.016)

Number of observations: 3227 .Figures in parentheses indicate the standard errors.
***Significant at 1% level

Table 2: Parameter estimates for the conditional distribution of "food share" for different cut-off points of "others share" (Equation 2.2)

Percentile cut-off for "others"	Cut off values of "others share"	Number of observations	Parameters	
			m	*p*
10th	*W*=0.27	3039	14.132*** (0.273)	0.925*** (0.017)
25th	*W*=0.31	2740	14.118*** (0.287)	0.984*** (0.019)
50th	*W*=0.36	2128	13.225*** (0.303)	1.060*** (0.023)
75th	*W*=0.42	1320	10.582*** (0.307)	1.105*** (0.031)

Figures in parentheses indicate the standard errors.***Significant at 1% level

Table 3: Parameter estimates for the unconditional distribution of "food share" (Equation 2.3)

Parameters	
Q	12.812*** (0.433)
R	8.918*** (0.273)

Figures in parentheses indicate the standard errors.***Significant at 1% level

Table 4: Value of c.d.f. under conditional and unconditional distribution

Budget share of food	Parameters of conditional distribution		Value of c.d.f		Absolute Difference with 1σ confidence interval ©
	m	*P*	Conditional	Unconditional*	
0.70	14.132	0.925	0.5158	0.8537	0.3379 (0.1022, 0.5736)
0.65	14.118	0.984	0.4230	0.7090	0.2860 (0.0503, 0.5217)
0.60	13.225	1.060	0.4525	0.5285	0.076 0 (-0.1597 0.3117)

*Parameters: q = 12.8122, r = 8.9183.

© The absolute difference between two uniformly distributed random variables follows a triangular distribution and the confidence interval does not contain zero except for food share 0.60.

Figure 1: Nonparametric and Unconditional Beta (12.8122, 8.9183) Density for Budget share of Food

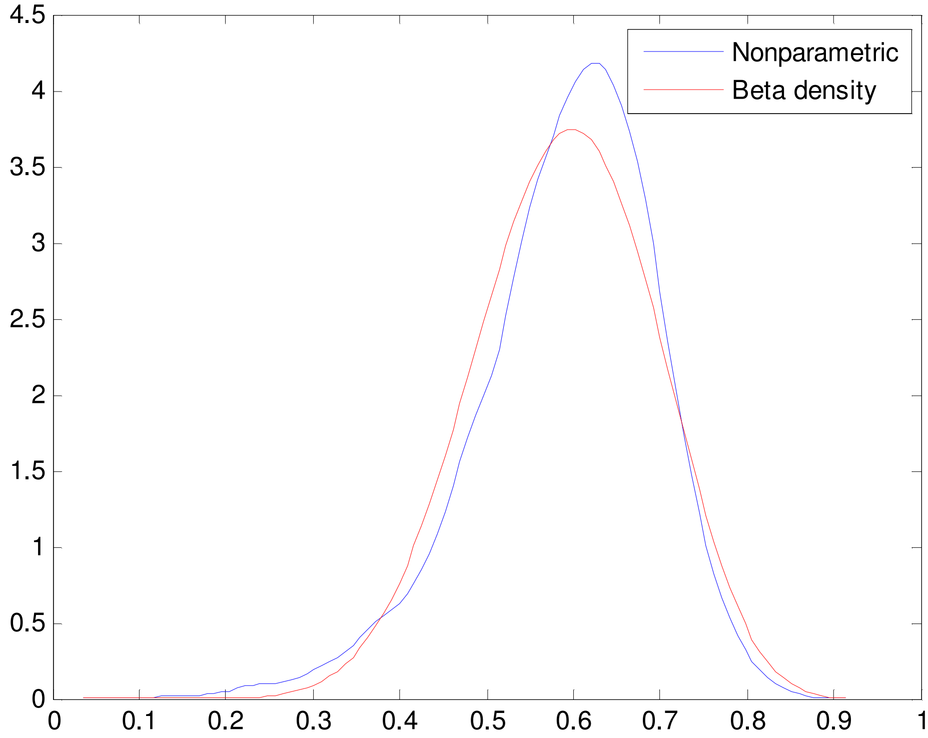


Figure 2(a):

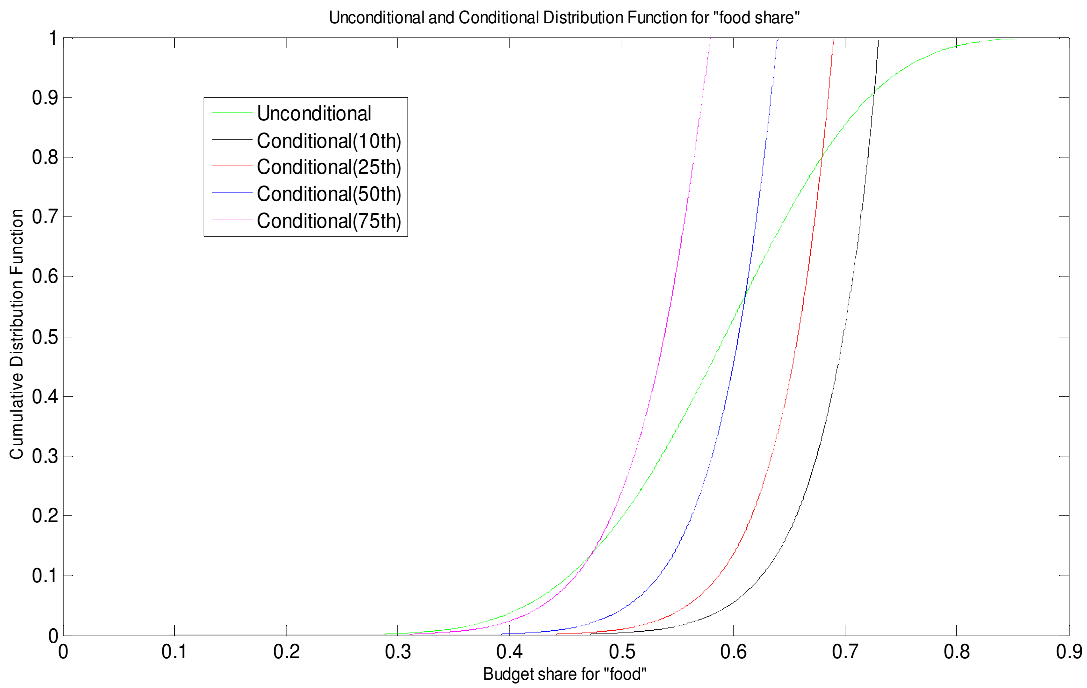
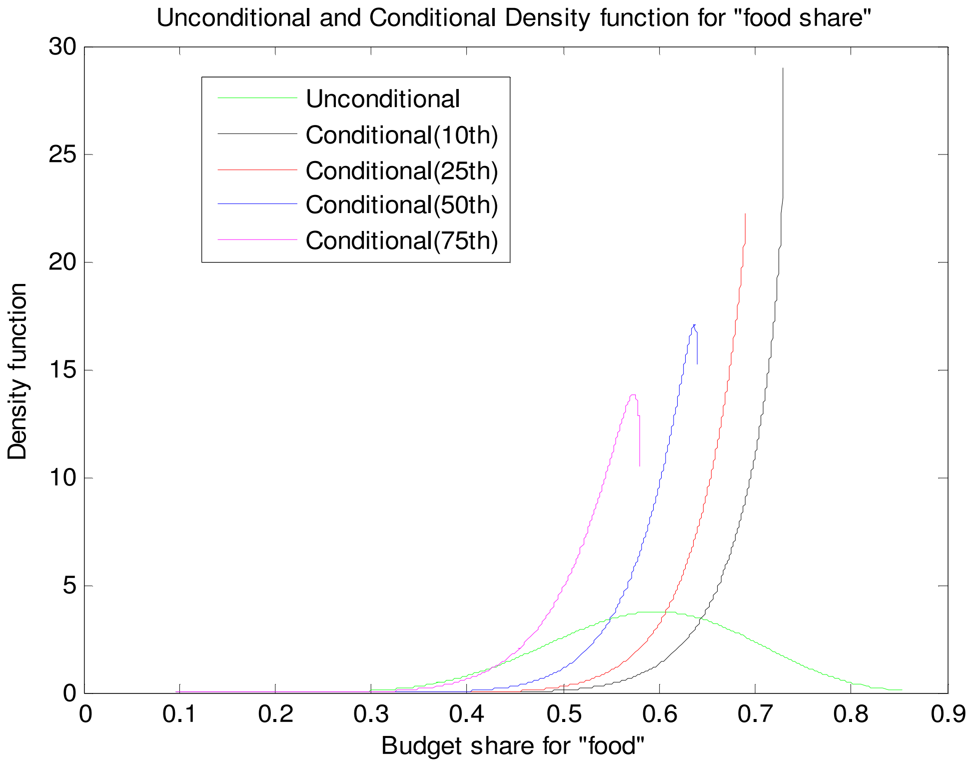


Figure 2(b)


Unconditional and Conditional Density function for "food share"
Unconditional
Conditional(10th)
Conditional(25th)
Conditional(50th)
Conditional(75th)
Density function
Budget share for "food"
0
5
10
15
20
25
30
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9