MEASURING TOTAL HOUSEHOLD SPENDING IN A MONTHLY INTERNET SURVEY:
EVIDENCE FROM THE AMERICAN LIFE PANEL

Michael D. Hurd
Susann Rohwedder

Measuring Total Household Spending in a Monthly Internet Survey: Evidence from the American
Life Panel
Michael D. Hurd and Susann Rohwedder
NBER Working Paper No. 17974
April 2012
JEL No. C81,C83,D12

ABSTRACT

Beginning in May 2009 we fielded a monthly Internet survey designed to measure total household
spending as the aggregate of about 40 spending components. This paper reports on a number of outcomes
from 30 waves of data collection. These outcomes include sample attrition, indicators of data quality
such as item nonresponse and the variance in total spending, and substantive results such as the trajectory
of total spending and the trajectories of some components of spending. We conclude that high-frequency
surveying for total spending is feasible and that the resulting data show expected patterns of levels
and change.

Michael D. Hurd
RAND Corporation
1776 Main Street
Santa Monica, CA 90407
and NBER
mhurd@rand.org

Susann Rohwedder
RAND
1776 Main Street
P.O. Box 2138
Santa Monica, CA 90407
Susann_Rohwedder@rand.org

1. Introduction

A very extensive battery of questions about spending along the lines of the Consumer Expenditure Survey (CEX) is not feasible in a general purpose household survey because of space limitations. In the absence of panel measures of total household spending a large number of empirical papers have been based on the panel measure of food consumption in the Panel Study of Income Dynamics (PSID).[1] However, food consumption as a proxy for total consumption has limitations for some research questions: in the CEX the fraction of total consumption accounted for by food varies with income and with age, making it difficult to estimate life-cycle models based on food consumption.[2] An example of measuring consumption with just a few questions – but more than one – comes from the Survey of Health, Ageing and Retirement in Europe (SHARE). The first wave of SHARE included these measures of consumption: Food consumed at home, food consumed outside the home, telephoning, and total expenditure on non-durable goods and services.[3] While the data on food consumption was useful, the data on total nondurable consumption was deemed unreliable (Browning and Madsen, 2005).

As an approach that tries to strike a middle ground between the extensive detail collected in the CEX and just collecting information on a handful of categories, the Health and Retirement Study introduced the Consumption and Activities Mail Survey (CAMS). Collecting total household spending in a self-administered format (mail survey), the CAMS asks about spending in 36-38 categories and allows respondents to choose the length of the recall period (last month, last 12 months) for most categories. The self-administered nature of the survey has the advantage that respondents can take the time to think about their answers, even consult records if they are so inclined, without the social pressure arising from the presence of an interviewer. Hurd and Rohwedder (2009) show that obtaining useful spending data in a reduced number of spending categories is feasible. The spending totals aggregate quite closely to CEX totals and the age patterns of saving derived from the CAMS data (taking total income minus taxes minus spending) are quite close to the age patterns of savings implied by data on wealth change (Hurd and Rohwedder (2011).

---

[1] See, among others, Hall and Mischkin (1982), Zeldes, (1989), Altug and Miller (1990), Shea (1995).
[2] Bureau of Labor Statistics, 1978.
[3] The last item on this list included a number of cues, prompting respondents to include 'groceries, utilities, transportation, clothing, entertainment, out-of-pocket medical expenses and any other expenses the household may have and to exclude housing payments (rent or mortgage), housing maintenance, and the purchase of large items such as cars, televisions, jewelry and furniture.' (p. 318, Browning and Madsen, 2005).

However, whenever respondents are asked to recall their spending over a long period of time (say, one year), recall error becomes very important (Hurd and Rohwedder, 2009). This observation gave rise to the idea of attempting high-frequency elicitation of spending. Building on our experience with the CAMS data, we designed a spending module that we administered as part of the *Financial Crisis Surveys* that we were conducting in early 2009 in the RAND American Life Panel. The timing for a high-frequency elicitation of spending seemed particularly suitable in view of the high volatility in the economic environment which would make it likely that spending would be more volatile than usual as well. In fact, back in November 2008, just shortly after the large and sudden drops in the stock markets, about 75% of all households interviewed in the ALP reported reductions in spending in response to the economic crisis.

In this paper we describe in detail our survey methods, including an important innovation – the spending reconciliation screen – designed to catch large outliers that can be more frequent in self-administered surveys, for example, due to typos, and no interviewer to verify unusually large numbers. The reconciliation screen allows respondents to review all of their entries and the resulting total on one screen. Beyond the catching and self-correction of outliers the reconciliation screen also allows respondents to fine tune their entries, most likely reducing the noise in the data and leading to more accurate reports overall. Section 2 provides background on the American Life Panel, the Financial Crisis Surveys and the specifics of the design of the spending survey module. Section 3 reviews unit and item response rates, and various other indicators of data quality. Section 4 concludes.

2. Data and Survey Design

2.1 The American Life Panel

The American Life Panel (ALP) is an ongoing Internet panel survey operated and maintained by RAND Labor and Population. It covers the U.S. population age 18 and over. Those who do not have access to the Internet at the time of recruitment are provided with a Web TV (www.webtv.com/pc/), including an Internet access subscription with an e-mail account. Accordingly the sample does not

suffer from selection due to a lack of Internet access.[4] Post-stratification weights are provided so that after weighting, the ALP approximates the distributions of age, sex, ethnicity, education, and income in the Current Population Survey. About twice a month, respondents receive an email request to visit the ALP website to complete questionnaires that typically take no more than 30 minutes to finish. Respondents are paid an incentive of about $20 per 30 minutes of survey time, and pro-rated accordingly for shorter surveys. Response rates are typically between 75 and 85% of the enrolled panel members, depending on the topic, the time of year, and how long a survey is kept in the field.

Since inception of the American Life Panel in 2006, there have been four sample recruitment efforts.   In this paper we report on high-frequency data collections that were part of the so-called *Financial Crisis Surveys*, covering the period from November 2008 following the onset of the turmoil in the U.S. financial markets, up to the latest survey that was completed in October 2011. Back in 2008 the majority of active ALP panel members had been recruited from the University of Michigan Survey Research Center's Monthly Survey (MS). The MS incorporates the long-standing Survey of Consumer Attitudes and produces the Index of Consumer Expectations. The MS survey is considered to have good population representation (Curtin, Presser, and Singer, 2005).

2.2 The Financial Crisis Surveys

The very large stock market declines in October 2008 prompted our first financial crisis data collection. We designed a survey that was administered to the ALP in November 2008. The survey covered a broad range of topics, including various dimensions of life satisfaction, self-reported health measures and indicators of affect, labor force status, retirement expectations, recent actual job loss and chances of future job loss, housing, financial help (received and given and expectations about these), stock ownership and value (including recent losses); recent stock transactions (actual and expected over the next 6 months); expectations about future stock market returns (one year ahead, 10 years ahead); spending changes; credit card balances and changes in the amounts carried over; impact of the financial crisis on retirement savings; and expectations about future asset accumulation. We followed up with a second longitudinal interview in late February 2009 covering approximately the same topics.

---

[4] This approach has been used successfully in the Dutch CentER panel for many years.

In our first survey (November 2008) 73 percent of households reported they had reduced spending because of the economic crisis. These spending reductions are of substantial policy and scientific interest, and so there is considerable value in a careful measurement of the magnitude of the reductions. For example the welfare implications of the crisis depend partially on the reduction in consumption. Furthermore, because of the lack of knowledge of how spending responds to economic shocks at high frequency, it is important to establish the empirical connection between the triggering events and the magnitude of consumption reductions. The wide-spread spending reductions prompted us to re-orient the survey, expanding the collection of information on the components of spending.

Beginning with the May 2009 interview we established a monthly interview schedule to reduce the risk of recall error about spending and to collect data at high frequency on items such as employment, satisfaction, mood, affect and expectations. An objective was to permit detailed sequencing of events and their consequences.[5]

Each month we ask about spending in 25 categories during the previous month. Every third month beginning in July 2009 we ask about spending during the previous three months on an additional 11 categories plus seven big-ticket items. Spending in these categories tends to be less frequent such as durables. Taken together, the monthly and quarterly surveys measure total spending over a three-month period. The categories that are queried monthly amount to about 60% of total quarterly spending and the categories that are queried every three months account for the remaining 40%. This three-month schedule of two shorter monthly surveys and a longer quarterly survey has continued to the present.[6]

These surveys have several unique aspects. The first and most obvious is that they are monthly panel surveys. This design permits the observation of the immediate effects of changes in the economic environment that cannot be captured in low frequency surveys via retrospection. A second unique aspect is our measurement of a large fraction of total spending on a monthly basis. This measurement reduces recall bias for high frequency purchases, yet because the surveys cover an entire year, this measurement also captures low frequency purchases. The use of a reconciliation screen in the consumption module, described in detail below, reduces noise in the spending data

---

[5] To further reduce recall error the survey is only available to respondents for the first 10 days of each month, with only minor variation (e.g. adding the weekend if the 10th falls on a Friday or to accommodate staff work schedules when the beginning or end of the survey coincides with a major holiday such as New Year). Thus state variables such as unemployment refer to approximately the first 10 days of a month, not the entire month.

[6] Information about the surveys is given in Appendix Table 1, including survey length, fielding schedule and response rates.

substantially, allowing meaningful analyses even in a small sample. Furthermore, the combination of spending data with a very rich set of covariates, elicited at high-frequency, allows for a wide variety of analyses, with much more careful information on timing and sequencing of events to investigate determinants and the effects.

A total of 2,693 respondents participated in at least one of the 30 interviews from November 2008 through August 2011. The wave-to-wave retention rate has been consistently high throughout this entire period, averaging 91.0 percent without showing any decline over time. Respondents are invited to continue to participate in the surveys even if they miss one or more interviews resulting in a higher retention rate across multiple waves than would be implied by the wave-to-wave retention rate. For example, 73.0 percent (N=1,966) responded to at least ten of the first 14 interviews. Beginning with wave 15, facing budgetary constraints, we had to restrict the sample and decided to exclude the most sporadic respondents, dropping those who had responded to less than five of the first fourteen interviews, leaving us with 2,338 eligible respondents. Since then another 30 respondents either requested to be dropped from the monthly surveys or they died.

In this paper we use data from 30 surveys covering the period May 2009 through October 2011. Calculated over the eligible sample of 2,338 respondents retained since wave 15, we obtain a unit response rate for the spending module that averages 81.7 percent (N=1911) in cross-section over 30 waves. In the interest of maintaining an adequate sample size while at the same time basing results on an approximate panel sample, we present results for the sample of respondents who missed at most four of the 30 interviews, resulting in an average sample size of 1,440 respondents per wave, translating into an average unit response rate for this sample of 61.6 percent per wave. Restricting the sample to those who completed the spending module in all 30 waves yields a unit response rate of 35.5 percent (N=829).

2.3. Eliciting Total Household Spending

Each month we asked about spending in 25 categories that are purchased at high to middle frequency every month. Then, every three months we asked about the purchase over the past three months of 11 less frequently purchased categories, and about seven big-ticket items. With possibly a few minor exclusions the total of the three monthly surveys and the quarterly survey add to total spending over the quarter.

The 25 categories queried in the monthly surveys are shown in Appendix Table 2 grouped as they would have been displayed.[7] The grouping by broad types of spending or by frequency of spending is meant to facilitate placement of reported amounts in the proper category: Respondents are sometimes unsure about category placement and they are helped by seeing other possibly relevant categories. The grouping should reduce the risk of either omission or double counting. For example, the following categories were displayed at the same time because they are associated with household operations.

Mortgage: interest & principal
Rent
Electricity
Water
Heating fuel for the home
Telephone, cable, Internet
Car payments: interest and principal

A major innovation was the development of a "reconciliation" screen. Outliers are a problem in self-administered data collection such as Internet interviewing because there is no interviewer to question extreme values. Therefore, we designed a new strategy for the ALP to help reigning in outliers: following the queries about spending last month on the 25 items we presented the respondent with a summary table which listed the responses and added them to produce the implied monthly spending total. The respondent was invited to review and edit any items. This produced two very favorable results: most importantly, there was a sharp reduction in outliers which has a large impact on standard errors of the total that is constructed as the sum of these 25 spending categories. Also, respondents had the opportunity to improve the accuracy of their entries, including previously missing entries which should reduce the noise in the data further. We give more details on these outcomes in the next section. See Appendix Table 4 for a display of the reconciliation screen.

A natural comparison with our method is the "Cash-Flow Reconciliation" method (Fricker, Kopp and To, 2012) or the "balance edit" that has been used in the Canadian expenditure survey (Brzozowski and Crossley, 2011). These approaches have the same objective as the reconciliation screen of catching and correcting reporting errors. They compare household spending to the sum of household income and net cash flows into and out of assets, and they challenge respondent reports

---

[7] In November 2010 (wave 21) we added another monthly category ("other transportation expenses") in reaction to some respondents indicating difficulties allocating some of their expenses.

when there are large discrepancies between the two. One disadvantage of this approach is that just a fraction of respondents are challenged; yet, measurement error on income, spending or on both could result in incorrectly challenging some respondents while incorrectly not challenging others. To the extent that respondents are induced by a challenge to modify possibly correct reports, the procedure will worsen the data of those incorrectly challenged. The reconciliation screen approach on the other hand invites all respondents to review and possibly edit their answers and therefore avoids the risk of misclassification due to measurement error. It is also worth noting that a comparison of spending with income is not practical when the reference period is one month as is the case in the ALP Financial Crisis Surveys. This is because the concept of after-tax-income in a month is not well defined, and income is more likely to fluctuate across short time periods leading to potential mismatches in the timing of the receipt of income and expenditures.

3. Results

3.1 Indicators of data quality

In the first 10 days of May, 2009, about 2100 people responded to the survey, which was the initial survey about spending. This number was fairly constant until wave 15 (May, 2010) when it declined by about 300 persons. The main reason for the decline is that due to budgetary constraints some infrequent responders were dropped from the survey.[8] Since then the number of observations has stabilized hovering between 1,750 and 1,850 observations. In a typical wave about 25 people begin the survey but fail to complete it.

The measurement of spending, which is the focus of this paper, is embedded in a longer survey of the effects of the great recession. For the spending part of the survey only, the median time for completion of the 25 monthly items, including time spent on the reconciliation screen, was about 3.3 minutes.[9] The median time to complete the additional quarterly items (10 items plus 7 big-ticket items) was about 2.3 minutes.

---

[8] Many analyses will use the panel aspect of the survey to study change: Infrequent responders have less or no value in such analyses.
[9] Mean times are not meaningful in a self-administered survey such as an Internet survey because respondents may interrupt the survey without disconnecting.

There is a very substantial age gradient: those age 65 or older take a little more than twice as much time as those less than 40. As for the variation over time, completion times in the first several waves were greater by roughly one minute than the typical median, but fairly quickly reached a steady level.

Item non-response is generally very low in the ALP and that holds also for the spending items. Respondents have two opportunities to fill in initially missing answers: first, if someone leaves a spending category blank an additional screen will appear asking the respondent to fill in the missing item(s). Then, should the respondent ignore this prompt he or she can still provide the missing value on the reconciliation screen. The average rate of remaining item non-response across all waves and all 25 monthly categories is 0.4 percent. Examining the rate of item non-response averaged across the 25 monthly categories by wave shows that this rate is fairly stable over time, ranging between 0.2 and 0.5 percent in almost all waves.[10] There is some variation across categories of spending, but it is rather small. For example the category with the highest rate of item non-response is "heating fuel for the home" (0.6 percent) compared to the lowest rate of 0.3 percent for "telephone, cable, Internet." In the reconciliation screen previously missing items are replaced with the value $0. Respondents have the opportunity to correct this value on the reconciliation screen. About 13 percent of initially missing values are updated from $0 on the reconciliation screen to a positive value. In some cases it could be that the initial missing entry was due to the respondent not having that type of spending or that the respondent left the entry blank, because s/he already accounted for that type of spending elsewhere in the survey.[11] If that was the case – and one could argue that respondents affirmed this view by not updating the zeros on the reconciliation screen – then no further imputation for missing values is required. This is the approach we apply in this study when calculating total spending. Because the rates of item non-response are so low, any other decision how to deal with missing information would not affect any of the statistics we present in a material manner.

---

[10] The highest average rate of item nonresponse was recorded in August 2009 at 0.8 percent and the lowest in February 2011 at 0.1 percent.

[11] A conscientious respondent may feel hesitant entering a zero when s/he had that type of spending (i.e. it was not truly zero), but had already included it elsewhere.

The reconciliation screen invites the respondent to correct entries. In the initial wave that elicited spending (May 2009), on average 1.8 percent of the entries were corrected (modified or updated) by respondents (Figure 1). The rate of correction declined steadily until the survey collected in February 2010. Since then it has fluctuated between 0.6 and 0.4 percent. Although this may seem like a small rate of correction the effect on outliers can be substantial if the corrections are for entries that are extreme. In the initial wave in May 2009, 17.0 percent of respondents altered at least one entry on the reconciliation screen for the monthly items. Four waves later (September 2009), this group had declined to 7.6 percent and has stayed about that size since then. The average rate of updating is smaller for the quarterly items when calculated over the entire population which is mainly due to the fact that many more respondents have zero spending in those less frequent quarterly items and those zeros are not usually updated.[12] For the quarterly items the fraction of respondents updating any quarterly spending items has hovered around its average size of just under 4 percent all along.

The frequency and magnitude of outliers can be a problem in self-administered surveys because there is no interviewer to question extreme values. The reconciliation screen is meant to reduce this problem. A measure of the extent of the problem is the standard deviation of spending: while some fraction of the measured standard deviation reflects true variation in spending across individuals, some fraction is the result of measurement error and often it is the result of extreme outliers. We compare the standard deviation of the sum of all monthly spending items based on respondents' initial reports prior to the reconciliation screen to the same measure based on respondents' updated reports. Averaging over all 30 waves we find that the standard deviation before the reconciliation screen is 44% larger than after the reconciliation screen. However, even after respondents' own corrections from the reconciliation screen there remain a small number of large outliers that dominate both the standard deviation before and the standard deviation after the reconciliation screen in some of the waves. We therefore reviewed some of the largest remaining outliers. For 22 very large values out of a total of 57,322 respondent-wave observations we then edited the spending report that produced the outlier using respondent-specific information from

---

[12] Calculating the rate of updating conditional on positive entries (i.e. excluding the zeros) for the quarterly spending items gives an average updating rate of 0.6 percent of entries per category.

adjacent waves. [13] We recomputed the standard deviation of total monthly spending based on initial reports before the reconciliation screen and also for the reports after the reconciliation screen. Figure 2 shows the standard deviation before and after the reconciliation screen conditional on the 22 edits.[14] In the first two waves the reduction due to the reconciliation screen was very substantial: from an average of $17.2 thousand to $3.0 thousand. In subsequent waves the reductions varied between being very small in some waves and sizeable in others, depending on whether respondents self-corrected some very large outliers. The average standard deviation (averaged over 30 waves and calculated after the 22 outliers were edited) was 112 percent higher before the reconciliation screen compared to after. This reduction will have a substantial effect on the standard errors in the estimation of models of spending.

Because the corrections induced by the reconciliation screen tend to involve large outliers, the corrections will reduce mean values of spending. We present the effects of the corrections on the mean (and the median) in Figures 3, 4 and 5 for the monthly items, the quarterly items and the quarterly totals (monthly plus quarterly items). The statistics in these figures are unweighted and cross-sectional to bring out the effect of the reconciliation screen.[15] The main observation from these figures is that the updates from the reconciliation screen do not affect measures of spending at the median, but they result in lower population averages at the mean in those waves where large outliers are caught.

Comparison with the Consumer Expenditure Survey

The CEX has the most authoritative survey measure of spending at the household level. Yet the level of spending as measured in the CEX is questioned because it is substantially less than household spending as measured by the National Income and Product Accounts. Nonetheless we will compare our measure of spending with that from the CEX because the CEX is a household survey

---

[13] The edited cases were identified usually as one particularly large outlier in a wave where comparisons with that respondent's reports in prior and/or subsequent waves suggested that there was an entry mistake. In most cases there was a decimal error (e.g., $1,000 instead of $100). Both the initial report prior to the reconciliation screen and the updated report were edited and the observation was included in the calculations for Figure 2.

[14] The standard deviation includes the total mortgage payment including repayment of principal because that is what is displayed to the respondent on the reconciliation screen. In calculating actual spending we remove principal repayments based on a follow-up question that queries the respondent about the different components of the mortgage payment (amount for principal, interest, and other expenses).

[15] We have edited the 22 outliers as described above both in the initial reports (prior to reconciliation screen) and in the reports after the reconciliations screen.

and aims for a complete measure of household spending using similar methods to ALP (although in much greater detail). We choose the calendar year 2010 for this comparison as this is the first complete year of monthly data on household spending in the ALP. 2010 is also the latest calendar year for which published tables from the CEX are available. For ALP we calculate spending over a year by summing all 25 monthly spending items from the 12 monthly surveys and the quarterly reported spending items from the quarterly surveys referring to 2010. Average spending in 2010 as reported in the CEX was $42,736.[16] Average weighted spending in the ALP was quite close at $41,553, or 97.2% of CEX spending. The similarity of these levels shows that it is possible to capture approximately the same amount of spending as in the CEX but using many fewer categories of spending and, therefore, imposing substantially less respondent burden and cost.

3.2 Trends in spending

To examine trends in spending over time we apply seasonal adjustments and weight the statistics.[17] We define three samples: a 30-wave panel sample composed of those who responded to all 30 waves of the spending surveys (wave 3-32 of the financial crisis surveys); a 26-wave panel sample composed of those who responded to 26 or more waves; and a cross-section sample composed of those who responded in a particular wave. When calculating population statistics we exclude from each wave a small number of observations where the respondent left many spending categories blank.[18] The respective size of these three samples is approximately 830, 1440 and 1900. We focus the discussion of the quantitative findings on the 26-wave panel sample as it maintains close to complete panel consistency, while substantially reducing potential selection effects that a strict panel definition would entail.

Figure 6 shows mean spending on monthly items, seasonally adjusted and weighted for the three samples. With the exception of the first four waves all samples produce remarkably similar results

---

[16] We excluded from the CEX published total for 2010 outlays for "Personal Insurance and Pensions," because we consider contributions to Social Security and pensions part of savings rather than spending. We also exclude life and other personal insurance payments because, except for insurance company profit, they represent transfers from one household to another.

[17] We calculate our own seasonal adjustment factors.

[18] Specifically we exclude a respondent's observation on the total of the monthly spending categories if he or she did not answer 6 or more of the 25 monthly categories. For the quarterly items we exclude respondents if they did not answer 6 or more out of the 11 (non-big ticket) quarterly items. Because of the very low rates of item non-response, the resulting reductions in sample size are small: the cross-sectional sample is reduced by 10 observations per wave on average; the 26-wave panel and the 30-wave panel are reduced by about 20 responses per wave on average.

both with respect to levels of spending and with respect to trends. Spending on the monthly items reached a minimum of about \$2,170 (nominal) in the May 2010 survey (which measures what households spent in the month of April) and recovered after that, increasing by about 9 percent from its minimum. When adjusting for inflation, the recovery since the trough in spending on the monthly items amounted to about 4.5 percent real.

There is substantial month-to-month variation most likely due to relatively high standard errors of the mean associated with our sample size. The 30 wave sample, which has the smallest number of observations, has a standard error of 74; the 26 wave sample, which has more observations, has a standard error of 60 and the cross-section, which has the largest number of observations, has a standard error of 53.[19]

Figure 7 shows median spending on the monthly items. The level is about \$500 lower then the level of the mean but the trends and patterns over time of the two measures are similar.

These figures do not include spending on durables and other low-frequency items whose pattern and trend may differ from spending on high frequency items. Figure 8 shows average total quarterly spending which is the sum of the monthly spending items for the three consecutive months in each quarter and the quarterly spending items. The statistics are seasonally adjusted and weighted. All three samples show a decline during the first four quarters of our period of study until reaching a minimum in the second quarter of 2010; after that they all show spending increases, indicating a recovery of household spending. Quantitatively there are some differences among the three samples. The most restrictive sample of those answering in all 10 quarters shows a smaller decline, but it is also the one that is most likely to suffer from sample selection, so we focus our attention on the results from the 8-quarter panel which are very similar to those based on the cross-sections. Average total spending declined from an initial level of \$11.1 thousand in the second quarter of 2009 by about \$1,100 or 9.9 percent until it reached the minimum in the second quarter of 2010. Adjusting for inflation the decline amounted to 12.0 percent real.[20] Over the recovery period from II/2009 through

---

[19] The standard errors illustrate the tradeoff between sample size and panel consistency.

[20] We note a particularly large decline during the first quarter or our survey period. This time coincides with the stock market reaching its lowest reading during this recession and rising unemployment. This large decline in spending in the first quarter of our surveys is therefore not implausible. However, the CEX disregards respondents' first quarterly reports for concerns about data quality. Our survey design and methods are very different so CEX data quality concerns do not

the end of our sample period, spending increased by 10.6 percent in nominal terms (as shown in the graphs) which corresponds to 6.3 percent real. Comparing the level of spending in the last quarter (III/11: $11,069) of our period of study with initial spending (II/09: $11,114), we find that despite the recovery, spending in the third quarter of 2011 was still lower compared to the second quarter of 2009, both in nominal terms (0.4 percent less) and in real terms (6.4 percent less).

Figure 9 has median total quarterly spending, seasonally adjusted, and weighted for the three samples. Compared to the statistics based on the mean, the median shows very similar qualitative patterns, but at a lower level reflecting the skewed distribution of spending. The medians also show a minimum in about II-10, and the recovery of spending observed until the third quarter of 2011 still remained short of the initial spending we recorded, even in nominal terms.

4. Conclusions

Before the initiation of our financial crisis surveys it was not clear whether high frequency surveying about a repetitive yet complex topic such as spending was feasible. We were unsure whether respondents would respond to the survey by attriting or by providing meaningless data. Neither of these fears appears to have been warranted. Attrition was modest, and, in fact, the main reduction in sample size occurred when we dropped some 300 respondents from our sample pool. Measures of data quality such as item nonresponse, time to complete the spending module, occurrence of outliers (after reconciliation) have been very stable over time suggesting a steady degree of respondent cooperation over time.

The reconciliation screen has been an important innovation, especially in a self-administered survey where there is no interviewer prompting the respondent in case of unusual values. It allows the respondent to efficiently review his/her prior response and make appropriate changes. Consequently our level of data cleaning and outlier adjustment in this paper was minimal.[21] Furthermore, the reconciliation screen also leads to corrections that do not concern outliers, but that consist of respondents' revisions of their initial entries to improve the accuracy of their reports. For

necessarily apply also to the ALP spending data collection. To find out for sure we would have to run a large survey experiment with different parts of the sample starting to take the survey in different months.

[21] Some surveys engage in a fairly involved "data editing" process prior to data release. In the data we presented in this paper we did not perform any data edits beyond about 22 most obvious outliers recorded over 30 waves of data collection.

example a household may correct a spending level from $1000 to $100. This reduction of noise in the data could not be achieved by ex-post data editing. Especially in small samples like ours this effect could be important. When fitting models of spending change, these minor corrections could have a substantial effect on significance levels of estimated coefficients. A topic of future research will be to investigate the quantitative importance of such changes.

References

Browning, Martin and Thomas F. Crossley (2009). Shocks, Stocks and Socks: Smoothing Consumption Over a Temporary Income Loss, *Journal of the European Economic Association*, December 2009, 7(6):1169 1192.

Brzozowski, Matthew and Thomas F. Crossley (2011), "Measuring the well-being of the poor with income or consumption: A Canadian perspective," *Canadian Journal of Economics*, Vol. 44, No. 1, pp. 88-106.

Fricker, Scott, Brandon Kopp, Nhien To, 2011, "Exploring the Feasibility of Implementing a Cash-Flow Reconciliation Approach in the Consumer Expenditure Interview Survey," Bureau of Labor Statistics.

Gruber, J. (1997). "The consumption smoothing benefits of unemployment insurance." *The American Economic Review*: 192-205.

Hurd, Michael D. and Susann Rohwedder, 2011, "Wealth Dynamics and Active Saving at Older Ages," presented at CRIW meeting in Washington D.C., December 2011.

Hurd, Michael D. and Susann Rohwedder, 2009, "Methodological Innovations in Collecting Spending Data: The HRS Consumption and Activities Mail Survey," *Fiscal Studies*, 30(3/4), 435-459.

Stephens Jr, Mel (2004). "Job loss expectations, realizations, and household consumption behavior." *Review of Economics and Statistics*, 86(1): 253-269.

Appendix Table 1

Basic monthly surveys are shorter, eliciting just a subset of variables. Every three months we administer a long survey (shaded in grey). The size of the initial eligible sample was N=2,693. Starting with wave 15 (May10) we restricted the sample to those respondents who had participated in 5 or more of the 14 prior surveys (N=2,338), excluding the most sporadic of respondents. This decision was motivated by budgetary constraints. As a result response rates among the (somewhat reduced) set of eligible respondents is higher on average in wave 15 and beyond.

Appendix Table 1: Survey schedule, survey length and response rates

Wave	Survey	Time to complete survey		Field period	Completed Responses	Response rate
		Mean	Median		N	[%]
1	Nov08	19.1	16.7	Initially unrestricted	2,052	76.2
2	Feb09	24.6	21.4	02/24-03/16/09 = 21 days	2,119	78.7
3	May09	14.6	11.8	05/01-05/10/09 = 10 days	2,080	77.2
4	Jun09	14.5	11.7	06/01-06/10/09 = 10 days	2,109	78.3
5	Jul09	26.8	22.7	07/01-07/10/09 = 10 days	2,100	78.0
6	Aug09	12.5	9.9	08/03-08/12/09 = 10 days	2,074	77.0
7	Sep09	12.4	9.7	09/01-09/10/09 = 10 days	2,123	78.8
8	Oct09	27.9	23.5	10/01-10/11/09 = 11 days	2,016	74.9
9	Nov09	13.9	11.1	11/02-11/11/09 = 10 days	2,056	76.3
10	Dec09	14.1	11.3	12/01-12/10/09 = 10 days	2,059	76.5
11	Jan10	28.5	24.3	01/01-01/10/10 = 10 days	2,069	76.8
12	Feb10	15.1	11.7	02/01-02/10/10 = 10 days	2,075	77.1
13	Mar10	14.0	10.7	03/01-03/10/10 = 10 days	2,057	76.4
14	Apr10	27.4	22.9	04/01-04/11/10 = 11 days	2,019	75.0
15	May10	10.4	7.9	05/03-05/12/10 = 10 days	1,861	79.6
16	Jun10	10.3	7.8	06/01-06/10/10 = 10 days	1,924	82.3
17	Jul10	25.9	21.5	07/01-07/11/10 = 11 days	1,814	77.6
18	Aug10	12.1	9.2	08/02-08/11/10 = 10 days	1,750	74.9
19	Sep10	11.8	9.2	09/01-09/10/10 = 10 days	1,836	78.5
20	Oct10	27.4	22.6	10/01-10/10/10 = 10 days	1,797	76.9
21	Nov10	12.0	9.3	11/01-11/10/10 = 10 days	1,851	79.2
22	Dec10	12.3	9.3	12/01-12/12/10 = 12 days	1,874	80.2
23	Jan11	34.5	29.1	01/03-01/12/11 = 10 days	1,836	78.5
24	Feb11	13.8	10.5	02/01-02/10/11 = 10 days	1,840	78.7
25	Mar11	12.8	9.8	03/01-03/10/11 = 10 days	1,845	78.9
26	Apr11	34.2	29.4	04/01-04/10/11 = 10 days	1,774	75.9
27	May11	16.4	12.4	05/01-05/10/11 = 10 days	1,768	75.6
28	Jun11	15.4	12.1	06/01-06/12/11 = 12 days	1,753	75.0
29	Jul11	31.0	26.3	07/01-07/14/11 = 14 days	1,838	78.6
30	Aug11	15.2	11.8	08/01-08/14/11 = 14 days	1,832	78.4
31	Sep11	14.8	11.5	09/01-09/11/11 = 11 days	1,785	76.3
32	Oct11	31.6	26.4	10/01-10/10/11 = 10 days	1,777	76.0

Notes: Time to complete the survey is calculated for completed survey responses, excluding any interviews that took 2 hours or longer. These respondents presumably interrupted the interview and returned to it later. The response rates for each wave are calculated over the initial eligible sample for the two periods and do not adjust for the fact that some few respondents declined future participation at some point and therefore were no longer part of the eligible sample. For example in Oct11, a total of 2,212 respondents were still eligible and active. The adjusted response rate for that wave would be 80.3% (=1,777*100/2212).

Appendix Table 2: Items queried each month, grouped by actual screen display

Screen 1:

Mortgage: interest & principal
Rent
Electricity
Water
Heating fuel for the home
Telephone, cable, Internet
Car payments: interest and principal

Screen 2:

Food and beverages: food and drinks, including alcoholic, that you buy in grocery or other stores	
Dining and/or drinking out: items in restaurants, cafes, bars and diners, including take-out food	
Gasoline	
Other transportation expenses: parking, tolls, public transport, taxi and similar (please exclude spending on trips and vacations)	Category added in wave 21 (Nov10)

Screen 3:

Housekeeping supplies: cleaning and laundry products
Housekeeping, dry cleaning and laundry services: hiring costs for housekeeping or home cleaning, and amount spent at dry cleaners and laundries
Gardening and yard supplies: yard, lawn and garden products
Gardening and yard services: hiring costs including materials they provided

Screen 4:

Clothing and apparel: including footwear, outerwear, and products such as watches or jewelry
Personal care products and services: including hair care, shaving and skin products, amount spent at hair dresser, manicure, etc.
Prescription and nonprescription medications: out-of-pocket cost, not including what's covered by insurance
Health care services: out-of-pocket cost of hospital care, doctor services, lab tests, eye, dental, and nursing home care
Medical supplies: out-of-pocket cost, not including what's covered by insurance

Screen 5:

Entertainment: tickets to movies, sporting events, performing arts, etc
Sports: including gym, exercise equipment such as bicycles, skis, boats, etc.
Hobbies and leisure equipment: such as photography, stamps, reading materials, camping, etc.

Screen 6:

Personal services: including cost of care for elderly and/or children, after-school activities
Education: including tuition, room and board, books and supplies
Other child or pet-related spending, not yet reported: including toys, gear, | equipment and veterinarian

Appendix Table 3

Additional items were queried quarterly beginning in the July survey about spending over the previous three calendar months. They include seven big ticket items and 11 other less frequent spending categories.

Screen 1:

Big ticket items
- Automobile or truck
- Refrigerator
- Stove and/or oven
- Washing machine and/or dryer
- Dishwasher
- Television
- Computer

Follow-up questions on big ticket items queried amounts, and in the case of cars how the purchase was financed.

Screen 2:

Homeowner's or renter's insurance
Property taxes
Vehicle insurance
Vehicle maintenance: parts, repairs, etc.
Health insurance

Screen 3:

Trips and vacations
Home repair and maintenance materials
Home repair and maintenance services
Household furnishings and equipment: such as furniture, floor coverings, small appliances, miscellaneous household equipment
Contributions to religious, educational, charitable, or political organizations
Cash or gifts to family and friends outside the household

Appendix 4: Selected Screen Shots from ALP Spending Module

Sample screen shot from the monthly spending survey module

Mortgage, rent, utilities, car

Please, provide your best estimate of how much in total your household spent in the following categories. Please include spending by all members of your household, that is, by you or anyone living with you. Even if the amount your household spent last calendar month was unusual, please report that amount.

	Amount spent last month	No money spent on this last month
Mortgage	$ _____ .00	OR ☐
Rent	$ _____ .00	OR ☐
Electricity	$ _____ .00	OR ☐
Water	$ _____ .00	OR ☐
Heating fuel for the home	$ _____ .00	OR ☐
Telephone, cable, internet	$ _____ .00	OR ☐
Car payments: *interest & principle*	$ _____ .00	OR ☐

<<Back   Next>>

RAND
American Life
Panel

Screen shot of the reconciliation screen

Your household's spending total last month: $2,838

According to your entries your household's spending in **January** on the described categories was: **$2,838**. Below is a summary of your entries. If you would like to make any changes to your entries, you can change the amounts in the table below and then click the 'Update total' button in the lower right corner of the screen to recalculate your total. Once you are satisfied with your entries, please just click 'Next'.

As a reminder, there are some categories of spending that we **DID NOT** ask you about in this survey, but we will ask these in a later survey. (Click here for a list of spending categories that we will ask about in a later survey.)

| Category	Amount spent last month
Mortgage	$ .00
Rent	$ .00
Electricity	$ .00
Water	$ .00
Heating fuel for the home	$ .00
Telephone, cable, internet	$ .00
Car payments	$ .00
Food and beverages	$ .00
Dining and/or drinking out	$ .00
Gasoline	$ .00
Housekeeping supplies	$ .00
Housekeeping, dry cleaning and laundry services	$ .00
Gardening and yard supplies	$ .00
Gardening and yard services	$ .00
Clothing and apparel	$ .00
Personal care products and services	$ .00
Prescription and nonprescription medications	$ .00
Health care services	$ .00

This screen shot displays the top portion of the reconciliation screen. The dollar amounts stated at the top give the sum of the respondent's answers to the monthly spending questions. All dollar amount fields are filled with the respondent's previously provided entries. Any missing categories are filled with a zero. Using the usual scroll bar to the right of the screen the respondent can scroll through the entire list of categories and edit any entries. At the bottom is a field that displays the "Total", an update button to have the total (displayed at the top and bottom) recalculated and the usual "Back" and "Next" buttons.

Figure 1. Percent of monthly spending categories updated, unweighted.
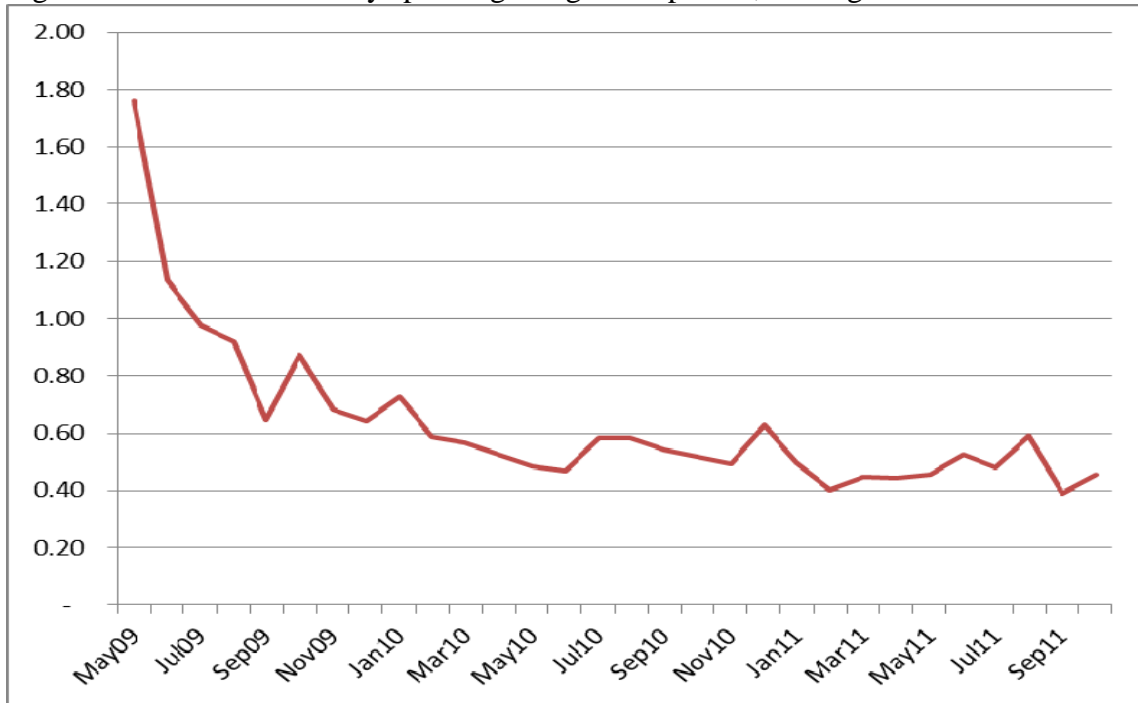
Average number of observations per wave =1,911.

Figure 2: Standard deviation of sum of 25 monthly items before and after the reconciliation screen, cross-section, unweighted.
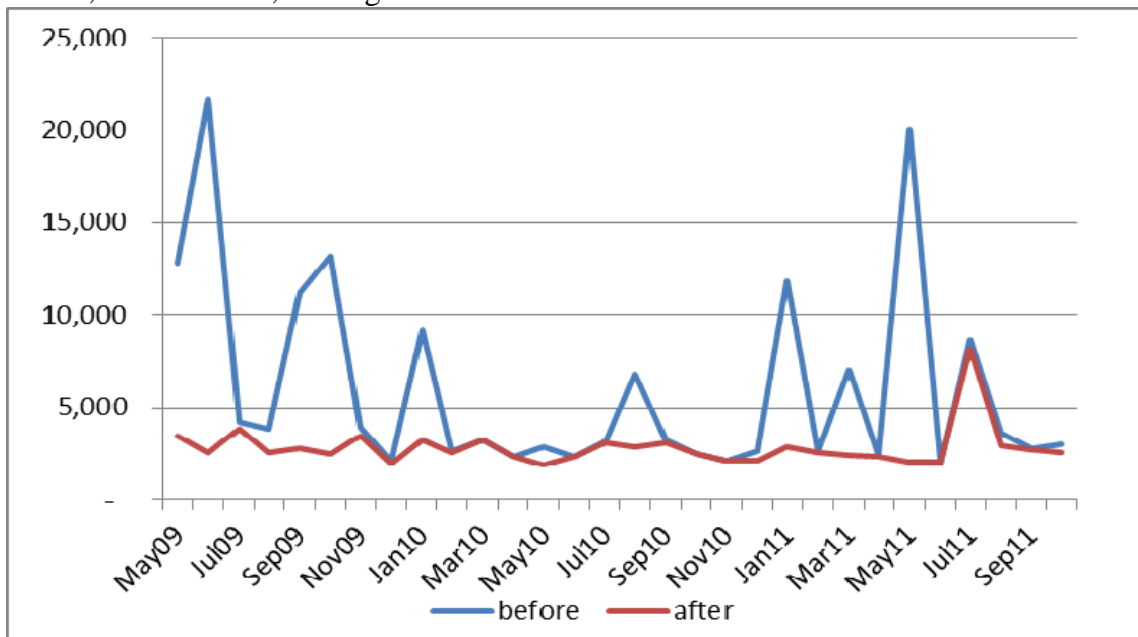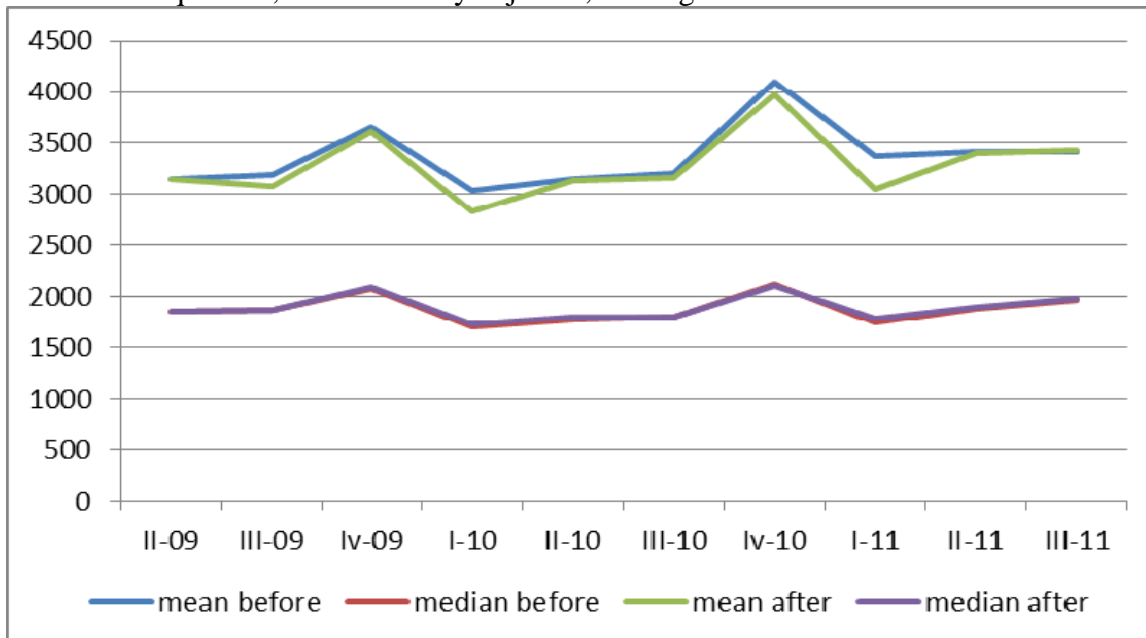

Note: Average number of observations per wave = 1,911. We edited 22 extremely large values out of 57,322 respondent-wave observations. Edits were subject to sufficient respondent-specific information being available from adjacent waves to support the determination that a value was indeed erroneous. Many consisted of shifted decimals (e.g., 10,000 instead of 100). There is a peak in July 2011 in the line depicting the standard deviation after the reconciliation screen. It is caused by two particularly large values that we decided *not to edit*. One pertains to a household that paid off the mortgage ($246k mortgage payment in July11 and no further mortgage payments in subsequent waves) and one pertains to a household that reports $200k in home repairs and $150k in home furnishings.

Figure 3: Mean and median spending on 25 categories before and after reconciliation screen, cross-section, not seasonally adjusted, unweighted
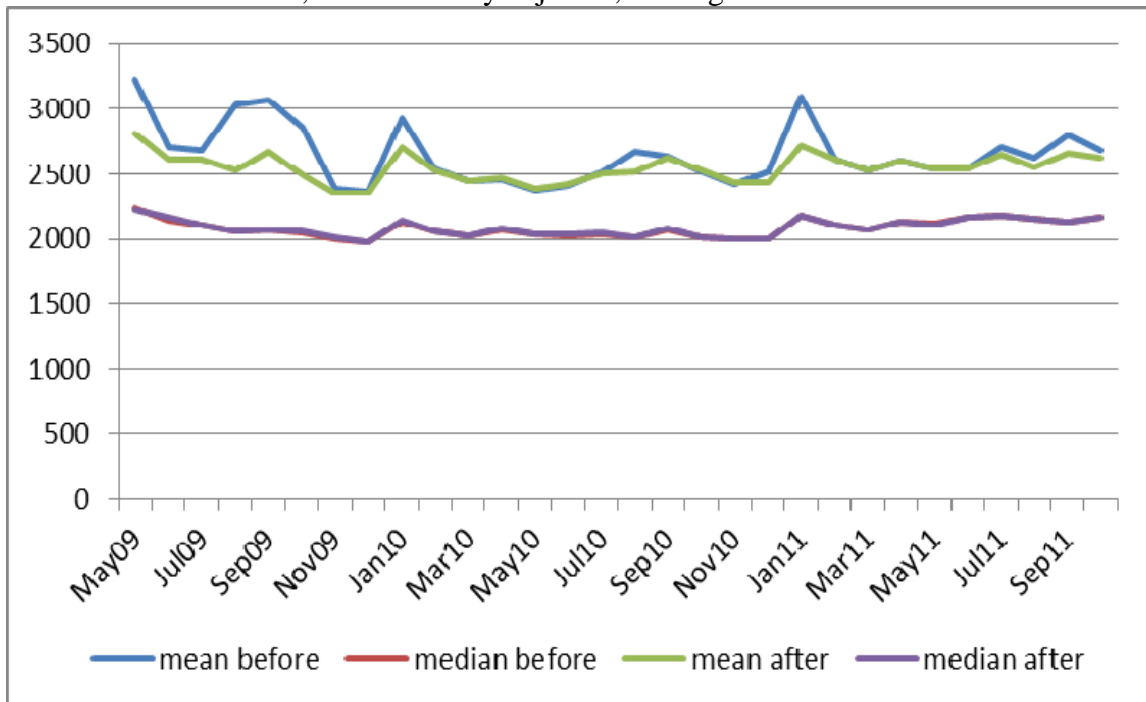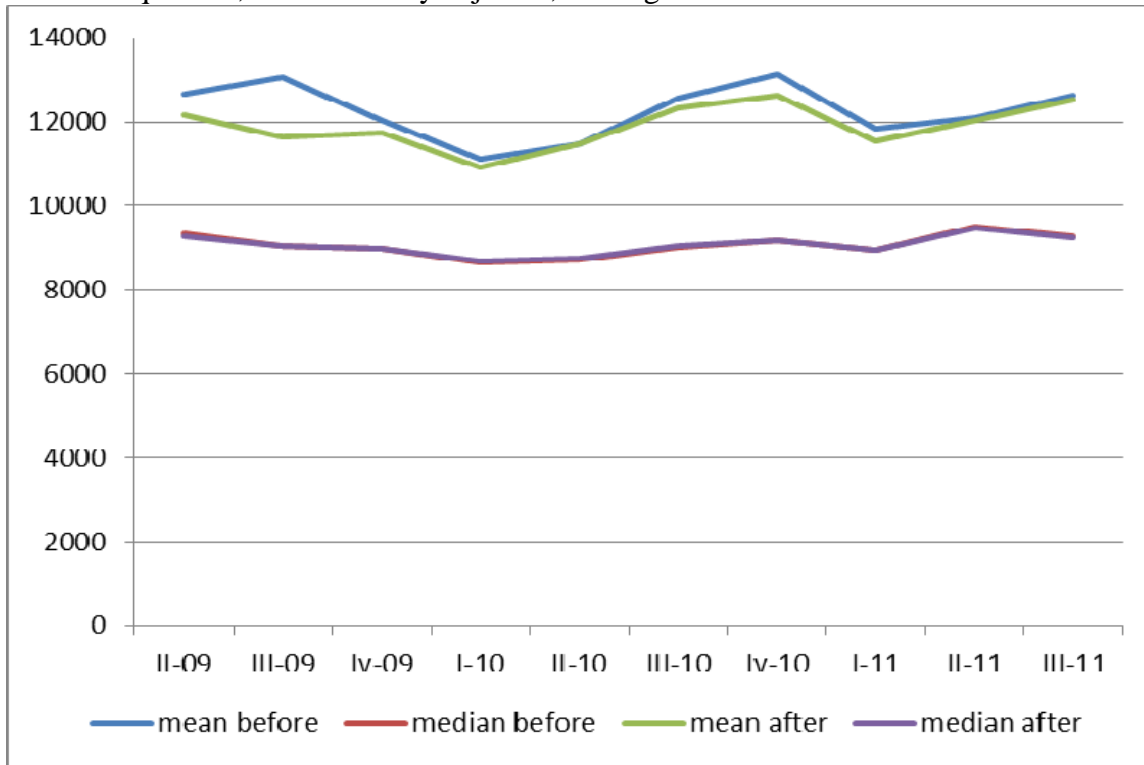

Average number of observations per wave = 1,907.

Figure 4. 17 quarterly spending items before and after reconciliation screen, cross-section of quarters, not seasonally adjusted, unweighted
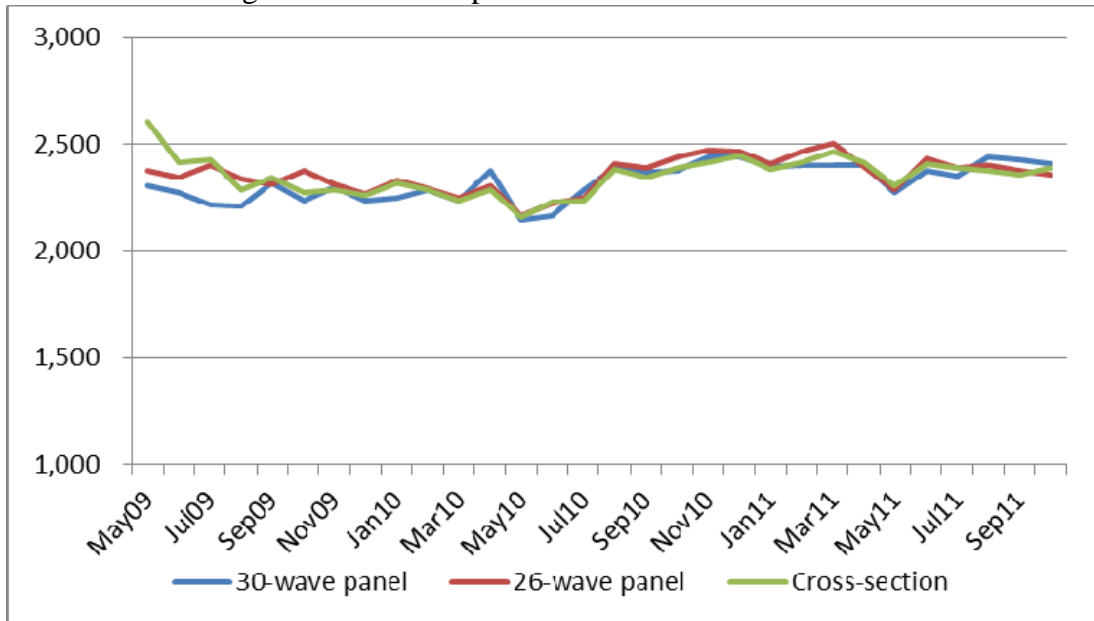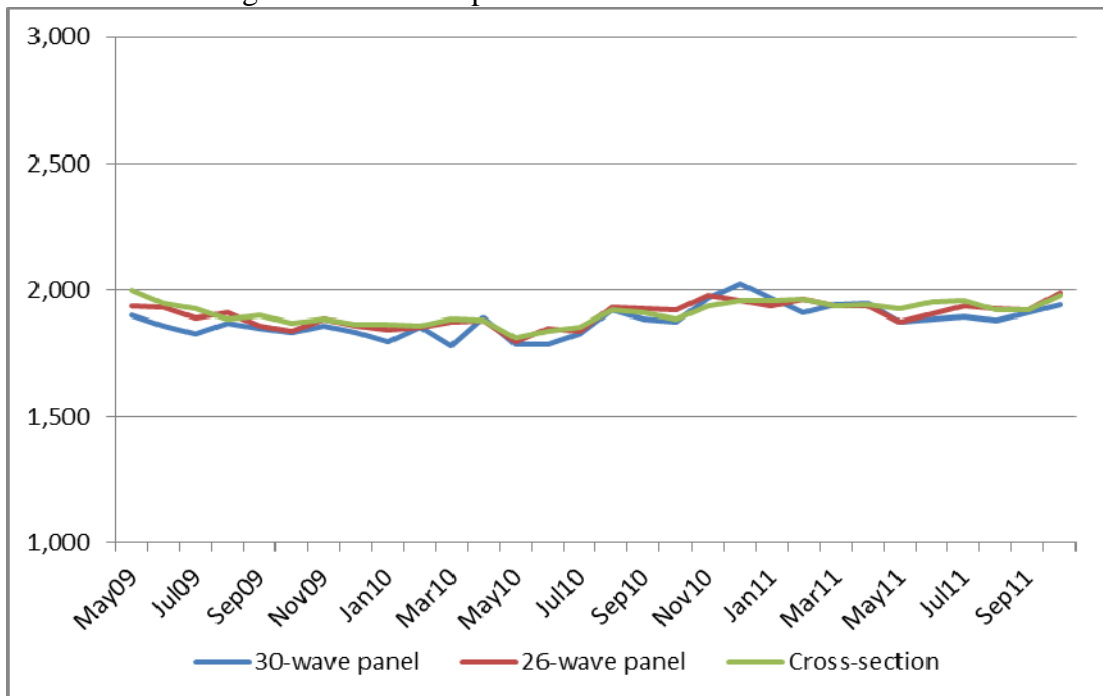

Note: Average number of observations per quarter = 1,621. Only respondents who participated in every survey in a quarter are included to allow calculation of statistics of before and after reconciliation screen. Respondents may miss a quarter.

Figure 5. Total quarterly spending before and after reconciliation screen, cross-section of quarters, not seasonally adjusted, unweighted.

Note: Average number of observations per quarter = 1,621. Only respondents who participated in every survey in a quarter are included to allow calculation of statistics of before and after reconciliation screen. Respondents may miss a quarter.

Figure 6. Mean spending on 25 monthly items (after correction), seasonally adjusted
and weighted. Three samples.

Average number of observations per wave = 1,901 in cross-section; 1,440 for 26-wave panel
and 829 for 30-wave panel.

Figure 7. Median spending on 25 monthly items (after correction), seasonally adjusted
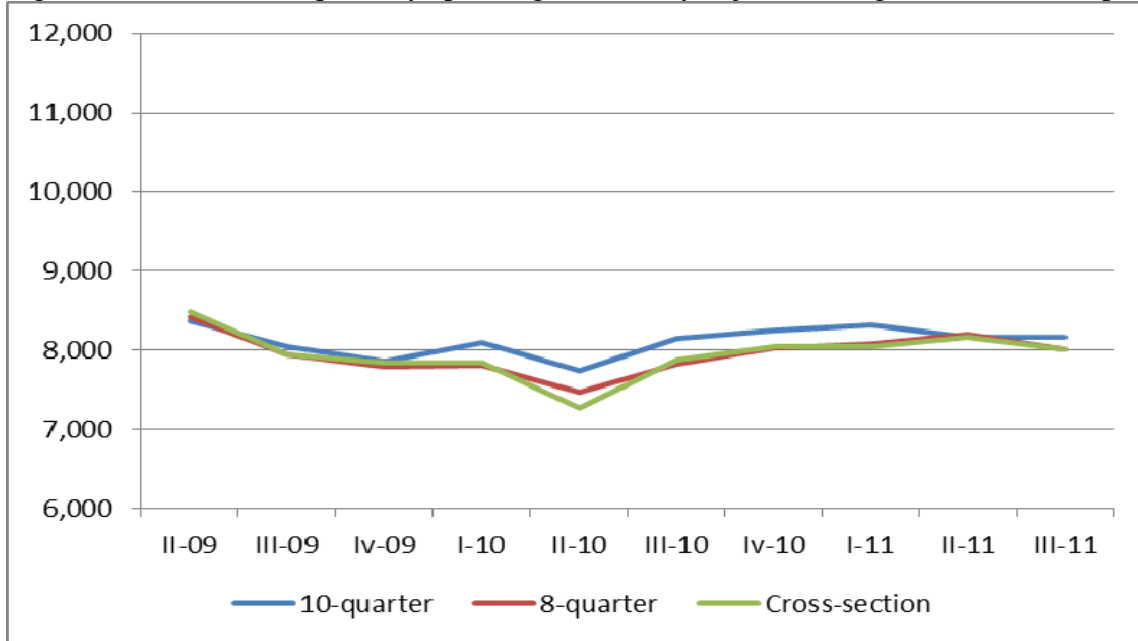and weighted. Three samples.

Average number of observations per wave = 1,901 in cross-section; 1,440 for 26-wave panel
and 829 for 30-wave panel.

Figure 8. Average total quarterly spending, seasonally adjusted, weighted. Three samples.
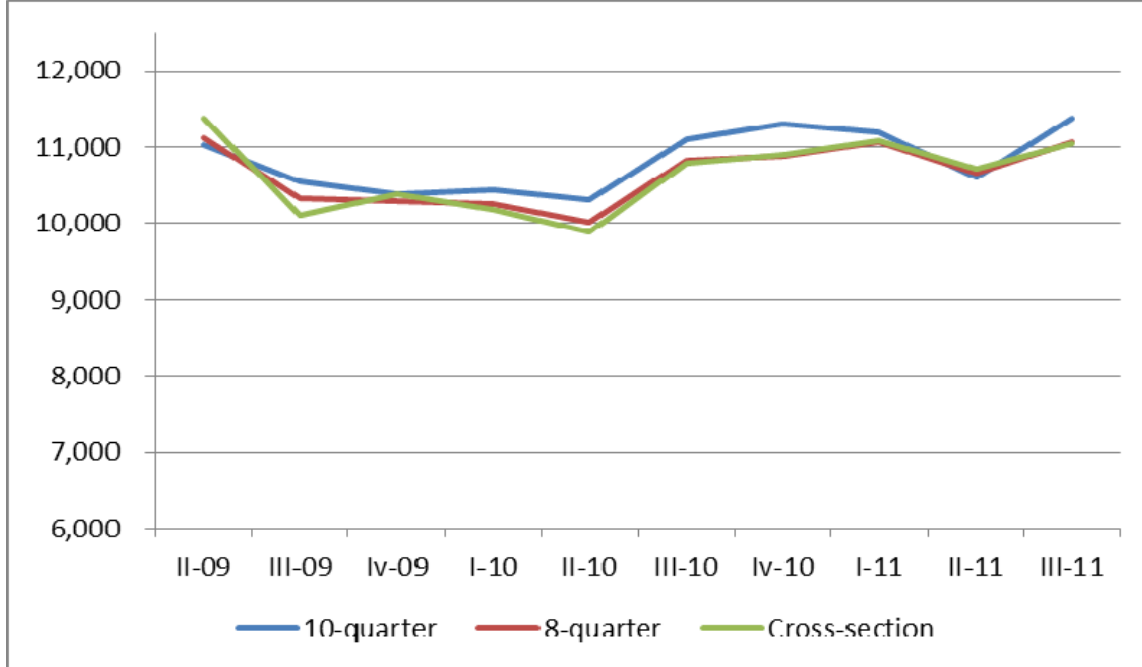
The 10-quarter panel has on average N= 1,290 observations per quarter. Respondents who completed at least 8 quarters or more yield a sample size that averages N=1,594 per quarter. Cross-sections of quarters average 1,812 observations per quarter.

Figure 9. Median total quarterly spending, seasonally adjusted, weighted. Three samples.

The 10-quarter panel has on average N= 1,290 observations per quarter. Respondents who completed at least 8 quarters or more yield a sample size that averages N=1,594 per quarter. Cross-sections of quarters average 1,812 observations per quarter.