MEASUREMENT ERRORS IN RECALL FOOD CONSUMPTION DATA

Naeem Ahmed
Matthew Brzozowski
Thomas F. Crossley

THE INSTITUTE FOR FISCAL STUDIES
WP06/21

Measurement Errors in Recall Food Consumption Data

Naeem Ahmed, Matthew Brzozowski and Thomas F. Crossley*

This Version: October 2006

---

* *Ahmed*: Department of Economics, McMaster University. *Brzozowski*: Department of Economics, University of Western Ontario. *Crossley:* Faculty of Economics, University of Cambridge, Institute for Fiscal Studies, and Department of Economics, McMaster University.
Author order is alphabetical. The authors gratefully acknowledge the support of 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). SEDAP is primarily funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) and also receives substantial additional support from Statistics Canada. The analysis reported in this paper was carried out in the Statistics Canada Research Data Centre at McMaster University and the authors thank James Chowan for his assistance in accessing the data. The authors also thank Sule Alan, Chris Carroll, Martin Dooley and Michael Veall for helpful comments. Any remaining errors are the responsibility of the authors.
Please check http://socserv.mcmaster.ca/crossley/research.htm for updates.
Correspondence: Thomas Crossley, Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge CD3 9DD, UK, phone: +44 (0) 1223 335225, fax: +44 (0) 1223 335475, email: Thomas.Crossley@econ.cam.ac.uk.

Measurement Errors in Recall Food Consumption Data

***Abstract*:** Recall food consumption data, which is the basis of a great deal of empirical work, is believed to suffer from considerable measurement error. Diary records are believed to be very accurate. We study a unique data set that collects recall and diary data *from the same households*. Measurement errors in recall food consumption data appear to be substantial, and they do not have the properties of classical measurement error. We also find evidence that the diary measures are themselves imperfect. We consider the implications of our findings for modelling demand, measuring inequality, and estimating inter-temporal preference parameters.

***Keywords*:** expenditure, consumption, measurement error, survey data

JEL classification: C81, D12

Executive Summary

- Information on household expenditure or “consumption” is crucial for a broad range of economic research, including research on consumption and savings behaviour, on poverty and inequality and on living standards.
- A great deal of existing research in these areas has been based on food consumption measures, and particularly, on recall questions about food consumption. Such questions can be asked cheaply and do not substantially increase the response burden of a survey. Collecting consumption or expenditure information via diaries is thought to be much more accurate. However, it is also more costly and substantially increases the burden on survey respondents.
- We study a unique data set that collects recall and diary data on food consumption *from the same households*. Assuming that the diary records are accurate, we can therefore calculate a recall error for each household. This allows us to study the distribution of these errors, to determine their magnitude, and whether they have the properties are usually assumed in econometric modelling (for example, the properties of “classical” measurement error.)
- Measurement errors in recall food consumption data appear to be substantial, and they do not have the properties of classical measurement error. However, we also find evidence that the diary measures are themselves imperfect.
- We consider the implications of our findings for modelling demand, measuring inequality, and estimating inter-temporal preference parameters.

I. Introduction

Information on household expenditure or "consumption" is crucial for a broad range of economic research, including research on consumption and savings behaviour, on poverty and inequality and on living standards.[1] Measurement error in consumption data has been an important concern in these literatures. A great deal of existing research in these areas has been based on food consumption measures. There are at least two reasons for this. First, there is a long tradition of treating food consumption as a welfare measure. Second, and more practically, response load considerations have led surveys that have a panel structure, or that collect other important information from households, to collect only limited consumption information. Such surveys usually do ask a recall food consumption question. Well-known examples are the Panel Study of Income Dynamics (PSID) and the British Household Panel Survey (BHPS). Thus measurement error in food consumption data is of particular interest. This paper provides new evidence on the extent and character of measurement error in food consumption data.

Concern with measurement error in consumption data has been prominent in the recent literatures on inequality, and on demand. For example, Battistin (2004) explores differences in the evolution of apparent consumption inequality between the diary and interview (recall) samples of the U.S. Consumer Expenditure Survey. He shows that the interview data suggest that consumption inequality rose during the 1980s but not during the 1990s, while data based on diaries alone or on optimal (under some assumptions) combination of recall and diary records suggest that consumption inequality continued to rise during the 1990s.

---

[1] We shall initially assume that food consumption and expenditure are the same, as often done for food, and other nondurables. We return to this issue, however, below.

Gibson (2002) analyzes a small survey from Papua New Guinea in which a random half of the respondents were posed a recall food consumption question and the other half asked to complete a food consumption diary. Gibson claims that a puzzle regarding the relationship between household size and food consumption that was first highlighted by Deaton and Paxson (1998) can largely be explained by measurement error in recall food consumption data that is correlated with household size.

The inter-temporal consumption literature is very largely based on food consumption data from the PSID. The belief that such consumption data contains significant measurement error (Altonji and Siow, 1987, Runkle, 1991), and the difficulty of estimating nonlinear models in the presence of measurement error (Amemiya, 1985), has led to the extensive use of linear (in log) approximations to the consumption Euler equation as a basis for estimation. However, the use of such approximations introduces other, equally difficult, problems, as emphasized by Carroll (2001), Ludvigson and Paxson (2001), Attanasio and Low (2004), and Alan and Browning (2003).[2] This has led some (see Carroll, 2001, for example) to call for the complete abandonment of Euler equation estimation. Recent approaches have returned to the exact (nonlinear) consumption Euler equation, but employed specific assumptions about the measurement error. For example Colera (1993) assumes the measurement error is multiplicative and log-normally distributed. Alan, Attanasio and Browning (2005) develop an estimator that requires only the assumption that the mean of the measurement errors be constant over time (for each household.)

[2] A key problem is that omitted higher order terms in the approximation enter the error term. Unlike the innovation in marginal utility between *t* and *t+1*, theory does not require that these approximation errors be orthogonal to time *t* information. Thus these terms make it much more difficult to find valid instruments.

Note that this implies that the errors are uncorrelated (again over time, for each household) with the true value of consumption.

The Canadian Food Expenditure Survey (FoodEx) asks respondents to first estimate their household's food expenditure over the past four weeks, and then to record food expenditure in a diary for two weeks. Thus this survey provides an ideal opportunity to directly compare recall and diary methods of collecting food consumption data. Existing research on measurement error in consumption often compares data from different surveys (for example, Battistin, Miniaci, and Weber (2001) and Browning, Crossley and Weber, (2003)) in which case corrections must be made for differences in sample design, etc. Battistin (2004) and Gibson (2002) both use a single survey, but different samples. While this allows for a comparison of distributions, it does not allow for an examination of the distribution of differences between recall and diary records. In contrast, the Foodex Data allow us to calculate a recall error for each household, and to examine the properties of those errors directly.[3]

[3] Gibson suggests (2002) that a possible problem with comparisons such as the one allowed by the FoodEx is that the beginning of the recall period is not marked by a visit from an interviewer, whereas the diary period is. This may lead to "telescoping errors" in the recall data. While we agree that this is a possible problem with the recall question, it seems to us that since almost all recall expenditure questions share this possible problem, the FoodEx allows the appropriate comparison: between diary collection and recall information *as usually collected*. A study of recall expenditure data from a survey in which the recall measure was marked by a visit from an interviewer would not be as informative about the recall expenditure data in, for example, the PSID.

In their *Handbook of Econometrics* survey, Bound et. al. (2001) emphasize that while econometric methods for dealing with measurement error typically assume that measurement errors are "classical", much of the available empirical evidence contradicts this assumption. They also emphasize the usefulness of validation data in characterizing the joint distribution of error-ridden measures and their true values, and for testing the assumption of classical measurement error or other assumptions about measurement error. Bound et. al. report evidence on measurement error in a variety of constructs (for example wages and earnings) but not food consumption. While the FoodEx is not a designed validation study, the fact that it has recall and diary measures from the same households makes it a good approximation to a validation study, and allows us to carry out similar analyses.

The next section of this paper describes the Canadian Food Expenditure survey as well as a second, more widely used Canadian expenditure survey (the Family Expenditure Survey or FamEx), which also collects recall food consumption data. This section also provides a preliminary analysis of the different food consumption measures available in the two surveys.

In Section 3, we calculate errors in recall food consumption, using the diary measures to construct "true" food consumption in a number of different ways. Under the assumption that true food consumption can be constructed from the diary records, measurement errors in recall food consumption data appear to be substantial, and they do not have the properties of classical measurement error. In particular, they are neither mean independent of true consumption nor homoskedastic. They are also not well approximated by a normal distribution. We also show evidence that diary measures are themselves imperfect. This suggests alternative interpretations for the differences between recall and diary consumption measures.

In Section 4 we compare estimates of income and household size elasticities of per capita food consumption based on the two kinds of consumption data. Here, we find a more positive result. In contrast to Gibson (2002), we find that the mode of data collection makes very little difference to estimates of income and household size elasticities. This in turn means that (in contrast to Gibson) we find the evidence of the “Deaton-Paxson puzzle” both in the diary and in the recall data.

Section 5 draws out the implications of the measurement error patterns we document for two other common applications of consumption data: the estimation of inter-temporal preference parameters and the measurement of inequality. Finally, Section 6 offers some concluding remarks.

II. Canadian Expenditure Surveys

The 1996 Canadian Food Expenditure Survey (FoodEx) is a large, nationally representative survey of Canadian households. Respondents were asked basic demographic questions and recall food consumption questions. In addition, they were asked to record every food purchase in a diary, for two contiguous weeks. Conducting the survey involved three visits to each household. At the initial visit, demographic and recall food consumption questions are asked. In addition, respondents were instructed on the proper technique for filling out the food consumption diaries. After a week the first diary was collected and the household received another second blank diary in which to record purchases made in the following week. This second diary was collected during the third visit. During the second and third visit the interviewers double-checked the diaries and verified the exactness and fullness of the responses. The survey was run continuously throughout the year so that the seasonality of purchases is not an issue. The initial response rate was 76 percent, and there were 10898 responding households.

Attrition between the first and second week was less than 2 percent.[4] Statistics Canada provides household weights that take account of the survey design and non-response, but not of attrition between the two weeks. Further details can be found in Statistics Canada (1999).

For the purposes of this paper, the key feature of the survey is that each household is asked recall food expenditure questions as well as recording food expenditure diaries. The exact wording of the key recall food expenditure questions is as follows:

In the last four weeks…

> *Q1. How much do you estimate this household spent on food and other groceries purchased from stores (including farmer stalls and home delivery)? Exclude periods away from home overnight or longer. Report bulk purchases of food for canning, freezing in question 3.*
>
> *Q2. About how much of this amount was for non-food items such as paper products, household supplies, pet food, alcoholic beverages, etc.?*

This differs somewhat from the question in the PSID, particularly in that it refers to the last four weeks, while the PSID refers to the amount the household "usually" spends on food at home. We construct recall food consumption as Q1 – Q2. From a total of 10898 respondent households, this quantity is available for all but 220 households, a very low rate of item non-response (2 percent).

Although comparison of recall and diary data within the FoodEx is the main focus of our analysis, we can also compare the FoodEx data to data from a second large Canadian survey. The 1996 Family Expenditure Survey (FamEx) is a full household expenditure survey (collecting

---

[4] To investigate the determinants of retention, we estimated a simple Probit model of week 2 response on the sample of households that responded in the first week. Although the overall retention rate was very high, we did uncover some statistically significant correlates of retention. In particular, retention was increasing in income and higher in the province of Quebec. Full results are available from the authors.

information on all categories of expenditure).[5] Unlike most national expenditure surveys, the FamEx does not have a diary component. Instead, face-to-face interviews are conducted in the first quarter to collect income and expenditure information for the previous year (Thus the 1996 data were collected in January, February and March of 1997 but refer to the 1996 year calendar year). The FamEx is therefore an unusual kind of recall survey. Considerable effort is made to ensure the quality of the data.[6] Statistics Canada also undertakes various checks of the data and the data are generally thought to be of very good quality. There are 10085 respondent households in the 1996 FamEx.[7]

Because the FamEx collects annual data and the FoodEx survey is run 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. Summary Statistics comparing the two data sets are presented in Appendix Table A1. The only significant obstacle to the direct comparison of the data stems from differences in the household income information included in the files. The FamEx file includes only net household income while the FoodEx file includes only gross household income. However, the FamEx also includes gross personal income for head and spouse, and where we use income information in our

[5] The FamEx (and its subsequent replacement, the Survey of Household Spending) are the surveys that are used to determine the weights for the Consumer Price Index in Canada. They have also been extensively used for demand analysis.

[6] Respondent households are asked to consult bills and receipts and income is carefully reconciled with expenditures and savings. In some cases, multiple visits to a household are made.

[7] Statistics Canada reports that the response rate to the FamEx surveys is about 75%.

analysis we use the sum of these two items as our income variable in the FamEx.[8] This obviously is an imperfect match to the FoodEx income information when there are additional earners in the household. A second minor difference between the data sets concerns the top coding of numbers of different types of persons (children, young adults, adults, seniors) in the household. For the Foodex these are recorded as 0,1 or (2 or more). In the FamEx, the top-coding is at 3. In both data sets total household size is top-coded at 6.

In summary then, we have four distinct data items that capture the distribution of food consumption in Canada in 1996. These are:

i. The "food at home" expenditure category in the FamEx
ii. The recall food consumption measure we construct for the FoodEx (described above)
iii. Food expenditures recorded in the first week diary of the FoodEx
iv. Food expenditures recorded in the second week diary of the FoodEx

We have multiplied the second by 13 and the third and fourth by 52 so that all are annual measures.

Figure 1 displays the empirical cumulative distribution of these four measures, while Table 1 reports the mean, median and coefficient of variation for these four measures as well as for budget shares and income in the two surveys.[9] Several features are notable. First, the diary records are considerably lower than the recall responses of the same individuals (in the FoodEx) or a second sample drawn from the same population (the FamEx). Second, diary records are

---

[8] The FoodEx file does not contain personal income data.

[9] Empirical cumulative distributions for income and budget shares are presented in Appendix Tables A1 and A2.

considerably more variable. Third, there is a notable drop off, of on average 10 percent, between the first and second week of the diary.

The drop off between the first and second week of the diary seems to be evidence of "diary fatigue" or "diary exhaustion". Statistics Canada (1999) concludes that diary exhaustion was a significant factor affecting accuracy of the responses. They report that, in addition to the between week differences, within week responses tended to be significantly larger for the earlier days of either week. Such exhaustion effects in expenditure diaries have been known for a long time (McWhinney and Champion, 1974.) Recently, Stephens (2003) reports similar phenomena in the diary sample of the U.S. Consumer Expenditure Survey (CEX) (also a two-week back-to-back panel.)

Tables 2 and 3 and Figure 2 provide some supplemental analysis of diary fatigue in the FoodEx. Table 2 shows that week-on-week changes in recorded food expenditure are largely unrelated to observable household characteristics. The one exception is that households from the Atlantic Provinces exhibit (on average) less diary fatigue. Table 2 examines the week-on-week change in recorded outlay by expenditure category and by store type. The results suggest that records of small items (coffee and tea, non-alcoholic beverages, sugar), and especially purchases from convenience stores decline from week one to week two. Figure 2 illustrates that week-on-week changes in recorded expenditures are both positive and negative, are highly variable, and roughly symmetric around the (negative) mean.

Because diary records are usually thought to be quite accurate, the usual interpretation of the gap between the diary and recall measures might be that the latter suffer from significant over-reporting. However, the significant diary fatigue evident in the diary records, suggests the possibility that the diary records (and even the first week diary records) suffer from significant under-reporting. This is in fact the conclusion reached by Statistics Canada who routinely inflate the diary information in publicly released data by the factor necessary to match the recall information.[10] (We have undone this adjustment for the purposes of our analysis.)

Figure 3 displays histograms of the four food consumption measures (note that, in this figure only, amounts are weekly rather than annual). These suggest that both diary and recall data may suffer from their own particular problems. In particular, the diary data exhibit significant numbers of zeros (as much as 10% of the sample). Since it is implausible that this large a fraction of the sample is fasting, a natural interpretation is that the diary data suffer from purchase infrequency. There is a small literature on methods for dealing with purchase infrequency, including Keen (1986), Pudney (1988 and 1989) and Meghir and Robin (1992). Note that this problem is not entirely resolved by combining the two weeks of diary data: the combined data still exhibit a significant spike at zero. On the other hand, Figure 3 also suggests that the recall data suffer from considerable heaping and rounding (note the "spikes" in the empirical distribution at round figures such as $50 and $100). The consequences of such heaping and rounding, and methods for dealing with it, are discussed in Battistin et al. (2003) and in Heitjan and Rubin (1990). We now turn to a more detailed analysis of the differences between the recall and diary data.

III. Measurement Errors in Recall Food Expenditures

Let c^* be true food consumption and c be an imperfect measure of that quantity. Define $\varepsilon = c - c^*$ so that:

$$c = c^* + \varepsilon$$

[10] The factor that Statistics Canada inflates by is 15.8%.

In order to work with c, it is common to make assumptions about the characteristics of ε. Typical assumptions include those that characterize "classical" measurement error (Bound et al., 2002): that the errors are mean zero and independent of the true level of consumption and all other variables in the model. In our notation:

i. ε is mean zero: $E[\varepsilon]=0$,

ii. ε is mean independent of (or uncorrelated with) c^* : $E[\varepsilon \mid c^*]=E[e]$. Note that a testable implication of this assumption is that a regression of c on c^* should give a coefficient (on c^*) of 1.

iii. ε is mean independent of other variables, X : $E[\varepsilon \mid X]=E[\varepsilon]$.

iv. ε is independent of c^*. This of course implies that higher moments of ε are not related to $c^*: E[\varepsilon^k \mid c^*]=E[\varepsilon^k], k=2,3....$, starting with conditional homoskedasticity: $E[\varepsilon^2 \mid c^*]=E[\varepsilon^2]$.

Sometimes a distributional assumption is added, in particular, that the measurement error is normally distributed:

v. $\varepsilon \sim N(0,\sigma^2)$,

Finally, it is useful to have a measure of the relative size of ε. A common measure is the signal-to-noise ratio of c, which is calculated as $R^2/1-R^2$ from a regression of c on c^*.

If c^* is observable, these things are all amenable to empirical investigation. On first thought, the FoodEx would seem to offer such a possibility. In particular, diary records of food expenditure are thought to be very accurate (Battistin, 2004.) Thus, a natural approach is to take the diary information in the FoodEx as true consumption. However, the analysis of the previous section suggests that the diary measures are not prefect. Nevertheless, it is still very informative to compare the recall data to a superior measure. As Bound et al. (2002) note, most validation studies do not have a "perfect" or true measure to which to compare survey responses as even administrative records contain some errors. The question is how to best use the diary information. What we do is to construct, from the diary records, three alternative measures of "true" food consumption, $c*$:

(A) The first week diary,

(B) The average of 1st and 2nd week diaries.

(C) The linear projection of the recall measure onto the two diary measures.

Arguments can be made for each of these measures. **(A)** has the virtue that it minimizes the effects of diary exhaustion. On the other hand, it will be affected more by infrequency than **(B)**. To construct **(C)** we regress the recall measure on the diary week records and take the predicted values from this regression as true consumption (and hence the regression error is interpreted as measurement error in the recall measure). **(C)** is a weighted average of the first and second week of the diary (plus a constant), where the weights are chosen in a way that assumes the "best case" for the recall measure: note that this procedure imposes the assumptions that measurement error is mean zero and uncorrelated with the true value.

Table 4 presents summary statistics for the measurement error in recall food expenditures. Each column corresponds to one of the assumptions outlined above (**A**, **B** and **C**) regarding the true value. The first panel shows that the measurement errors have a positive mean if we take either the first week of the diaries or the average of the two weeks as $c*$ ($301 and $512 respectively.) In either case, the errors have negative skew (-0.71 and -0.14 respectively), and have much thicker tales than the normal distribution (with measures of kurtosis of 10.0 and 12.1 respectively, where the normal distribution would be 3). Our third procedure (C), which imposes a mean of zero on the measurement errors results in a distribution of measurement errors that is positively skewed, but again with thick tails. Kernel density estimates of all three distributions are presented in Figure 4.

The third and fourth panel of Table 4 present tests for mean independence and homoskedasticity of the error terms. These tests are implemented by regressing c on $c*$. If the measurement errors are mean independent (uncorrelated with $c*$), then the coefficient, β, on $c*$ should be 1. We present a t-test of this hypothesis. We then use a standard Breusch-Pagan test to test whether the second moment of the measurement errors is independent of $c*$ (that is, to test for heteroskedasticity in the measurement errors).

If we use the first week of the diary or the average of the two weeks as true food consumption, then the measurement errors in the recall measure of food consumption are strongly and negatively correlated with the true value. Mean independence is rejected with t-statistics of -55.8 and -32.2 respectively. Recall that true measure (C) assumes mean independence. By any measure of true food consumption, homoskedasticity is strongly rejected, with p-values for the Breusch-Pagan test less than 0.001. Thus even if we impose mean independence (as in (C)), we reject independence.

The Breusch-Pagan tests uses residuals from the regression of recall consumption on $c*$. Squares of those residuals are regressed on $c*$ and $c*$ squared. Regardless of the choice of $c*$, the linear term is always strongly negative and the quadratic term positive; in each case the estimated elasticity of the squared measurement error with respect to $c*$ is negative at the mean of $c*$. The nature of the heteroskedasticity seems to be that the measurement error variance falls with value of “true” consumption, but a decreasing rate.

In the next (5th) panel of Table 4 we present Kolmogorov-Smirnov tests of normality of the implied measurement errors. In all three cases, normality is strongly rejected, with p-values less than 0.001.

Finally, we calculate the signal-to-noise ratio for c under each of our assumptions about $c*$. These suggest that the measurement errors in c are very substantial. If we take the first week diary record to be $c*$, the signal-to-noise ratio in c is only 0.22. With either of the other two measure of $c*$ the signal to noise ratio in c rises to 0.36 (differing only beyond the fourth decimal place.) Equivalently, 70 to 80% of the cross-sectional variance in recall consumption is measurement error. This is a very large number, but it is not unprecedented. For example, on the basis of serial correlation in the errors in consumption growth equations, Runkle (1991; pg 86) concludes “that approximately 76 percent of that portion of the variance in the growth rate of consumption unexplained by family-specific real interest rates is the result of measurement error” (where consumption is food consumption as measured in the PSID).[11]

Table 5 presents the results of regressing the implied measurement errors on variables typically used in the modelling of consumption: income, and demographic variables. If we take either the first week diary measure (A) or the un-weighted average of the two weeks (B) as true consumption, then these income and demographic variables do not seem to be significant determinants of the implied measurement errors, except perhaps for the presence of youths in the household. The measurement errors implied by our third procedure (C) appear to be more

---

[11] Note though that first differencing usually removes signal, so that typically one would expect measurement error to be a smaller fraction of the variance in *levels*.

strongly related to variables such as income, household size and the presence of children and youths.[12]

Table 6 presents the results of regressing the squares of the implied measurement errors on the same set of variables, in order to further investigate heteroskedasticity in those errors. Again, the results seem sensitive to the measure of $c*$ used to construct the measurement errors. The variance of the measurement errors constructed by either procedure (B) or (C) is significantly related to household demographics.

To summarize, this analysis suggest that the measurement errors in food consumption are large, do not satisfy the "classical" measurement error assumptions, and are not normally distributed.

In the inter-temporal consumption literature it is common to work with the logarithm of expenditure and to model the measurement error as multiplicative rather than additive. In this case assumption i. is replaced by $E[\mathrm{e}^{\varepsilon}] = E[\frac{c}{c*}] = 1$ and e^{ε} is typically assumed to be log-normally distributed. Thus ε, which is now the difference between $\ln c$ and $\ln c*$ is normally distributed (but not with mean 0): $\varepsilon \sim N(\frac{-\sigma^2}{2}, \sigma^2)$. The assumption of independence of $c*$ (and hence $\ln c*$) is maintained.

[12] The analyses reported in Tables 4, 5 and 6 were repeated but with observations for which "true" food consumption was zero deleted from the sample. The results do not differ significantly differ from those reported in Tables 4, 5, and 6. In particular, it is *not* the case that the rejection of normality is driven by these zeros.

Accordingly, we repeated the analysis described above, but working in logarithms, rather than levels, of food consumption. The results are presented in Tables 7, 8 and 9 (which parallel the format of Tables 4, 5 and 6 respectively) and in Figure 5.

The results for logarithms are quite similar to those for levels. We find evidence of negative correlation between the measurement errors and true values, except where it is zero by construction. We also reject homoskedasicity, and normality of the errors. The signal-to-noise ratios are again quite low. The coefficients of the linear (in $\ln c^*$) terms of the Breusch-Pagan regressions are again strongly negative and their absolute value is larger by orders of magnitude than the positive coefficients on the quadratic terms. Thus the elasticity of the measurement error variance with respect to $\ln c^*$ is estimated to be negative at the mean of $\ln c^*$ (and, indeed, even at its 99^{th} percentile.) We find more evidence in logarithms (than in levels) that the mean of the measurement errors is systematically related to income and demographics, as reported in Table 8. Table 9 reveals considerable evidence that the measurement error variance is also related to household demographics. Correlations between household size and measurement errors in reported food consumption play a central role in the application we take up in the next section.

IV. A Demand Application

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) $-\frac{\beta}{\alpha_1}\times 100\%$. Economies of scale imply that the per capita income required to keep living standards constant should fall with household size. Empirically, α_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 β be negative. 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 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 β (and the Engel method) with an underlying model of household economies of scale. 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 (normal) private goods, such as food.[13] Thus, holding per capita income constant, the per capita quantity of food, and hence the budget share, should rise. Thus β (and $\frac{\beta}{w_f}$, 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 sometimes referred to as the "Deaton-Paxson puzzle."

---

[13] This assumes that there is limited substitution between food and the public good.

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 higher is the chance for systematic underreporting of food consumption. Gibson shows that such a negative correlation between the measurement error and household size imparts a negative bias on 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 to test the validity of this prediction. Households were randomly divided into two subsamples and one subsample was asked to keep a diary while the other was asked recall questions. His results suggest that while recall surveys underestimate the household size elasticities, estimates based on the diary do not exhibit the Deaton-Paxson puzzle.

Two features of our data seem to pose an immediate challenge to the generality of the Gibson result. First, in the FoodEx, the recall data on food expenditure on average *exceed* the diary measure (implying over- rather than under-reporting).[14] Second, the evidence on the correlation between household size and measurement error in recall food consumption is mixed. The sign and statistical significance of this correlation depend on exactly how we use the diary

---

[14] This could be because of the “telescoping” problems referred to in Footnote 2. Because of Gibson’s experimental design, his recall data is not subject to such problems. Other obvious potential differences include larger households in PNG, and differences in shopping behaviour between PNG and Canada.

information to construct "true" food consumption. If we simply use the sum of the two weekly diaries, the correlation is positive but not significant.

To further explore these issues, we estimate food share equations that are a quadratic extension of the familiar, Working-Leser form. In particular, we estimate:

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

Where w_f is the budget share of food at home,[15] $\ln pcy$ 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 10.

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 dairy data, -0.023 with the FoodEx recall data and -0.003 with the FamEx recall data (3rd row, 2nd panel, Table 10.) The first two estimates are different from zero at conventional levels of statistical significant, while the third is not. Although the estimates are of the same sign and similar magnitude, F-tests do indicate that the FamEx recall estimates are statistically different from both FoodEx estimates (2nd and 4th row, 3rd panel, Table 10.) The implied elasticities are presented in the last row of the 4^{rth} panel of

[15] We define the food at home budget share as expenditure on food at home divided by gross income. This is both somewhat unusual and not entirely satisfactory – the preferred and more common denominator being total outlay. But gross income is the measure of resources that we have in both surveys.

Table 10. 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.

Turning to income effects, we find that the three implied elasticities have the same sign and are of similar magnitude. The estimated income elasticity of food expenditure (evaluated at the means of the data) is 0.239 with the FoodEx dairy data, 0.175 with the FoodEx recall data and 0.225 with the FamEx recall data (1^{st} row, 4^{rth} panel, Table 10.) 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.

V. Other Implications

With respect to the estimation of inter-temporal preference parameters, we are limited by the fact that we have only cross-sectional data. We do note that the signal-to-noise ratio in recall food expenditure data is quite low, and that unless the measurement errors made by households are more persistent than true consumption, the signal-to-noise ratio will be worsened by differencing the data. We also note that the cross-sectional distribution of errors does not have the properties that authors such as Colera (1993) and Alan, Attanasio and Browning (2003) assume to hold for time series distribution of errors (for each household.) The assumptions made by those authors could hold if, for example, *all* of the mean dependence that we document results from time-invariant, household-specific components of the measurement error. While this seems unlikely, we cannot rule it out with the data at hand.[16]

[16] Data that would allow us to examine repeated measurement errors from the same households would obviously be extremely valuable. Perhaps a future combination of scanner and recall data will make this possible.

Turning to the literature on the evolution of consumption inequality, we are again limited by our single cross-section of data. Nevertheless, our analysis suggests potentially serious problems. The magnitudes under dispute in this literature are not large. For example, Krueger and Perri (2006) argue that the variance of the log of nondurable consumption rose by about 2.5% between 1986 and 2000 in the U.S., whereas Attanasio, Battistin and Ichimura (2005) report a number about twice as large for the same period. Both numbers are considerably smaller than the increase in the variance of log income over the same period, which Krueger and Perri report exceeded 12%. Krueger and Perri employ the interview sample of the Consumer Expenditure Survey (CEX), so that every component of nondurable consumption, including food, is measured by recall. Attanasio, Battistin and Ichimura optimally combine the interview and diary samples of the CEX and in particular, draw information on food expenditures from the diary sample.

Our results (Table 7) suggest that if diary records are accurate, then with recall food consumption data perhaps 70% of the cross-sectional variation in log food consumption is due to measurement error. In addition, at least in cross section, the variance of this measurement error is significantly declining in the level of true food expenditures, and also related to the demographic characteristics of respondent households. If these cross-sectional relationships hold over time as well, then it easy to show that movements in the level of consumption (or perhaps even in demographics) over the relevant period could generate spurious changes in the cross-sectional variance of log consumption of the magnitudes under debate.

VI. Conclusion

Measurement error is a ubiquitous feature of micro data, and a major challenge to empirical work. A first step in dealing with this challenge is to learn as much as possible about the characteristics of the measurement error in different kinds of data. In this paper, we have used an unusual Canadian survey to investigate the nature of measurement error in food consumption data.

Direct inspection of the measurement errors suggests that they are large, and that they do not have the properties of "classical" measurement error. In particular, the evidence suggests that the measurement errors are negatively correlated with the true values.

In an application drawn from demand analysis, we compared estimates of income and particularly 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 Gibson's resolution of the Deaton-Paxson puzzle.

The fact that measurement errors do not appear to be independent of true values has important implications for the literature on estimating inter-temporal preference parameters and for the literature on the evolution of consumption inequality. If the mean dependence we document holds for a given household over time, then it violates a key assumption underlying strategies recently suggested for estimating exact (nonlinear) consumption Euler equations in the presence of measurement error. Similarly, the heteroskedasticity we find would provide a mechanism by which growth in average consumption could lead to spurious changes in the cross-sectional variance of (mis-measured) consumption.

Finally, we note that our analysis has followed the literature in assuming that the diary information on food consumption is very accurate – much more so than the recall data. However, our preliminary analysis of the data (Section 2) documented evidence suggestive of several kinds of problems with the diary data (including infrequency and diary exhaustion.) If one is open to the possibility that the diary data contain substantial measurement error, or even that they measure expenditure well but over the period usually covered by diaries (one to two weeks) there can be substantial deviation between expenditure and consumption, then our results are subject to alternative interpretations. In that case, what we have studied is the sum (at the household level) of the measurement errors in the recall and diary data. Some of the measurement error properties we have documented might be attributable to the diary records. For example, significant purchase infrequency in the diary records would generate the (negative) mean dependence we observe. This suggests to us that the superiority of diary data may not be as obvious as the literature suggests. This is an issue that could bear further scrutiny.

References

Alan, S., O. Attanasio and M. Browning (2005), "Estimating Euler equations with Noisy Data: Two Exact GMM Estimators," working paper.

Alan, S., and M. Browning (2003), "Estimating Intertemporal Allocation Parameters Using Simulated Residual Estimation." CAM Working Paper 2003-03, University of Copenhagen.

Altonji, J., and A. Siow (1987), "Testing the Response of Consumption to Income Changes with Noisy Panel Data," *Quarterly Journal of Economics*, 102, 293-328.

Amemiya, T. (1985), "Instrumental Variables Estimator for the Nonlinear Error-in Variables Model," *Journal of Econometrics*, 28(3):273-289.

Attanasio, O. and H. Low, (2004), "Estimating Euler Equations," *Review of Economic Dynamics*, 7:405-435.

Attanasio, O., E. Battistin and H. Ichimura, (2005), "What Really Happened to Consumption Inequality in the U.S.?" in E. Berndt and C.Hulten, (Eds.) *Measurement Issues in*

Economics – Paths Ahead: Essays in Honour of Zvi Griliches (Chicago: University of Chicago Press).

Battistin, E. (2004). "Errors in Survey Reports of Consumption Expenditures." Institute for Fiscal Studies Working Paper 03/07

Battistin, E., R. Miniaci, and G. Weber, (2001). "What Do We Learn From Recall Consumption Data?" *Journal of Human Resources*, 38(2):354-385.

Bound, J., C. Brown and N. Mathiowetz, (2001). "Measurement Error in Survey Data." In Heckman, J.J. and E. Leamer, (Eds.), *Handbook of Econometrics*, Volume 5, Elsevier Science.

Browning, M., T.F. Crossley and G. Weber, (2003). "Asking Consumption Questions in General Purpose Surveys." *Economic Journal*, 113:F540-F567.

Carroll, C. , (2001). "Death to the Log-Linearized Consumption Euler Equation (and Very Poor Health to the Second-Order Approximation)," *Advances in Macroeconomics* 1(1): Article 6.

Colera, E.V., (1993). "A Note on Measurement Error and Euler Equations." *Economic Letters*, 45:305-308.

Deaton, A. and Paxson, C. (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.

Heitjan, D. F. and Rubin, D.B. (1990). 'Inference from Coarse Data Via Multiple Imputation With Application to Age Heaping.' *Journal of the American Statistical Association*, vol. 85, pp. 304-314.

Income Statistics Division, Statistics Canada, (1999). *1996 Food Expenditure Survey, Public-use Microdata Files.* Statistics Canada: Ottawa.

Keen, M. (1986), "Zero expenditures and the estimation of Engel curves", *Journal of Applied Econometrics*, 1, 277-86.

Krueger, D. and F. Perri, (2006), “Does Income Inequality Lead to Consumption Inequality? Evidence and Theory,” *Review of Economic Studies*, 73, 163-193.

McWhinney, I. and H.E. Champion, (1974), “The Canadian Experience with Recall and Diary Methods in Consumer Expenditure Surveys.” *Annals of Economic and Social Measurement*, 3(2):411-435.

Meghir, C. and J.M. Robin, (1992), “Frequency of Purchase and the estimation of Demand Systems.” *Journal of Econometrics* 53:53-85

Pudney, Stephen (1988), "Estimating Engel curves: a generalisation of the P-Tobit model", *Finnish Economic Papers*, 1.

Pudney, Stephen (1989), *Modelling Individual Choice*, Basil Blackwell, Oxford.

Runkle, D.E., (1991), “Liquidity Constraints and The Permanent Income Hypothesis,” *Journal of Monetary Economics*, 27:73-98.

Stephens, M. (2003). ““3rd of the Month”: Do Social Security Recipients Smooth Consumption Between Checks?” A*merican Economic Review* 93(1): 406-422

TABLES

Table 1: Summary Statistics:

Annual Household Food Consumption, Income, and Budget Shares.

			FoodEx		
		FamEx	**Diary Week 1**	**Diary Week 2**	**Recall Measure**
Sample Size		10085	10876	10719	10678
Food at home Consumption	**Mean**	4336	3854	3432	4156
	Median	3900	3261	2839	3911
	Coefficient of variation	0.58	0.82	0.88	0.58
Food at home Budget Share	**Mean**	0.15	0.12	0.11	0.12
	Median	0.10	0.08	0.07	0.10
	Coefficient of variation	2.70	1.57	2.69	2.22
Income Before Taxes	**Mean**	45716	44016		
	Median	38500	37200		
	Coefficient of variation	0.73	0.75		

Notes:

1. The 1996 FOODEX contains 10898 observations (households). 22 did not submit a first week diary while 179 did not submit a second week diary. The attrition rate (from week 1 to week 2) was 1.6%. 220 households did not provide a recall food expenditure estimate.
2. Statistics are calculated using survey weights.

Table 2: Regression Analysis:

Week on Week Change in Food Consumption - Diary Measure

Dependent Variable: (Week 1 Diary –Week 2 Diary) x 52

	Coef.	(Standard error)
ln pcy	54.45	(59.39)
$(\ln pcy)^2$	-0.46	(0.58)
Log household size	-753.66	(687.31)
Presence of children (0-15)	137.26	(171.33)
Presence of youths (16-24)	-3.22	(126.61)
Presence of seniors (65+)	6.23	(102.91)
2^{nd} Earner in Household	-108.77	(125.83)
Constant	*-418.97	(43.60)
R-squared	0.001	

Notes:

1. Regressors are all measured as deviations from means.

Table 3: Ratio of Mean Week 2 Food Consumption over Mean Week 1 Food Consumption

(By Broad Food Categories and Store Types)

All food at home	0.91
By category:	
Meat	0.91
Fish and other marine products	0.94
Dairy products and eggs	0.91
Bakery and cereal products	0.91
Fruits and nuts	0.91
Vegetables	0.92
Condiments spices and vinegar	0.92
Sugar and sugar preparations	0.86
Coffee and tea	0.88
Fats and oils	0.92
Other food	0.93
Non alcoholic beverages	0.84
By Store Type:	
Food from specialty stores	0.83
Food from convenience stores	0.75
Food from supermarkets	0.93
Food from other stores	0.83

Table 4: Errors in Recall Food Consumption - Descriptive Statistics (1996 Can $ per year)

		A	B	C
Mean		301	512	0
Variance		9198159	6057782	4297449
Skewness		-0.71	-0.14	1.30
Kurtosis		9.97	12.07	9.50
Percentiles	**5%**	-4431	-3071	-2572
	10%	-2998	-2007	-2101
	25%	-1117	-720	-1360
	50%	367	428	-307
	75%	1913	1741	1024
	90%	3560	3223	2490
	95%	4797	4390	3696
Test of Mean Independence ($\beta = 1$)				
$\beta - 1$		-0.67	-0.52	$\beta = 1$ by construction
[t-stat]		[-53.9]	[-32.8]	
Test of Conditional Homoskedasticity				
B-P test, Chi2		194	558	566
df		2	2	2
Prob > Chi2		<0.01	<0.01	<0.01
K-S test for Normality, p-value		< 0.01	< 0.01	< 0.01
R^2		0.19	0.27	0.27
Signal to Noise Ratio		0.23	0.36	0.36

Notes:

1. **(A)** Assumes first week diary measures "true" food consumption. **(B)** assumes the average of 1st and 2nd week diaries measures "true" food consumption. **(C)** Assumes the linear projection of the recall measure onto the two diaries measures "true" consumption.
2. Signal to Noise Ratio is calculated as $R^2/1-R^2$ from a regression of the recall measure on the assumed "true" measure.
3. Linear Regression of the recall measure on the two diary week records yields:

$$\text{Recall} = 2391.6 + \underset{(0.012)}{0.239}\ \text{Week1} + \underset{(0.015)}{0.245}\ \text{Week2} + \text{error}$$

**Table 5: Errors in Recall Food Consumption –
Regression on Covariates
(1996 Can $ per year)**

	A		B		C	
	Coef	**(Std Err)**	**Coef**	**(Std Err)**	**Coef**	**(Std Err)**
ln pcy	1.64	(55.29)	-25.59	(40.63)	*139.41	(31.42)
(ln pcy)^2	< 0.01	(0.54)	0.24	(0.38)	*-0.82	(0.29)
Log household size	-181.58	(635.54)	195.25	(475.72)	*-900.01	(363.99)
Presence of children (0-15)	214.70	(160.08)	146.06	(120.68)	*-198.22	(89.99)
Presence of youths (16-24)	*373.79	(114.29)	*375.40	(92.38)	*181.72	(71.86)
Presence of seniors (65+)	-142.65	(97.89)	-145.76	(76.84)	-48.11	(60.69)
2nd Earner in Household	-91.88	(119.51)	-37.50	(94.61)	-51.16	(74.54)
Constant	*291.03	(40.01)	*500.51	(31.85)	-7.12	(24.79)

**Table 6: Squared Errors in Recall Food Consumption –
Regression on Covariates
(1996 Can $ per year)**

	A		B		C	
	Coef	**(Std Err)**	**Coef**	**(Std Err)**	**Coef**	**(Std Err)**
ln pcy	358423	(671381)	*794247	(283436)	*406701	(232427)
(ln pcy)^2	1707	(7184)	-4293	(2608)	-431	(2257)
Log household size	-1920796	(7401677)	*-7366240	(3419272)	*-6403234	(2654562)
Presence of children (0-15)	-2854017	(1851392)	*-2022081	(762104)	*-1206404	(462913)
Presence of youths (16-24)	-395027	(941823)	663679	(547176)	*942777	(368182)
Presence of seniors (65+)	-1141612	(858817)	-376535	(479492)	-150472	(351927)
2nd Earner in Household	-1154663	(1009227)	*-1573653	(637308)	*-941136	(507728)
Constant	*9245786	(348872)	*6121834	(208103)	*4160967	(151546)

Notes to Tables 5 and 6:

1. **A)** Assumes first week diary measures “true” food consumption. **(B)** assumes the average of 1st and 2nd week diaries measures “true” food consumption. **(C)** Assumes the linear projection of the recall measure onto the two diaries measures “true” consumption.
2. All explanatory variables have been mean differenced.
3. Errors based on annualised household consumption (1996 Canadian $).

Table 7: Errors in Recall Log Food Consumption – Descriptive Statistics (1996 Can $ per year)

		A	B	C
Mean		0.18	0.20	0
Variance		0.76	0.57	0.30
Skewness		0.88	1.09	-0.64
Kurtosis		5.75	7.82	4.91
Percentiles	**5%**	-1.02	-0.83	-0.96
	10%	-0.73	-0.57	-0.67
	25%	-0.33	-0.23	-0.29
	50%	0.07	0.12	0.05
	75%	0.58	0.52	0.34
	90%	1.23	1.04	0.60
	95%	1.80	1.46	0.80
Test of Mean Independence ($\beta = 1$)				
$\beta - 1$		-0.70	-0.60	$\beta = 1$
[t-stat]		[-64.0]	[-45.4]	by construction
Test of Conditional Homoskedasticity				
B-P test,				
Chi2		355	714	401
df		2	2	2
Prob > Chi2		<0.001	<0.001	<0.001
K-S test for normality, p-value		< 0.001	< 0.001	< 0.001
R^2		0.19	0.26	0.27
Signal to Noise Ratio		0.23	0.35	0.38

Notes:

1. **(A)** Assumes first week diary measures “true” food consumption. **(B)** assumes the average of 1st and 2nd week diaries measures “true” food consumption. **(C)** Assumes the linear projection of the recall measure onto the two diaries measures “true” consumption.
2. Signal to Noise Ratio is calculated as $R^2/1\text{-}R^2$ from a regression of the recall measure on the assumed “true” measure.

Table 8: Errors in Recall Log Food Consumption – Regression on Covariates (1996 Can $ per year)

	A		B		C	
	Coef	(Std Err)	Coef	(Std Err)	Coef	(Std Err)
ln pcy	*-0.03	(0.01)	*-0.03	(0.01)	*0.02	(0.01)
(ln pcy)^2	*<0.01	(<0.01)	*<0.01	(<0.01)	*<0.01	(<0.01)
Log household size	0.24	(0.16)	0.20	(0.14)	0.14	(0.10)
Presence of children (0-15)	*0.10	(0.04)	*0.08	(0.03)	-0.02	(0.02)
Presence of youths (16-24)	*0.13	(0.03)	*0.12	(0.03)	*0.03	(0.02)
Presence of seniors (65+)	*-0.10	(0.03)	*-0.09	(0.03)	*-0.04	(0.02)
2nd Earner in Household	-0.02	(0.03)	-0.03	(0.03)	0.01	(0.02)
Constant	*0.18	(0.01)	*0.20	(0.01)	-0.01	(0.01)

Table 9: Squared Errors in Recall Log Food Consumption – Regression on Covariates (1996 Can $ per year)

	A		B		C	
	Coef	(Std Err)	Coef	(Std Err)	Coef	(Std Err)
ln pcy	*-0.05	(0.03)	-0.02	(0.02)	<0.00	0.01
(ln pcy)^2	*<0.01	(<0.01)	<0.01	(<0.01)	<0.00	<0.00
Log household size	0.53	(0.33)	0.07	(0.28)	*-0.30	0.12
Presence of children (0-15)	*0.14	(0.07)	0.03	(0.07)	0.02	0.03
Presence of youths (16-24)	*0.24	(0.07)	*0.18	(0.06)	*0.08	0.02
Presence of seniors (65+)	*-0.14	(0.07)	*-0.13	(0.06)	*0.06	0.03
2nd Earner in Household	-0.09	(0.06)	*-0.16	(0.07)	-0.02	0.02
Constant	*0.80	(0.02)	*0.60	(0.02)	*0.29	0.01

Notes to Tables 8 and 9:

1. **(A)** Assumes first week diary measures “true” food consumption. **(B)** assumes the average of 1st and 2nd week diaries measures “true” food consumption. **(C)** Assumes the linear projection of the recall measure onto the two diaries measures “true” consumption.
2. All explanatory variables have been mean differenced.
3. Errors based on (log) annualised household consumption (1996 Canadian $).

Table 10: Food at Home Budget Share Regressions

	FoodEx Diary		FoodEx Recall		FamEx Recall	
Food Budget Share $\left(w_f\right)$	.106		.124		.125	
Coefficients (Standard Errors)						
ln pcy	-0.44	(0.037)	-0.71	(0.037)	-0.616	(0.031)
(ln pcy)^2	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)
F-Test for common coefficients (p-value)						
ln pcy, (ln pcy)^2 – 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)^2 – vs FoodEx Recall					2.27	(0.103)
ln household size – vs FoodEx Recall					27.85	(<0.001)
Elasticities						
ln pcy $\partial \ln pce_f / \partial \ln pcy = \left(\partial w_f / \partial \ln pcy \bullet 1/w_f \right) + 1$	0.239		0.175		0.225	
ln household size $\partial \ln pce_f / \partial \ln n = \partial w_f / \partial \ln n \bullet 1/w_f$	-0.073		-0.183		-0.020	

Notes:

1. Regressions of the form $w_f = \alpha_0 + \alpha_1 \ln pcy + \alpha_2 \left(\ln pcy\right)^2 + \beta \ln hhsize + X\gamma + \varepsilon$
2. FoodEx Diary is average of 2 weeks
3. Additional control variables (X) include regional dummies, dummies for presence of children, youth and seniors, and presence of a 2nd earner in the household. Full estimation results available from the authors.
4. Survey weights are used in all estimation. White (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.)
5. Elasticities calculated at the means of the data.

Figure 1. Food Consumption, Empirical CDFs

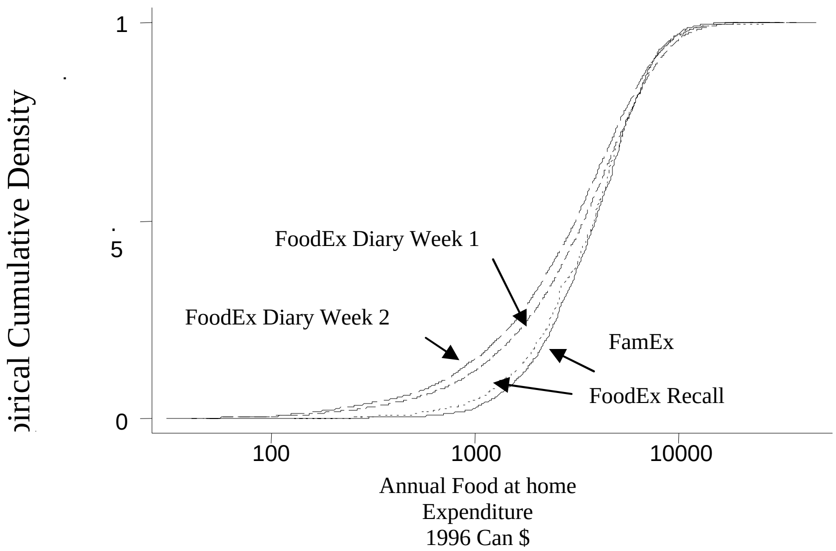


Figure 2: Changes in Reported Food Consumption Diary Week 1 to Week 2

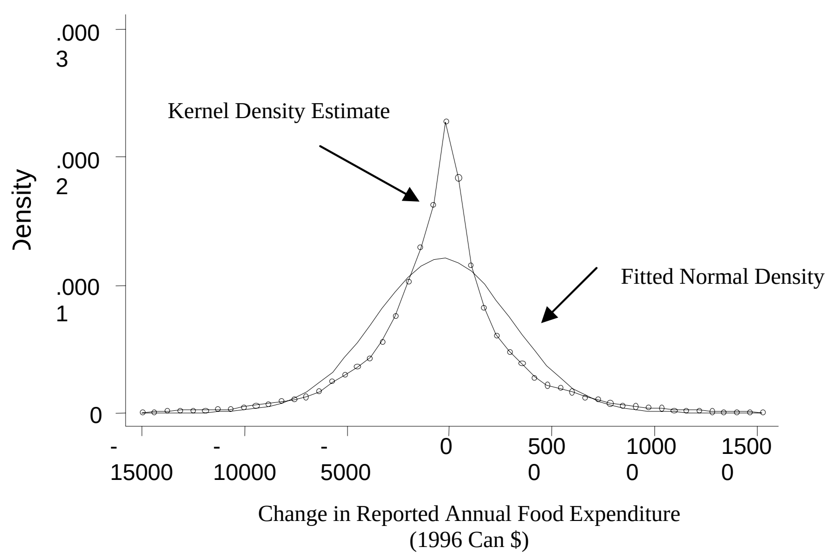


Figure 3: Food Consumption, Histograms

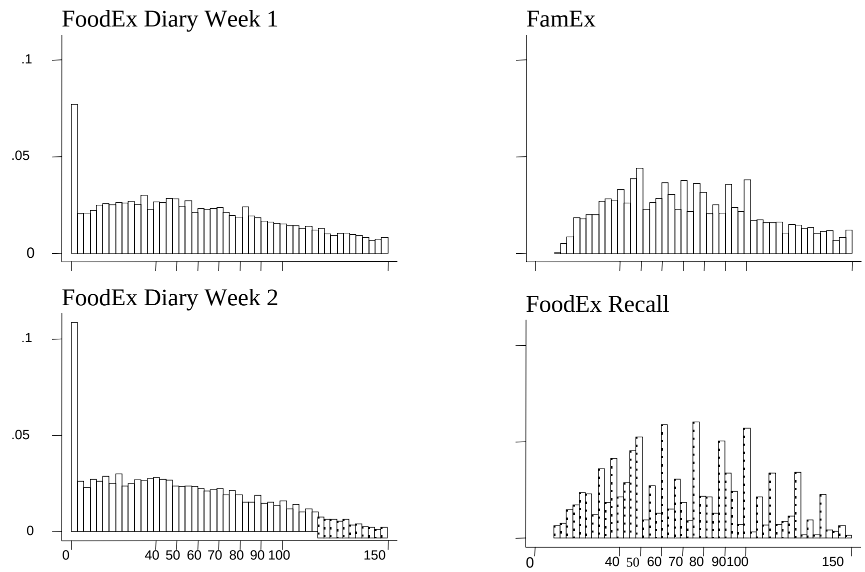
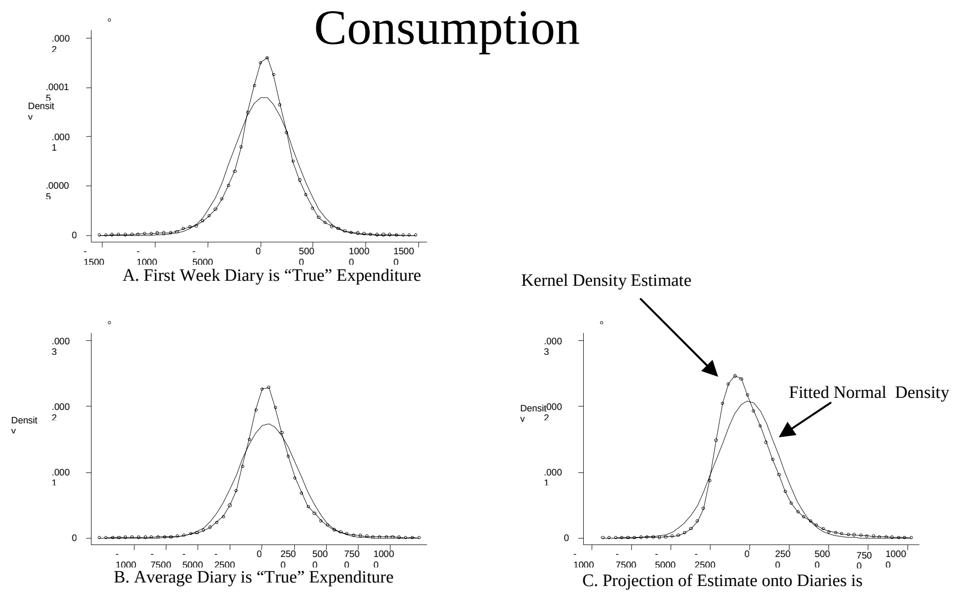


Figure 4. Errors in Recall Food Consumption

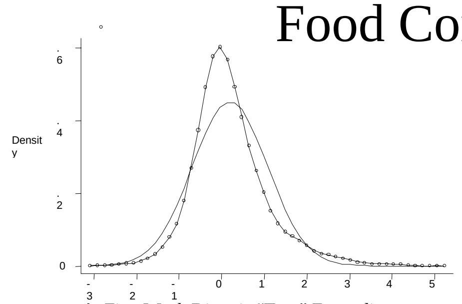


A. First Week Diary is "True" Expenditure

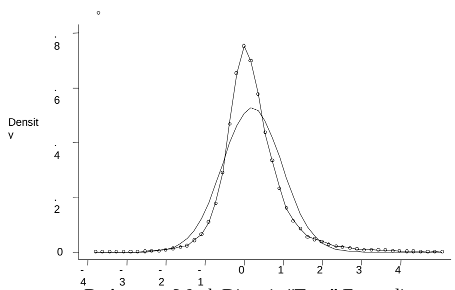
B. Average Diary is "True" Expenditure

C. Projection of Estimate onto Diaries is "True" Expenditure

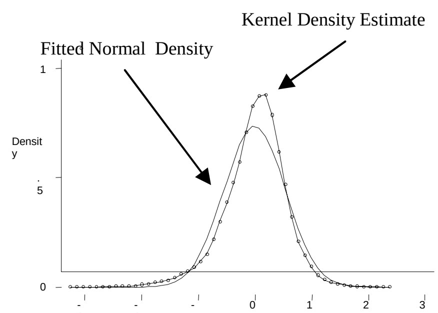
Figure 5. Errors in Recall Log Food Consumption



A. First Week Diary is “True” Expenditure



B. Average Week Diary is “True” Expenditure



C. Projection of Estimate onto Diaries is “True” Expenditure

APPENDIX: ADDITIONAL TABLES AND FIGURES

Table A1: Demographic Characteristics

		FamEx	**FoodEx**
Atlantic Provinces	%	7.7	7.7
Quebec	%	26.3	26.2
Ontario	%	36.9	36.8
Prairies	%	16.0	16.0
B.C.	%	13.0	13.3
Age	Mean	48.0	47.8
	Min	24.0	24.0
	Max	80.0	80.0
H-hold Size	Mean	2.62	2.6
	Min	1.0	1.0
	Max	6.0	6.0
Children (<15) Present	%	32.4	29.8
Youths (15-24) Present	%	23.4	24.9
Adults (25-64) Present	%	81.6	81.0
Seniors (65+) Present	%	23.0	22.5
2^{nd} Earner in Household	%	44.0	45.8

Figure A1. Household Income, Empirical CDFs

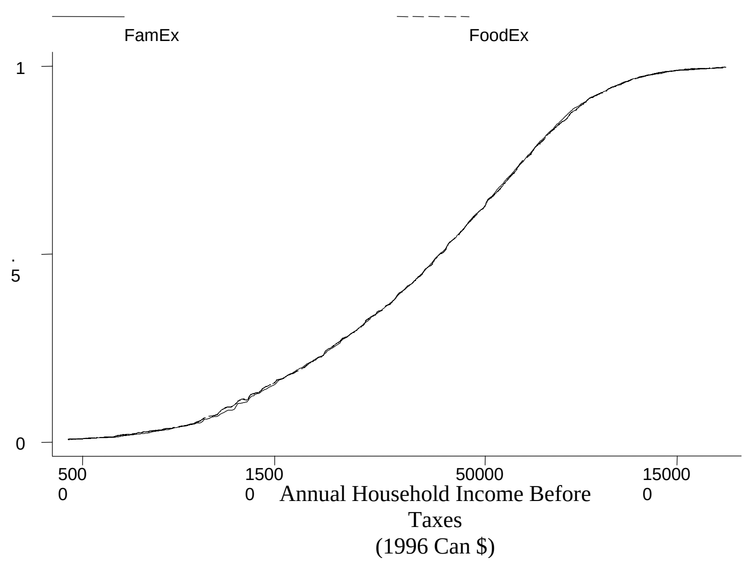


Figure A2. Food Budget Share, Empirical CDFs