Harnessing Commercial Data for Measurement: The Mint Bills Project

Michael Gelman, Michigan    Yuriy Gorodnichenko, UC Berkeley
Shachar Kariv, UC Berkeley   Matthew Shapiro, Michigan
Dan Silverman, Arizona State Steve Tadelis, UC Berkeley

FESAC
June 2015
Challenges of Measuring Spending, Income, and Assets

Hard to get large samples to provide information that is:

- Accurate
- Comprehensive
  - Across individuals
  - Within individual
- High frequency
- Timely
Our Approach – Primary Advantages

- Partner with Intuit and its mobile banking app Mint Bills
- Data generated in the course of business
- Made available for research at no incremental cost
- High frequency and timely: daily updates from population of the app’s users.
- Accurate: administrative records of transactions and balances
- Comprehensive
  - Across individuals: users are diverse
  - Within individual: all transactions and balances from an individual’s checking, saving, credit card, investment accounts
Our Approach – Primary Challenges

- No direct information on demographics
  - What is a user, and how does it relate to the household?
  - How do users vary by gender, age, education, occupation, etc.?
- Data are raw, not organized for research
  - Spending, income, and assets are not pre-categorized
- Sample is not randomly selected
Background – What is *Mint Bills*?
Mint Bills

- Mint Bills is an app for mobile phones, tablets, and the web.
- Has registered more than 12 million registered users since its founding in 2007.
- Users can integrate information from nearly any financial account with a web-based portal.
- Users provide Mint Bills with the credentials necessary to access these portals and, every day, Mint Bills automatically logs into and scrapes the associated webpages.
Mint Bills

Overview

- **Cash**: $3,832
- **Payments due**: -$1,159
  - min -$89
- **Investments**: $12,230
- **Credit card debt**: -$964
- **Offers**: 2 new offers (3 total)
- **Credit guard**: 729 (2 new alerts)

Your Alerts Stream 0
### Mint Bills

#### Overview

<table>
<thead>
<tr>
<th>Cash</th>
<th>Bills</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3,832</td>
<td>-$1,432</td>
</tr>
<tr>
<td><strong>Investments</strong></td>
<td><strong>Credit Card Debt</strong></td>
</tr>
<tr>
<td>$27,279</td>
<td>-$894</td>
</tr>
<tr>
<td><strong>Offers</strong></td>
<td><strong>Credit guard</strong></td>
</tr>
<tr>
<td>2 new offers (9 offers)</td>
<td>790 (2 new alerts)</td>
</tr>
</tbody>
</table>

#### Total Cash

- **Pay Comcast now**

- **Chase**
  - Checking (...5586) $653

- **Chase**
  - Checking (...2345) $2,369

- **Wells Fargo**
  - Savings (...3699) $810
Mint Bills
Who is in the *Mint Bills* data?
Table 1: Mint Bills vs. ACS Demographics

<table>
<thead>
<tr>
<th></th>
<th>Mint Bills</th>
<th>ACS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female</strong></td>
<td>40.07</td>
<td>51.41</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-20</td>
<td>0.59</td>
<td>5.72</td>
</tr>
<tr>
<td>21-24</td>
<td>5.26</td>
<td>7.36</td>
</tr>
<tr>
<td>25-34</td>
<td>37.85</td>
<td>17.48</td>
</tr>
<tr>
<td>35-44</td>
<td>30.06</td>
<td>17.03</td>
</tr>
<tr>
<td>45-54</td>
<td>15</td>
<td>18.39</td>
</tr>
<tr>
<td>55-64</td>
<td>7.76</td>
<td>16.06</td>
</tr>
<tr>
<td>65+</td>
<td>3.48</td>
<td>17.95</td>
</tr>
<tr>
<td><strong>Highest degree</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than College</td>
<td>69.95</td>
<td>62.86</td>
</tr>
<tr>
<td>College</td>
<td>24.07</td>
<td>26.22</td>
</tr>
<tr>
<td>Graduate School</td>
<td>5.98</td>
<td>10.92</td>
</tr>
<tr>
<td><strong>Census Bureau Region</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>20.61</td>
<td>17.77</td>
</tr>
<tr>
<td>Midwest</td>
<td>14.62</td>
<td>21.45</td>
</tr>
<tr>
<td>South</td>
<td>36.66</td>
<td>37.36</td>
</tr>
<tr>
<td>West</td>
<td>28.11</td>
<td>23.43</td>
</tr>
</tbody>
</table>

Note: The sample size for Mint Bills is 59,072, 35,417, 28,057, and 63,745 for gender, age, education and region, respectively. The sample size for ACS is 2,441,532 for gender, age, and region, and 2,158,014 for education.
What’s in the *Mint Bills* data?
Table 2: Transactions and Accounts

<table>
<thead>
<tr>
<th></th>
<th>Mint Bills</th>
<th>SCPC</th>
<th>P_5</th>
<th>P_25</th>
<th>P_50</th>
<th>P_75</th>
<th>P_95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily Transactions</td>
<td>4.54</td>
<td>2.33</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Credit card</td>
<td>1.23</td>
<td>0.50</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Checking account</td>
<td>3.03</td>
<td>0.91</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Saving account</td>
<td>0.22</td>
<td>n/a</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Accounts</td>
<td>5.84</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Credit card</td>
<td>3.58</td>
<td>4.00</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Checking account</td>
<td>1.35</td>
<td>1.60</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Saving Account</td>
<td>0.79</td>
<td>1.80</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Note: In total, 72,902 Mint Bills users generated 57,731,354 transactions over the 300 days of the study period. The Boston Fed's Survey of Consumer Payment Choice (SCPC) is conducted with a sample of approximately 2,100 members of the online survey, the American Life Panel.
Balance Data

Table 3: Account Balances

<table>
<thead>
<tr>
<th>Panel (a) Bank</th>
<th>Mint Bills</th>
<th>SCF</th>
<th>Mint Bills</th>
<th>SCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>$2,200</td>
<td>$4,000</td>
<td>$14,415</td>
<td>$30,493</td>
</tr>
<tr>
<td>Checking</td>
<td>$1,400</td>
<td>$1,900</td>
<td>$6,969</td>
<td>$8,338</td>
</tr>
<tr>
<td>Saving</td>
<td>$400</td>
<td>$3,500</td>
<td>$6,476</td>
<td>$17,312</td>
</tr>
<tr>
<td>Money Market</td>
<td>$900</td>
<td>$22,500</td>
<td>$12,076</td>
<td>$90,033</td>
</tr>
<tr>
<td>C.D.</td>
<td>$500</td>
<td>$16,000</td>
<td>$12,734</td>
<td>$63,859</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel (b) Credit Card</th>
<th>Mint Bills</th>
<th>SCF</th>
<th>Mint Bills</th>
<th>SCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Balance</td>
<td>$3,600</td>
<td>n/a</td>
<td>$7,228</td>
<td>n/a</td>
</tr>
<tr>
<td>Credit Limit</td>
<td>$11,900</td>
<td>$15,000</td>
<td>$23,019</td>
<td>$40,197</td>
</tr>
<tr>
<td>Utilization Ratio</td>
<td>0.45</td>
<td>n/a</td>
<td>0.48</td>
<td>n/a</td>
</tr>
<tr>
<td>Revolving Debt</td>
<td>$3,500</td>
<td>$2,500</td>
<td>$5,828</td>
<td>$5,707</td>
</tr>
<tr>
<td>APR</td>
<td>18.00%</td>
<td>13.90%</td>
<td>18.46%</td>
<td>13.57%</td>
</tr>
</tbody>
</table>

Notes: Mint Bills data are from the sample of 75,000, in the study period covered by Gelman et al. (2014). Survey of Consumer Finance analogues are weighted with population weights. Amounts are conditioned on a user/respondent having that type of account. Last three rows of Panel (b) condition on having revolving debt.
Measuring Income and Expenditure from Transactions
Distributions of User Income
Total Weekly Spending

Total (well-linked)

Total (all users)
Integrating Income and Spending Measures
Integrating Measures: The Spending Rate

Annual after tax average in the 2013 CE Survey: 0.91

- 1.14 for hhlds earning <$70K
- 0.73 for hhlds earning ≥ $100K
Integrating Measures: Spending Responses to the Arrival of Anticipated Income

- Gelman et al., *Science* (2014)
- Does spending respond to the arrival of anticipated income?
- If they know the money is coming, why don’t consumers smooth their spending?
Response of Total and Non-Recurring Spending to Paycheck Income
Further Integration: Response Heterogeneity By Liquidity

Bottom Third of the Liquidity Distribution

Top Third of the Liquidity Distribution
Responses to Income Shock: The 2013 Govt. Shutdown

Paycheck Income

Total Spending
Making These Data More Useful for Official Statistics

- Mint Bills data have proved useful for research
- Work in progress should make the data more useful for official statistics
- Survey Mint Bills users
  - Assess representativeness
  - Understand the relationship between users and hhlds.
  - Evaluate integration with the app
- Surveys will thus facilitate weighting and aggregation
- Allow real-time income, spending, and debt indices to complement official statistics
Some Challenges of Generating Time Series

- Complex “seasonals”
  - Day of week
  - Point in month
  - Paycheck effects
  - Trading day and reconciliation effects
  - Interactions of all of the above

- Short sample (Mint can help)
- Changing composition of users
- Evolving features and uses of the app
Making These Data More Useful for Official Statistics

- Use the Mint Bills apparatus as a tool of survey research
- Bring the apparatus to representative samples
  - American Life Panel
  - PSID
  - NLSY
- Complement existing income and spending measures
- Decrease the burden on survey respondents without sacrificing accuracy.
- Use app to replace components of CE Survey or Diary.
Making These Data More Useful for Official Statistics

- Providing responsible access to the data is a challenge.
- Priority is on protecting the confidentiality of users and the business interests of the firm
- Careful aggregation has allowed registered researcher access to data associated with Gelman et al. (2014).
- How can we facilitate better cooperation between researchers, firms, and official statistics?