

## Does Survey Recall Error Explain the Deaton-Paxson Puzzle?

Matthew Brzozowski  
York University, Toronto  
brzozows@yorku.ca

Thomas F. Crossley  
University of Essex  
and  
Institute for Fiscal Studies, London  
tcross@essex.ac.uk

Joachim K. Winter  
University of Munich  
winter@lmu.de

April 2017

**Abstract:** Using recall and diary food expenditure data from Canada, we compare estimates of the household size elasticity of per capita food expenditure. In contrast to Gibson (2002), we find negative elasticities in both recall and diary data. This in turn means we find evidence of the “Deaton-Paxson puzzle” in both diary and recall data. Recall error cannot be the sole explanation of the puzzle.

**Keywords:** Household economics, economies of scale, consumer demand

**JEL classification:** C81, D12

## Introduction

In applied demand analysis, the income and household size elasticities of food expenditure play an important role, particularly in thinking about the economies of scale in household consumption. An assertion due to Engel is that households of different size with the same food budget share have the same standard of living. This leads to the “Engel method” of calculating economies of scale in household consumption. Suppose, for the purposes of illustration, that the food budget share is adequately modelled by

$$w_f = \alpha_0 + \alpha_1 \ln pcy + \beta \ln n + \varepsilon$$

where  $w_f$  is the food share,  $\ln pcy$  is the logarithm of per capita income, and  $\ln n$  is the logarithm of household size. Thus to hold living standards (the food share) equal as household size doubles (increases by 100%), per capita income should change by (approximately)  $-(\beta / \alpha_1) \times 100\%$ . Economies of scale imply that the per capita income required to keep living standards constant should fall with household size. Empirically,  $\alpha_1$  is always negative (this is “Engel’s Law”). Thus, if the food share can be taken as a welfare measure (as Engel asserted), economies of scale require that  $\beta$  be negative (the budget share should fall with increasing household size, holding  $pcy$  constant). Empirically, this turns out to be the case. For example using Thai, Pakistani, South African, US, French and British data, Deaton and Paxson (1998) find that, holding per capita income constant, the food share varies inversely with household size. The Engel method delivers estimates of the economies of scale in consumption that many researchers find plausible.

Against this, Deaton and Paxson (1998) demonstrate that it is quite difficult to reconcile a negative  $\beta$  (and the Engel method) with an underlying model of household economies of scale (see Deaton, 2010). They note that, if there are public goods in the household, then holding per capita income constant, a larger household is better off. This

should lead them to consume more of private goods, such as food.<sup>1</sup> Thus, holding per capita income constant, the per capita quantity of food, and hence the budget share, should rise. Thus  $\beta$  (and  $\beta / w_f$ , the elasticity of food expenditures with respect to household size) should be positive. The fact that this compelling piece of analysis is empirically contradicted is referred to as the “Deaton-Paxson puzzle.”

Gibson (2002) suggests that one possible explanation for the Deaton-Paxson puzzle is measurement error in recall food expenditure data that is negatively correlated with household size. For larger households, it becomes an increasingly cumbersome task to accurately recall all food related purchases made over even a modest time period. Thus the larger the household, the more likely is systematic underreporting of food expenditure. A negative correlation between the measurement error and household size imparts a negative bias on the estimated relationship between the food share and household size.

Many of the surveys examined by Deaton and Paxson do employ recall methods to collect food expenditures, and Gibson suggests that the Deaton and Paxson puzzle might be resolved by using diary based food expenditures. He uses data from Papua New Guinea (PNG) to test the validity of this prediction. Households were randomly divided into two subsamples; one was asked to keep a diary while the other was asked recall questions. His results suggest that while recall data underestimates household size elasticities, estimates based on diary data do not exhibit the Deaton-Paxson puzzle.

## **Data**

The 1996 Canadian Food Expenditure Survey (FoodEx) provides a unique opportunity to study how food expenditure measures constructed from recall questions compare to those obtained from expenditure diaries. This nationally representative survey first asked respondents to estimate their household’s food expenditure over the past four

---

<sup>1</sup> This assumes limited substitution between food and the public good.

weeks, along with basic demographic questions. They were then asked to record daily food expenditure in two consecutive weekly diaries. The survey involved three visits to each household. At the initial visit, demographic and recall food consumption questions were asked. The weekly diaries were collected at subsequent visits. The interviewers double-checked diaries and verified the quality of the responses. The survey was run throughout the year. The initial response rate was 76 percent, and there were 10898 responding households. The non-response rate to the recall question was less than 2 percent. Attrition between the first and second week of the diary was less than 2 percent. Weights are provided that account for the survey design and non-response, but not for the attrition between the two weeks.

We can also compare the FoodEx to data from a second large Canadian survey. The 1996 Family Expenditure Survey (FamEx) is a full household expenditure survey (collecting information on all categories of expenditure).<sup>2</sup> Face-to-face interviews were conducted in the first quarter of 1997 to collect income and expenditure information for the previous year. Statistics Canada undertakes various checks of the data and the data are generally thought to be of very good quality.<sup>3</sup> There are 10085 respondent households in the 1996 FamEx.<sup>4</sup> Because the FamEx collected annual data and the FoodEx survey ran continuously over the year, they refer to the same time period. The surveys were based on the same (Labour Force Survey) sampling frame. Thus these two surveys readily lend themselves to comparison.<sup>5,6</sup>

---

<sup>2</sup> The FamEx surveys were used to determine the weights for the Consumer Price Index in Canada.

<sup>3</sup> Further details on the quality of this data are in Brzozowski and Crossley (2011).

<sup>4</sup> The response rate to the FamEx surveys is about 75%.

<sup>5</sup> FoodEx measures were converted to annual values. For detailed comparison of the consumption measures used in this paper, see Brzozowski et al. (2016). Also see that paper

## Results

We estimate food share equations that are a quadratic extension of the Working-Leser form,

$$w_f = \alpha_0 + \alpha_1 \ln pcy + \alpha_2 (\ln pcy)^2 + \beta \ln n + \gamma X + \varepsilon$$

where  $w_f$  is the budget share of food at home,<sup>7</sup>  $\ln pcy$  is the logarithm of per capita income,  $\ln n$  is the logarithm of household size, and  $X$  are other variables. We estimate this equation using two data sets and three measures of the food share. First, we use a food share based on the average of the diary weeks in the FoodEx. Second, we use a food share based on the (1 month) recall measure in the FoodEx. Third, we use a food share based on the (1 year) recall measure in the FamEx. The results are presented in Table 1.

---

for a discussion of differences across surveys in the construction of the household income variable and in top coding of household size.

<sup>6</sup> To deal with potential outliers we trimmed the top and bottom 2% of expenditure reports.

<sup>7</sup> We define the food at home budget share as expenditure on food at home divided by gross income. While total outlay is the preferred denominator, gross income is the measure of resources that we have in both surveys. In demand analysis, it would be common to use total outlay both to construct the budget share and as an explanatory variable but then to instrument total outlay with income to mitigate endogeneity and attenuation due to measurement error. Measurement error in income may lead to some attenuation bias in our estimates but this would be common to diary and recall food expenditure measures.

TABLE 1  
Food at Home Budget Share Regressions

	FoodEx Diary		FoodEx Recall		FamEx Recall	
Food Budget Share ( $w_f$ )	.106		.124		.125	
Variable	Coef	(Std Err)	Coef	(Std Err)	Coef	(Std Err)
ln pcy	-0.44	(0.037)	-0.71	(0.037)	-0.616	(0.031)
(ln pcy) <sup>2</sup>	0.019	(0.002)	0.032	(0.002)	0.027	(0.002)
ln household size	-0.007	(0.003)	-0.023	(0.003)	-0.003	(0.002)
Test for common coefficients			F- stat	(p-value)	F- stat	(p-value)
ln pcy, (ln pcy) <sup>2</sup> – vs FoodEx Diary			31.79	(<0.001)	25.78	(<0.001)
ln household size – vs FoodEx Diary			1.80	(0.180)	12.83	(<0.001)
ln pcy, (ln pcy) <sup>2</sup> – vs FoodEx Recall					2.27	(0.103)
ln household size – vs FoodEx Recall					27.85	(<0.001)
Household Size Elasticities (computed at mean budget share)						
$(\partial \ln pce_f / \partial \ln n) =$ $(\partial w_f / \partial \ln n) \cdot (1 / w_f)$	-0.066 (0.03)		-0.185 (0.02)		-0.024 (0.02)	

## Notes:

- FoodEx Diary is average of 2 weeks
- Additional control variables ( $X$ ) include regional dummies, dummies for presence of children, youth and seniors, and presence of a 2<sup>nd</sup> earner in the household.
- Survey weights are used in all estimation. Robust standard errors are reported in parentheses in rows one through three. (In rows four through seven the number in parentheses is the p-value of the corresponding F-test.)
- Elasticities calculated at the means of the data.

We find that the food share varies inversely with household size in all three cases. The coefficient on log household size is -0.007 with the FoodEx diary data, -0.023 with the FoodEx recall data, and -0.003 with the FamEx recall data (3<sup>rd</sup> row, 2<sup>nd</sup> panel, Table 1.) The first two estimates are different from zero at conventional levels of statistical significance, while the third is not. Although the estimates are of the same sign, F-tests do indicate that the FamEx recall estimates are statistically different from both FoodEx estimates (2<sup>nd</sup> and 4<sup>th</sup> row, 3<sup>rd</sup> panel, Table 1).<sup>8</sup> The implied elasticities are presented in the 4<sup>th</sup> panel of Table 1. The bottom line is that we find the Deaton-Paxson puzzle with both recall and diary data. Thus our data are incongruent with Gibson's resolution of the puzzle.

## **Discussion**

In response to an early version of our analysis of the Foodex, Gibson and Kim (2007) propose an explanation for the contrast with the results reported in Gibson (2002) and Gibson and Kim (2007). They postulate that because Foodex respondents are asked a broad question about total household food expenditure, they are more likely to employ an estimation-based response strategy rather than enumeration (of actual purchases). In contrast, the surveys

---

<sup>8</sup>To implement these tests we treat the diary and recall expenditure reports data as separate observations, effectively a panel of two observations on each household in the Foodex, and then pool the data, including the Famex. This gives an unbalanced panel, with two observations on some households and one observation on others. We then estimate a regression model with full interactions between "Foodex Recall" and "Famex" dummies and all other variables. In estimating this model, we calculate cluster robust standard errors with clustering at the household level, to allow for the obvious correlation between the responses of the same households. We then tested the interaction terms (jointly where appropriate) using the cluster-robust covariance matrix.

studied in Gibson (2002) and Gibson and Kim (2007) ask more detailed questions about expenditure on different food categories, and so respondents may be more likely to enumerate actual purchases. Errors in estimation may be less strongly related to household size than errors in recalling actual purchases. This is an appealing argument, and it would clearly be useful to better understand respondents' use of estimation and enumeration strategies when asked recall questions, and the nature of the errors associated with each. Moreover, other differences between surveys and settings, including the definition of food expenditure, the mix of food consumed inside and outside the household, the role of home-produced food and average household sizes, could mean that the nature of recall error is quite different in Canada and PNG. However, neither the suggestion that Foodex respondents employ an estimation-based response strategy to recall questions, nor other factors that might matter for recall error, explain why we find evidence of the Deaton-Paxson puzzle in *diary* data. Our results suggest that the Deaton-Paxson puzzle must arise, at least in some instances, for reasons other than (or in addition to) recall error.

Finally, if we employ the "Engel Method" to estimate returns to scale, the FamEx recall data imply that a doubling of household size allows a 3% cut in per capita income, while the FoodEx diary and recall data give estimates of 9% and 24% respectively.<sup>9</sup>

## **Conclusion**

In an application drawn from demand analysis, we compared estimates of household size elasticities of food expenditure based on recall and diary food expenditure data. We find negative household size elasticities with both kinds of data. This leads us to doubt the generality of a resolution of the Deaton-Paxson puzzle proposed by Gibson (2002).

## **Acknowledgements**

---

<sup>9</sup> Computed as  $\frac{\beta}{\alpha_1 + 2 \alpha_2 \ln pcy}$  at the mean value of  $\ln pcy$ .



Author order is alphabetical. This paper draws on material that previously circulated in unpublished form as Ahmed, Brzozowski and Crossley (2010), which in turn drew on material from Brzozowski's PhD. Dissertation at McMaster University. The analysis was carried out in the Statistics Canada Research Data Centre at McMaster University and the authors thank James Chohan for his assistance in accessing the data. The authors gratefully acknowledge financial support from the Social and Economic Dimensions of an Aging Population (SEDAP) Research Program at McMaster University, the Arts Research Board of McMaster University, and the Social Sciences and Humanities Research Council of Canada (SSHRC). Crossley also gratefully acknowledges support from the ESRC-funded Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (grant number RES-544-28-5001.)

## References

- Ahmed, N., M. Brzozowski, and T. F. Crossley (2010). "Measurement errors in recall food consumption data." Institute for Fiscal Studies Working Paper 06/21, London.
- Brzozowski, M. and T. F. Crossley, (2011). "Measuring the well-being of the poor with income or consumption: A Canadian perspective." *Canadian Journal of Economics*, 44(1), 88–106.
- Brzozowski, M., T. F. Crossley and J. Winter (2016). "Comparison of recall and diary food expenditure data", *Food Policy*, forthcoming.
- Deaton, A. (2010). "Understanding the mechanisms of economic development." *Journal of Economic Perspectives*, 24(3), 3–16.
- Deaton, A. and C. Paxson (1998). "Economies of scale, household size, and the demand for food." *Journal of Political Economy*, 106(5), 897–930.
- Gibson, J. (2002). "Why does the Engel method work? Food demand, economies of size and household survey methods." *Oxford Bulletin of Economics and Statistics*, 64(4), 341–359.
- Gibson, J. and B. Kim (2007). "Measurement error in recall surveys and the relationship between household size and food demand." *American Journal of Agricultural Economics*, 89(2), 473–489.