Optimal Recall Period Length in Consumer Payment Surveys

Marcin Hitczenko

Consumer Payment Research Center
Federal Reserve Bank of Boston

September 4, 2014
Motivation

CPRC estimates frequency of use of payment instruments.

- Average # of cash payments per week among U.S. adults.
- Also credit, debit, check, ...

Reliable and comprehensive records for individuals ...

- may not exist (cash).
- may pose significant respondent burden (privacy and credit card statements).
- are relatively expensive to obtain.

Instead, we rely on consumer surveys.

- Ask respondent for # of payments made.
- Involve inherent cognitive biases.
Survey Design

Virtually every aspect of the survey will affect responses [3, 8, 13, 15].

- Survey mode: **web**, telephone, in-person [5, 7, 14].

- Type of recall: [1, 4]
  - **Specific**: How many payments made in last week?
  - **Typical**: How many payments made in typical week?

- **Recall period**: day, week, month, year? [6, 9, 10, 12].

In this work, we focus on recall period.

**Q1** Which recall period gives optimal results in estimating population means for cash, credit, debit, and check use?

**Q2** Can we improve estimates by assigning different recall periods to different respondents?
Consider a hypothetical researcher . . .

- Interested in population parameter $\omega$. **Ex:** weekly average.

- Selects $N$ individuals and asks for $\#$ payments made in last $\ell$ days.

- If no recall error, collects $A_\ell = \{A_{1\ell}, \ldots, A_{N\ell}\}$, where $A_{i\ell}$ is actual $\#$ of payments by respondent $i$.

- $\hat{\omega}(A_\ell)$ is estimate of $\omega$. **Ex:**

  \[
  \text{Weekly data (}\ell = 7) \quad \text{Yearly data (}\ell = 365) \\
  \hat{\omega}(A_7) = N^{-1} \sum_{i=1}^{N} A_{i7} \quad \hat{\omega}(A_{365}) = N^{-1} \sum_{i=1}^{N} \frac{7A_{i,365}}{365}
  \]

- Wants sampling design so that estimator is unbiased:

  \[E[\hat{\omega}(A_\ell)] = \omega.\]
What if recall data is used in unbiased estimator instead of actual data?

- $R_{\ell} = \{R_{1\ell}, \ldots, R_{N\ell}\}$ represents *reported* data.
- Evaluate $\hat{\omega}(R_{\ell})$ through mean-squared error:

$$
\text{MSE}(\hat{\omega}) = E[\hat{\omega} - \omega]^2
= \text{Var}(\hat{\omega}) + \text{Bias}^2(\hat{\omega}).
$$

- Generally, $\lim_{N\to\infty} \text{MSE}(\hat{\omega}) = \text{Bias}^2(\hat{\omega})$.

Focus is on population estimate, *not* individual recall.

- Perfect recall $\not\Rightarrow$ perfect estimates. **Ex:** Perfect recall for year, but interested in Thanksgiving week.
- Imperfect recall $\not\Rightarrow$ poor estimates. **Ex:** Regression to mean.

\[ A_{i\ell} \sim F(\text{mean} = \mu_{i\ell}) \text{ and } R_{i\ell} = pA_{i\ell} + (1 - p)\mu_{i\ell} \implies E[R_{i\ell}] = \mu_{i\ell}. \]
Q1: Which recall period gives optimal results in estimating population means for cash, credit, debit, and check use?

We rely on two datasets:

- **2012 Diary of Consumer Payment Choice (DCPC)**
  - 2,547 individuals from RAND’s American Life Panel (ALP).
  - Track payment activity for three consecutive days in October 2012.
  - Provides direct insight into $\omega$.
  - Patterns in data help define reasonable estimator forms.

- **2011-2012 Payment Recall Survey (PRS)**
  - 3,369 individuals from RAND’s American Life Panel (ALP).
  - About 1,850 individuals participated in both surveys.
  - Fielded in five phases between May 2011 and September 2012.
  - Recall the # of payments made for day, week, month, and year for all four major payment instruments.
  - Provides insight into quality of recall for different recall periods.
### DCPC Data

<table>
<thead>
<tr>
<th>Day</th>
<th>$ Value</th>
<th>PI Used</th>
<th>Other Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/1</td>
<td>13.39</td>
<td>cash</td>
<td>10:15AM, grocery store, ...</td>
</tr>
<tr>
<td>10/1</td>
<td>45.00</td>
<td>credit</td>
<td>4:00PM, restaurant, ...</td>
</tr>
<tr>
<td>10/3</td>
<td>200.00</td>
<td>credit</td>
<td>12:30PM, automobile, ...</td>
</tr>
</tbody>
</table>

**Table:** Data for one individual.

- 3-day periods randomly distributed in month.
- Provides the daily number of payments made with each payment instrument.
What does DCPC data look like?

Figure: Means and 95 percent confidence intervals for mean daily use.
We fit a mixed-effects log-linear model for each instrument:

\[ \# \text{ payments on day } t \sim \text{Poisson}(\mu_{it}). \]

- \( \log(\mu_{it}) = \mu_i + f(t) \)
- \( \mu_i \): random effect corresponding to individual.
- \( f(t) \): fixed effects corresponding to day-of-week or day-of-month.

Comparison of models finds

- Strong day-of-week effects for all four instruments.
- Evidence of monthly cycle for checks.
Back to our hypothetical researcher . . .

- $\omega = \text{mean \# payments per week in October 2012}.$
- SRS among target population (defined by ALP).
- Seemingly reasonable linear estimators:

$$\hat{\omega}_\ell = \sum_{i=1}^{N} w_{i\ell} R_{i\ell}$$

- $\ell = 1$: $w_{i1} = (N_d)^{-1}$, $N_d = \#\text{ reporting for day-of-week } d$.
- $\ell > 7$: $w_{i\ell} = \frac{7}{N}\ell$.

- Possible limitations:
  - Monthly ($\ell = 30$) and yearly ($\ell = 365$) recall is not quite right; intervals of 30 and 365 days do not have equal representation of each day of week.
  - Yearly recall ($\ell = 365$) extends to periods outside of October 2012.
PRS Data

- Respondents participate in 1-3 phases (3-9 months between surveys).

- In each phase of survey:
  - Sequence of payment instruments is randomized.
  - Order of day, week, and month is randomized; year is always last.
  - Day is randomly assigned within past week.

Data for one individual (in each phase of survey)

<table>
<thead>
<tr>
<th></th>
<th>Day in Last Week</th>
<th>Past Week</th>
<th>Past Month</th>
<th>Past Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash</td>
<td>2</td>
<td>8</td>
<td>30</td>
<td>350</td>
</tr>
<tr>
<td>Credit</td>
<td>1</td>
<td>7</td>
<td>25</td>
<td>200</td>
</tr>
<tr>
<td>Debit</td>
<td>0</td>
<td>2</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>Check</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table: Reported # of payments.
Example of timing of PRS and DCPC:

Figure: Timing for individual who took DCPC on October 15th, 2012.
We want to estimate bias of estimator based on recall period length $\ell$:

$$\text{Bias}(\hat{\omega}_{\ell}) = E[\hat{\omega}_{\ell}] - \omega.$$ 

- Linear estimator depends on $E[R_{i\ell}]$ and $\omega$.
- Use DCPC data to estimate $\omega$.
- Use PRS data to estimate $E[R_{i\ell}]$.
  - Use only PRS data from after August 15th 2012.
  - Adjust for any lag effect with daily recall (not found to be significant).
  - Randomization in PRS helps with various survey-specific effects.
    **Ex:** Dependence of response errors (weekly value should limit possible daily values).
- Bootstrap respondents to determine distribution of bias estimate:
  - Sample within respondents who took both surveys.
  - Sample within respondents who only took DCPC.
  - Sample within respondents who only took PRS.
For each bootstrapped sample, estimate

- Bias for each $\ell$.
- Which recall period minimizes absolute bias.

**Figure:** Bootstrapped bias (lines) and probability of minimizing absolute bias (bars).
Conclusions

- Optimal recall periods differ across payment instruments:
  - Week for cash
  - Month for credit, debit.
  - Year for check.

- Hurd and Rohwedder [9] suggest that optimal recall periods relate to the frequency of behavior.

- Survey of Consumer Payment Choice (SCPC):
  - Taken by those who took DCPC; also in October 2012.
  - Respondents choose recall period (week, month, year) to report typical # of payments.
  - Correspondence between DCPC data and reported SCPC results matches these results.

  **Ex:** Respondents who report cash on weekly basis show most consistency between SCPC (recall) data and DCPC (diary) data.
Q2: Can we improve estimates by assigning different recall periods to different respondents?

- Recall for individual $i$ is based on recall period $\ell_i$.
- $\omega_i = \text{weekly mean for individual } i$.
- If $\mathbb{E}[\omega_i] = \omega$ with respect to sampling scheme,

$$
\mathbb{E}[\hat{\omega} - \omega] \leq \sum_{i=1}^{N} \mathbb{E} \left| w_{i\ell_i} R_{i\ell_i} - \omega_i \right|.
$$

- Minimizing discrepancy between recall-based estimate of $\omega_i$ and true $\omega_i$ likely improves population estimate.

- Can optimal recall periods be predicted for individuals based on demographic information known ahead of survey?
For any individual $i$

- $R_{is\ell} = \#$ payments in last $\ell$ days reported on day $s$ (i.e. phase $s$).

- $B_{is} =$ recall period that produces closest approximation to $\omega_i$:

$$B_{is} = \arg\min_\ell |w_{is\ell} R_{is\ell} - \omega_i|.$$ 

- If we know $\omega_i$, we can determine $B_{is}$ from PRS data. **Ex:** If $\omega_i = 5$:

<table>
<thead>
<tr>
<th>Recall Period</th>
<th>Response</th>
<th>Scaled Estimate of $\omega$</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week</td>
<td>7</td>
<td>$7 \times \frac{7}{7} = 7$</td>
<td>2</td>
</tr>
<tr>
<td>Month</td>
<td>20</td>
<td>$20 \times \frac{7}{30} = 4.67$</td>
<td>$-0.33$</td>
</tr>
<tr>
<td>Year</td>
<td>200</td>
<td>$200 \times \frac{7}{365} = 3.83$</td>
<td>$-1.17$</td>
</tr>
</tbody>
</table>

- Sampling $\omega_i$ allows us to sample $B_{is}$; want to sample from

$$P(\omega_i \mid \text{DCPC,PRS}) \propto P(\omega_i \mid \text{DCPC})P(\text{PRS} \mid \omega_i)$$
A simple model:

- Distribution of $\omega_i \mid$ DCPC provided from random-effect models; related to $\mu_i$.

- Distribution of $R_{is\ell} \mid \omega_i$ takes form:

$$R_{is\ell} \mid \omega_i, \lambda_{is\ell} \sim \text{Poisson} \left( \lambda_{is\ell} \times \frac{\ell}{7\omega_i} \right)$$

- $\lambda_{is\ell}$ represents degree of reporting bias:
  - $\lambda_{is\ell} = 1 \implies$ unbiased recall
  - $\lambda_{is\ell} > 1 \implies$ overestimation
  - $\lambda_{is\ell} < 1 \implies$ underestimation

- Special case of model based on idea that recall is done via enumeration or rate-based estimation [2, 3].

- $\omega_i = 0 \implies P(R_{is\ell} = 0) = 1.$

Marcin Hitczenko (CPRC)
We run a MCMC procedure:

- Restrict data to individuals who took DCPC and participated in PRS after July 2012.
- Use cash only:
  - only instrument with very high adoption rates.
  - issue of non-adoption ($\omega_i = 0$) presents modeling computations.
- Compare weekly, monthly, and yearly recall:
  - currently adding daily recall.
- Assume $\lambda_{i,s,\ell} \sim \text{Gamma}(k_{\ell}, \tau_{\ell})$, independent across $i$, $s$ and $\ell$:
  - currently loosening independence assumptions (especially across $s$).
- Use non-informative hyper-priors: $P(k_{\ell}, \tau_{\ell}) \propto 1$.
- Generate draws of $\omega_i \mid \text{DCPC,PRS}$. 
Example 1: Prior vs. posterior estimates of $\omega_i$.

![Graph showing prior and posterior distributions of $\omega_i$.](image)

Based on PRS: \{2, 0.7, 0.7\}, \{1, 0.7, 0.2\}

Scaled DCPC average (x 7/3): 9.3

**Figure:** Prior (dashed) and posterior(solid) distributions of $\omega_i$. PRS estimates are ordered according to \{W, M, Y\} recall.
Example 2: Prior vs. posterior estimates of $\omega_i$.

Figure: Prior (dashed) and posterior (solid) distributions of $\omega_i$. PRS estimates are ordered according to \{W,M,Y\} recall.
Example 3: Prior vs. posterior estimates of $\omega_i$.

Figure: Prior (dashed) and posterior(solid) distributions of $\omega_i$. PRS estimates are ordered according to $\{W,M,Y\}$ recall.
In each posterior draw from MCMC algorithm:

- Given $\omega_i$, determine $B_{is}$:

<table>
<thead>
<tr>
<th>Individual ($i$)</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>2</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase ($s$)</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>...</td>
</tr>
<tr>
<td>$\omega_i$</td>
<td>5.4</td>
<td>5.4</td>
<td>5.4</td>
<td>1.2</td>
<td>...</td>
</tr>
<tr>
<td>$B_{is}$</td>
<td>week</td>
<td>month</td>
<td>week</td>
<td>month</td>
<td>...</td>
</tr>
</tbody>
</table>

- For the generated set $\{B_{is}\}$ fit models:
  - $P(B_{is} = \ell) \propto \exp(\alpha\ell)$
  - $P(B_{is} = \ell) \propto \exp(demo_i^T\beta\ell)$.

- Second model suggests that optimal recall period for individual relates to demographic information ($demo_i$).

- Demographic information includes age, gender, education, and income.
For each draw calculate deviance between models.

Figure: Fitted probabilities, $\hat{P}(B_{is} = 7)$ for three draws from MCMC.

- Averaging over draws, find little evidence that demographics predict the optimal recall period (p-value = 0.16).
- For all demographic combinations, the weekly recall period is always most likely to be best.
Conclusions

- Important to think carefully about what parameters we are trying to estimate, and whether sampling design is suited to optimize results.

- Evidence that optimal recall periods depend on what is being measured; linked to frequency of behavior?

- Not (yet?) enough evidence of heterogeneity in optimal recall lengths to justify assigning different recall periods to different respondents.

Limiting Factors/Future Work

- Diary data is not necessarily the truth [11].
  - Get more accurate records (if possible).

- Modeling assumptions may not be correct.
  - Expand analysis and flexibility of models.

- Results may not hold for broader populations; the ALP is not representative of US.


Models for $f(t)$

Let $\text{dow}(t) =$

\[
\begin{cases}
1 & \text{t is a Sunday} \\
\vdots & \vdots \\
7 & \text{t is a Saturday}
\end{cases}
\]

index the day of the week and

\[
\text{pom}(t) = \frac{\sum_{t'} 1 [t' \leq t, \text{ and } t', t \text{ in same month}]}{\sum_{t'} 1 [t', t \text{ in same month}]}
\]

define location within a month. Ex: $\text{pom} \left( \text{October 15}^{th} \right) = \frac{15}{31}$.

We consider three models for $f(t)$:

A $f(t) = \sum_{j=1}^{7} \beta_j 1 [\text{dow}(t) = j] + \alpha_1 \text{pom}(t) + \alpha_2 \text{pom}^2(t)$

B $f(t) = \sum_{j=1}^{7} \beta_j 1 [\text{dow}(t) = j]$

C $f(t) = \nu$. 
The first term in posterior, $P(\omega_i \mid \text{DCPC})$:

- Represents posterior estimate of $\omega$ given DCPC data.
- Defined via estimates of $f(t)$ and predictions of $\mu_i$ in model fits:

$$
\omega_i \mid \text{DCPC} = \sum_{j=1}^{7} \exp(\mu_i + \beta_j)
= \exp(\mu_i) \sum_{j=1}^{7} \exp(\beta_j)
$$

with $\mu_i \mid \text{DCPC} \sim \text{Normal}(\hat{m}_i, \hat{v}_i)$.

- Can be approximated with $\omega_i \mid \text{DCPC} \sim \text{Gamma}(k_i, \tau_i)$ with parameters $(k_i, \tau_i)$ determined by matching first two moments of distribution implied by $(\hat{m}_i, \hat{v}_i)$. 
Model for Recall Data

The model

\[ R_{i \ell | \omega_i, \lambda_{i \ell}} \sim \text{Poisson}\left( \lambda_{i \ell} \times \frac{\ell}{7} \omega_i \right) \]

is a special case of more general class of models:

\[ R_{i \ell | \omega_i, \lambda_{i \ell}} = \begin{cases} 
\lambda_{i \ell} A_{i \ell} & \text{w.p. } p(\ell) \\
\text{Poisson}\left( \frac{\ell}{7} \times \gamma_i \omega_i \right) & \text{w.p. } 1 - p(\ell)
\end{cases} \]

- \( p(\ell) \) defines probability of using enumeration (presumably decreases as \( \ell \) increases).
- \( \lambda_{i \ell} \) defines the bias in the enumeration estimation.
- \( \gamma_i \) defines the bias in the rate-based estimation.