Effect of Computer-Assisted Personal Interviews in the U.S. Consumer Expenditure Interview Survey December 2006

Moon J. Cho and Carolyn M. Pickering
U.S. Bureau of Labor Statistics
Moon J. Cho, 2 Massachusetts Avenue NE, Washington, DC 20212
(Cho.Moon@bls.gov)

Key Words: Mode effect, Data quality, Mean expenditure, Mean counts, Respondent characteristics.

1. Introduction

In the U.S. Consumer Expenditure (CE) Interview Survey, consumer units (CUs) are asked to provide month-by-month reports of expenditures. Since April 2003, the CE Interview Survey began utilizing the Computer Assisted Personal Interview (CAPI) instrument for data collection. The CAPI conversion is expected to improve data quality by shortening the data processing time, reducing the invalid responses, increasing accuracy by enforcing skip patterns, and by checking consistency and range as data are being entered. In this paper, the CAPI is compared to the conventional paper-and-pencil personal interview (PAPI) method by the mean expenditure, count, and expenditure per count with respect to expenditure categories, and respondent demographic variables. We present an overview of the CE Interview Survey and CAPI application. Finally, we describe the comparison study and present statistical tools for the study. The proposed methods are applied to selected subsets of items from the CE Interview Survey.

2. The Consumer Expenditure Survey

The CE Survey collects data relating to family expenditures for goods and services used in day-to-day living (BLS Handbook, 1997, p.160). One major use of the data is to provide the basis for revising weights and associated pricing samples for the Consumer Price Index (CPI). In addition, the Bureau of Labor Statistics (BLS) uses the data to produce estimates of mean expenditures and to produce public data sets of expenditures and income. The current survey consists of two separate surveys, the Interview Survey and the Diary Survey, each with a different data collection technique and sample. The principal reason for this use of multiple collection modes is that some expenditures (generally small or frequently purchased items) are believed to be more readily captured through a diary, while other items (generally purchases that are larger, less frequent, or otherwise more salient) are more readily captured through a periodic in-person interview (Eltinge et al., 2000). The remainder of this paper will consider only data from the CE Interview Survey. The purpose of the CE Interview Survey is to obtain detailed data on relatively large expenditure items such as property, automobiles, or major appliances, or on expenses which occur on a fairly regular basis, such as rent, utility bills, and insurance premiums. The CE Interview Survey includes rotating panels: each CU in the sample is interviewed every 3 months over five calendar quarters and then is dropped from the survey. Approximately 20 percent of the addresses are new to the Survey each quarter. The interviewer uses a structured questionnaire to collect both demographic and expenditure data in the Interview survey. See Cho et al.(2004) for more detailed information on the CE Interview Survey.

3. Section Description of CE Interview Survey

The following list contains general descriptions of expenditures collected in selected sections of the CE Interview Survey in 2002 and 2004. We compared all these sections at their aggregate levels between 2002 and 2004. Note that these descriptions do not list all items collected in the sections.

Section 4A - Telephone services including cellular service

Section 4B - Pre-paid phone card and pay phone expenses

Section 4C - Utilities and fuels such as electricity, natural or utility gas, piped-in water, cable/satellite TV, and internet access

Section 4D - Utilities and fuels for rented vacation properties
Section 5 - Property Construction, repairs to the home, and maintenance for the home
Section 6A - Major household appliances such as refrigerators and dishwashers
Section 6B - Small appliances such as toasters and other household items such as computers, electronics, and tools for the home
Section 7A - Maintenance or repair of household items such as electronics and appliances
Section 7B - Expense for service contracts for the maintenance or repair of household items
Section 8A - Expenses for furniture, infant's equipment, outdoor furniture and decorative items, decorative items for the home, storage and travel items, glassware and dinnerware, linens, and floor and window coverings
Section 8B - Rental, leasing, or repair of furniture
Section 9A - Clothing and accessories for persons age 2 and older
Section 9B - Clothing and accessories for persons under age 2 and watches, jewelry, and wigs or hair pieces
Section 9C - Repair of clothing, shoes, watches, or jewelry and clothing rental and clothing storage
Section 9D - Expenses for sewing materials and notions
Section 12A - Vehicle maintenance and repair, parts, and equipment
Section 12B - Licensing of drivers and vehicles and registration and inspection of vehicles
Section 12C - Other vehicle expenses such average gasoline expenditures and expenditures on parking, towing, and auto clubs
Section 15A - Out-of-pocket medical expenses
Section 16 - Educational expenses
Section 17A - Subscriptions to newspapers or magazines, season tickets to sporting events or theater, book clubs, CD and video clubs, reference books, encyclopedias, club memberships, credit card fees, and shopping club fees
Section 17B - One time or single admission expenses for the items in 17A except memberships plus photographic film and film developing
Section 19A - Miscellaneous expenditures
Section 19B - Contributions
Section 20A - Food and beverage expenses
Section 20B - Expenses for selected goods and services

4.   PAPI to CAPI Conversion

The CE Interview Survey introduced the CAPI instrument to the field in the data collection month April 2003. As of that date, field representatives (FRs) began collecting all cases in the CAPI instrument even if the FR previously interviewed the CU on paper. At the time of implementation, the CAPI interviewing experience levels of the FRs varied. Many FRs conducted CAPI interviews for other surveys using DOS-based software such as CASES. The Census Bureau programmed the CAPI instrument in Windows-based software called Blaise.

The questions changed very little from the last paper form in 2001 and the first CAPI instrument in 2003. However, the approach to asking the questions changed. On paper, the FR introduced the section, presented the reference period, listed each item collected in the section, and then collected detailed expenditure information on any items purchased by the CU. On paper, the FR marked 'no' next to any item not purchased. In CAPI, the FR introduces the section, presents the reference period, begins listing each item collected in the section, but instead of waiting until the respondent has heard the whole list, the FR immediately collects detailed expenditure information on any items purchased by the CU then continues asking about the remaining items. In CAPI, the FR makes no entries for items or services not purchased by the CU.

In addition, the authors note the following change in Section 4C: The focus shifted from the grouping the utility expenses by property in PAPI to focusing on the individual utilities in CAPI. The CE Interview Survey still collects the property associated with the utility, but the questionnaire no longer uses property as a recall tool for received utility bills.

5.   Previous Findings

The CAPI test team (2003) of the CE division conducted a preliminary comparison study between CAPI and PAPI data in terms of respondents, response rates, mean expenditure, and interview time. CAPI data were collected on approximately 3000 addresses per quarter in the first three quarters of 2002 in the study. CAPI data were then compared with PAPI data for the normal production during the same time period. The team found that demographic characteristics of respondents, expenditure levels, and overall length of interview time were quite similar between CAPI and PAPI. However, the team observed difference in the average number of reported expenditures per CU between two modes. The comparison study by The CAPI test team

provided important results. However, the CAPI instrument was not quite established during the testing time. Insurance expenditure counts and dollar amounts collected in the CAPI dress rehearsal, for example, were too low because the CAPI instrument did not carry forward some previous insurance policy information. In addition, interviewers who conducted CAPI were not as experienced as the ones with PAPI. The team compared PAPI and CAPI data that were partially processed data with the first three quarters, and with the first three interviews. We used the Phase 3 processed data (from the second interview through the fifth interview) from all four quarters.

See Bradburn et al (1991), Fuchs et al. (2000), Lamas et al. (1996) and references cited therein for more information on the CAPI data collection.

6.   Comparison Study

6.1   Comparison Data

We selected PAPI data from the 2nd through 5th interviews of the Year 2002 and CAPI data from the Year 2004. Both 2002 and 2004 share the same sample design. That means that the primary sampling units (PSUs) stayed the same and PSU groupings for the variance estimation also stayed the same. The data used for this analysis was generated from the monthly expenditures (MTABQ) files and the CU characteristics and income (FMLYQ) files of the Phase 3 databases. We didn't include Pension and Social Security expenditures which were stored in the income (ITAB) file. There are 22 sections in the CE Interview Survey. We chose twelve sections out of 22, and those twelve sections are the nearly exact translations of PAPI version to CAPI version. We avoided sections with significant changes. Numerically, we can conduct the same analyses for those sections. However, those sections provide less of a simple methodological conclusion when we observe the differences. In addition, they would provide less insight into prospective effects of further changes. Note that we considered only the components of the mean monthly expenditure of the CE Interview Survey that contribute to current CE production estimates. In other words, we used only the universal classification codes (UCCs) which were collected through the CE Interview Survey and used for publishing expenditures. Consequently, the "Overall Mean" entries are based on data from the UCCs for which publication is based on the interview reports. In addition, the entries for "Food" are based on food UCCs that are published from interview data; similarly for other subgroups included.

6.2   Notations

The remainder of this paper presents methodology and empirical results for expenditures, counts, expenditures per count, and other estimators observed in the CE Interview Survey. Define $\hat{\theta}_{cm}$ as the expenditure obtained from the CU c for month m and w_{cm} associated weight. Then a weighted estimator of the mean monthly expenditure is:

$$\bar{\hat{\theta}} = \frac{\sum\limits_{c}\sum\limits_{m} w_{cm}\hat{\theta}_{cm}}{\sum\limits_{c}\sum\limits_{m} w_{cm}}$$

where m is from January to December. The mean monthly count and the mean monthly expenditure per count are defined in a similar way. We made a CPI-adjustment for each selected section of 2004 expenditures in order to compare the 2004 expenditure with the 2002 expenditure accurately. The BLS provides CPIs for both overall and detailed expenditure categories on the CPI website. We noted that five out of twelve selected sections had the opposite direction from the overall CPI. Define $\hat{\theta}_{2004}^{CAPI}$ to be the original 2004 dollar values and $\hat{\theta}^{CAPI}$ to be the CPI-adjusted values from the CAPI method. The conversion formula from 2004 money to 2002 money is the following:

$$\hat{\theta}^{CAPI} = \hat{\theta}_{2004}^{CAPI} \times \frac{CPI_{2002}}{CPI_{2004}}$$

We examined the statistical significance of the following test statistic:

$$\frac{\hat{d}}{se(\hat{d})}$$

where

$\hat{d} = \left(\hat{\theta}^{CAPI} - \hat{\theta}^{PAPI}\right),$

$se(\hat{d}) = \sqrt{\hat{V}_{BRR}\left(\hat{d}\right)},$

$\hat{V}_{BRR}\left(\hat{d}\right) = (I_{13}, -I_{13})\,\hat{V}_{BRR}\,(I_{13}, -I_{13})',$

$\hat{V}_{BRR} = \frac{1}{44}\sum\limits_{i=1}^{44}\left(\left[\begin{array}{c}\hat{\theta}_{\alpha}^{CAPI}\\\hat{\theta}_{\alpha}^{PAPI}\end{array}\right] - \left[\begin{array}{c}\hat{\theta}^{CAPI}\\\hat{\theta}^{PAPI}\end{array}\right]\right)$

$\times\left(\left[\begin{array}{c}\hat{\theta}_{\alpha}^{CAPI}\\\hat{\theta}_{\alpha}^{PAPI}\end{array}\right] - \left[\begin{array}{c}\hat{\theta}^{CAPI}\\\hat{\theta}^{PAPI}\end{array}\right]\right)',$

$\hat{V}_{BRR}$:   the variance estimates by the balanced repeated replication method,
$\hat{\theta}_{\alpha}$: αth replicate estimator which is 13×1 vector,
I_{13}: 13×13 identity matrix.

Note that the variance formula takes into account the covariance structure between the PAPI and the CAPI versions.

Table 1: **Reporting Consumer Unit**

Section	PAPI	PAPI	CAPI	CAPI
Overall	31322	1.00	31077	1.00
Section 4	30748	0.98	30420	0.98
Section 5	5321	0.17	4455	0.14
Section 6	11266	0.36	10374	0.33
Section 7	2245	0.07	2612	0.08
Section 8	7516	0.24	7226	0.23
Section 9	15144	0.48	13898	0.45
Section 12	28067	0.90	28061	0.90
Section 15	20092	0.64	18432	0.59
Section 16	8372	0.27	7743	0.25
Section 17	23584	0.75	22066	0.71
Section 19	22790	0.73	21417	0.69
Section 20	27831	0.89	25890	0.83

Table 3: **Weighted Mean Monthly Counts**

Section	PAPI	CAPI	Diff	Test Stat
Overall*	19.70	18.49	-1.20	-6.30
Section 4	5.08	5.06	-0.01	-0.37
Section 5*	0.12	0.11	-0.01	-2.95
Section 6*	0.24	0.22	-0.02	-3.06
Section 7*	0.04	0.05	0.01	4.12
Section 8	0.15	0.14	0.00	-0.29
Section 9*	0.47	0.39	-0.07	-6.58
Section 12*	1.76	1.65	-0.11	-5.50
Section 15*	0.95	0.91	-0.04	-2.57
Section 16	0.27	0.27	0.01	0.93
Section 17*	2.71	2.31	-0.40	-9.96
Section 19*	1.44	1.25	-0.18	-5.90
Section 20*	1.82	1.59	-0.23	-12.39

7. Numerical Result

7.1 Comparing Point Estimators

In the CE Interview Survey, each selected CU is asked to participate in a total of five interviews. The first interview collects data for only bounding purposes, and was not considered in the paper. The second through fifth interviews are conducted at three-month intervals. In each of these interviews, the CU is asked to report expenditures for the past three months. The Table 1 presents sections, total number and proportion of reporting CUs of PAPI, total number and proportion of reporting CUs of CAPI. We observed that the CAPI had smaller proportion of reporting CUs in ten sections out of twelve.

We compared the mean monthly expenditure per CU of overall and selected sections between PAPI and CAPI. The Table 2 presents sections, mean monthly expenditure estimates of the PAPI and the CAPI data, difference between mean monthly expenditure estimates, standard error estimates of the PAPI data and the CAPI data, standard error estimates of the difference, and test statistics. The cut-off point using the student t test at $\alpha = 0.05$ is 2.021. We observed that CAPI data had smaller mean expenditure in seven sections out of twelve. Among those seven sections, three sections were statistically significant. Note that the sections with statistically significant differences are marked with the asterisks.

In addition, we observed that CAPI data had larger standard error estimates in eight sections out of twelve. Standard error estimate of overall expenditure in CAPI is also larger than the one in PAPI.

We compared the mean monthly counts per CU of overall and selected sections between PAPI and CAPI. The Table 3 presents sections, mean monthly count estimates of the PAPI and the CAPI data, difference between mean monthly count estimates, standard error estimates of difference, and test statistics. We observed that CAPI data had smaller counts in ten sections out of twelve. Among those ten sections, nine sections were statistically significant: overall counts in CAPI also decreased, and it was statistically significant.

We also studied the unweighted expenditure and counts. The unweighted expenditure and counts are slightly larger than the weighted ones in each section of both the PAPI and the CAPI. The unweighted expenditure and counts followed the similar patterns as the weighted ones. That means in cases where weighted estimator showed a decrease in the CAPI, unweighted ones also showed a decrease in CAPI. In addition, we looked into non-zero expenditures. For non-zero expenditures, mean expenditures of sections such as 'Education Expense' were increased a great deal while sections such as 'Utilities' stayed the same. That is because many CUs did not have 'Education Expense' expenditures and therefore did not report it, while most CUs had 'Utilities' expenditures and it is hard not to report to interviewers.

We compared the expenditure per count per CU of overall and selected sections between PAPI and CAPI. The Table 4 presents sections, expenditure per count estimates of the PAPI and the CAPI data, difference between expenditure per count estimates, and test statistics. We observed that CAPI data had larger expenditure per count in ten sections out of twelve. Among those ten sections, eight sections were statistically significant. Overall expenditure per count was also increased and it was statistically

Table 2: **Weighted Mean Monthly Expenditure**

Section	PAPI	CAPI	Diff	se02	se04	se(Diff)	Test Stat
Overall	2359.89	2329.63	-30.26	22.98	28.42	29.55	-1.02
Section 4*	260.28	256.82	-3.46	1.75	1.92	1.44	-2.40
Section 5	52.19	55.24	3.05	2.11	2.76	3.12	0.98
Section 6	53.24	56.28	3.04	1.23	1.56	1.90	1.60
Section 7*	3.37	4.53	1.16	0.16	0.21	0.33	3.52
Section 8*	43.14	47.78	4.64	1.79	1.57	1.91	2.43
Section 9	33.66	32.62	-1.03	0.94	1.19	1.33	-0.78
Section 12*	158.31	173.80	15.48	1.37	1.68	2.25	6.89
Section 15	87.38	83.67	-3.71	2.03	2.13	2.33	-1.59
Section 16	92.03	86.62	-5.42	4.94	4.68	3.86	-1.40
Section 17*	50.56	45.85	-4.72	0.92	0.90	1.10	-4.29
Section 19	191.92	191.12	-0.80	5.90	8.75	9.67	-0.08
Section 20*	70.63	63.11	-7.52	1.22	0.88	1.01	-7.45

Table 4: **Expenditure per Count**

Section	PAPI	CAPI	Diff	Test Stat
Overall*	119.81	125.97	6.16	5.38
Section 4	51.28	50.75	-0.53	-1.53
Section 5*	445.06	520.51	75.45	2.88
Section 6*	224.72	258.77	34.05	5.56
Section 7	87.90	90.55	2.65	0.41
Section 8*	297.02	331.97	34.95	3.30
Section 9*	72.14	83.25	11.11	4.55
Section 12*	90.13	105.58	15.45	18.04
Section 15	92.20	92.30	0.10	0.04
Section 16*	345.95	315.68	-30.26	-3.01
Section 17*	18.64	19.85	1.21	3.55
Section 19*	133.63	152.69	19.06	2.77
Section 20*	38.76	39.67	0.91	2.43

significant.

7.2 Comparing Distributions

Figure 1 displays a quantile-quantile (QQ) plot of total monthly expenditures per CU in the CAPI against total monthly expenditures per CU in the PAPI, where expenditures were log-transformed and the circles represented percentiles. The pth percentile is the value where p percent of the data lay below or equal to the value. Under regularity conditions, if the CU-level sample expenditures from both the PAPI and the CAPI are the same, then the percentiles from the CAPI would approximately follow the percentiles from the PAPI and the QQ plot in Figure 1 should have its points arranged along a line with a slope of 1 and an intercept of 0. We did not

observe any substantial deviation from this line except at the 100th percentile.

We also plotted a QQ plot of monthly expenditure data which only included the 95th to 100th percentiles, where expenditures were log-transformed and the circles represented percentiles. Figure 2 showed how tail distributions differed between the PAPI and the CAPI. We observed that expenditures from the CAPI were slightly higher in high spending CUs even after accounting for inflation.

Similarly, we plotted QQ plots of percentiles in the original scale for the selected sections, such as Section 7, Section 12, Section 17, and Section 20. Figure 3 showed that Section 7 expenditures from the CAPI were higher than the ones from the PAPI. We did not observe any substantial deviation from the line in Section 12, Section 17, and Section 20.

We compared family incomes between the PAPI and the CAPI respondents. A QQ plot of family income showed that the percentiles of family income per CU from the CAPI approximately followed the percentiles from the PAPI except at the 100th percentile. A QQ plot of family income data which only included the 95th to 100th percentiles, however, showed that family income from the CAPI were slightly higher in high income CUs even after accounting for inflation.

Figure 4 presents side-by-side boxplots of the unweighted overall monthly counts separately for PAPI and CAPI. The middle line in the box corresponds to the median, the upper and lower bounds of the box correspond to the sample upper and lower quartiles, and the upper and lower whiskers correspond to the largest and smallest sample values, respectively. The median values were 18 for PAPI data

and 17 for CAPI data.

8. Summary

The preliminary results show a decrease in expenditures and counts in the CAPI data when compared to the PAPI data. However, we observed an increase in the expenditure per count in the CAPI data. Sections whose expenditures were thought to be somewhat underestimated with PAPI such as Section 7 showed increase in CAPI, while sections whose expenditures were thought to be somewhat overestimated with PAPI such as Section 20 showed a decrease in CAPI. The CAPI does not allow interviewers much liberty in assigning the counts, especially for the combined purchases. Therefore, the CE considers the count from CAPI to be more accurate. The PAPI and the CAPI data had similar distributions of both expenditure and counts. However, we observed an increase in variability in the CAPI, and noted a considerable number of extreme values in the CAPI data. It would be of interest to consider extensions to this work. For example, one could consider the analysis of extreme values in the CAPI data and investigate how they affect the variance and the bias of the estimators. In addition, one could further explore the sections with significant differences and conduct a more detailed analysis at the UCC level.

This study is not a full-scale cross sectional study. Although we selected sections with nearly exact translation, and then adjusted expenditures according to the CPI values in each section for higher comparability, we are aware of the fact that there are other factors we were not able to adjust.

9. References

Bradburn, N., Frankel, M., Baker, R., and Pergamit, M. (1991). A Comparison of Computer-Assisted Personal Interviews (CAPI) with Paper-and-Pencil Interviews (PAPI) in the National Longitudinal Study of Youth. *NLS Discussion Paper Series.* U.S. Bureau of Labor Statistics.

Bureau of Labor Statistics (1997). *BLS Handbook of Methods.* Washington, DC: U.S. Department of Labor.

Cho, M. and Eltinge, J. (2005). Modeling of Survey Response Rates and Reporting Rates in the U.S. Consumer Expenditure Interview Survey. *Proceedings of the American Statistical Association.* Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.

Cho, M., Eltinge, J., and Swanson, D. (2003). Inferential Methods to Identify Possible Interviewer Fraud Using Leading Digit Preference Patterns and Design Effect Matrices. *Proceedings of the American Statistical Association.* Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.

Cho, M., Eltinge, J., and Steinberg, B. (2004). Temporal Patterns of Survey Response Rates and Reporting Rates in the U.S. Consumer Expenditure Interview and Other Panel Surveys. *Proceedings of the American Association for Public Opinion Research.* Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.

Dietz, R., Duly, A., Groves, S., Hackett, C., Henderson, S., James, L., Kincaid, N., Paszkiewicz, L., Ryan, J., Shapiro, J., and Swanson, D. (2003). Final Report of the Interview CAPI Test Analysis Team. Manuscript, U.S. Bureau of Labor Statistics.

Eltinge, J., and Cho, M. (2003). Nonidentified Item Nonresponse and Reporting Rate in the U.S. Consumer Expenditure Interview Survey. *Proceedings of the 14th International Workshop on Household Survey Nonresponse.* Belgium. Available at http://www.nonresponse.org

Eltinge, J.L., Sukasih, A. and Weber, W. (2000). Feasibility of Constructing Combined Estimators using Consumer Expenditure Interview and Diary Data. Manuscript, Office of Survey Methods Research, U.S. Bureau of Labor Statistics.

Fuchs, M., Couper, M., and Hansen, S. (2000). Technology Effects: Do CAPI or PAPI Interviews Take Longer? *Journal of Official Statistics.* Vol. 16, No. 3, 2000, pp.273-286.

Lamas, E., Palumbo, T., and Eargle J. (1996). The Effect of the SIPP Redesign on employment and Earnings Data. No. 9606, U.S. Bureau of the Census.

Swanson, D., Cho, M., and Eltinge, J. (2003). Detecting possibly fraudulent or error-prone survey data using Benford's Law. *Proceedings of the American Statistical Association.* Survey Research Methods Section [CD-ROM]. Alexandria, VA: American Statistical Association.

Wolter, K.M. (1985). *Introduction to Variance Estimation.* Springer-Verlag, New York.
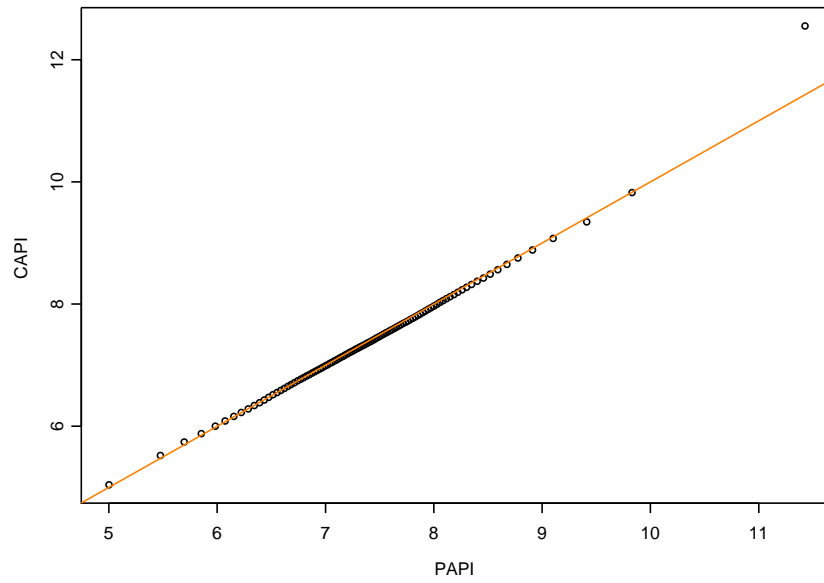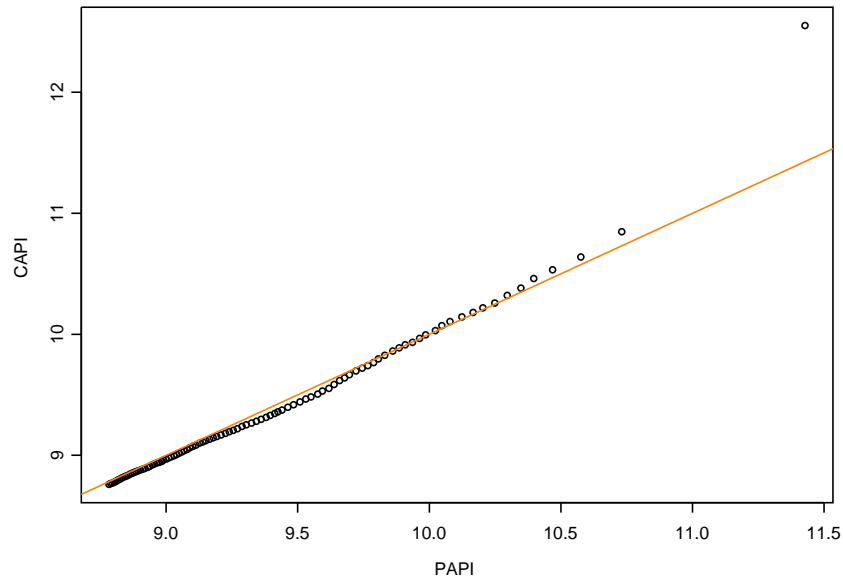
Figure 1: QQ plot of Log (Monthly Exp)

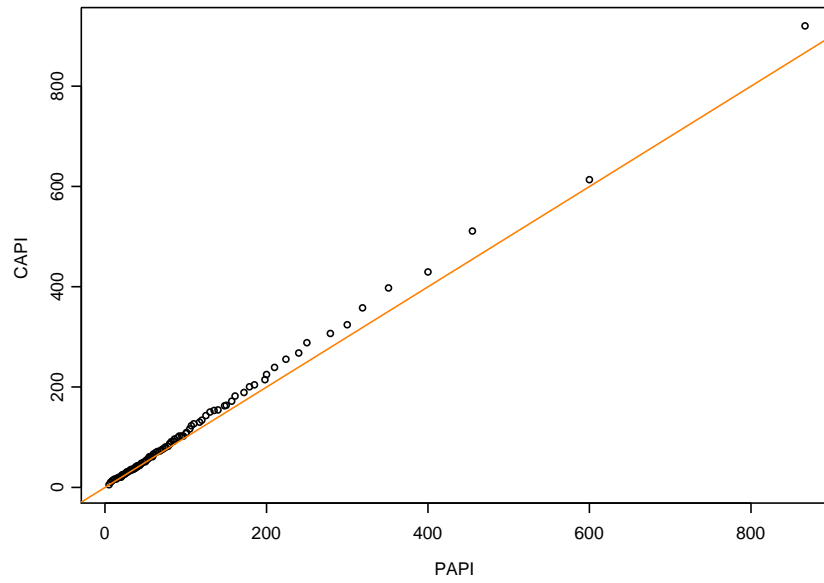Figure 2: QQ plot of Log (Monthly Exp) in upper Tail
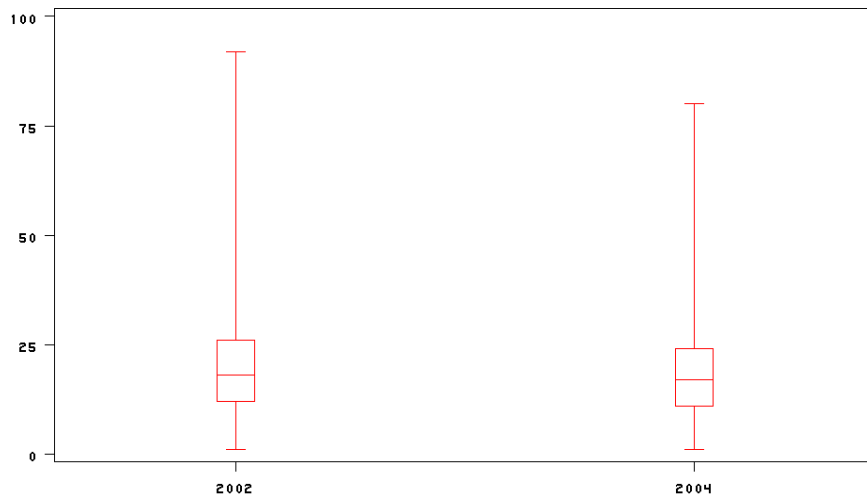
Figure 3: QQ plot of Monthly Exp: Section 7 (non zero, up to 99th percentile)

Figure 4: Box plot of Monthly Counts for a CU