The Process of NAS Panels

David Betson

Presentation at CE Workshop

October 16, 2012

Initiation of the Project

- *Projects undertaken by the NRC are initiated by federal agencies to pursue a question they have or have been directed by Congress to answer*
- *The importance of the "Charge to the Panel"*
- *Selection of Study Director*

Selection of the Panel of Experts

- *Nomination of Chair*
- *Nomination of Panel Members to create diversity of expertise and perspective*
- *Confirmation and Appointment of Chair and Panel Members*

The Panel on Redesigning of the CE

The Panel was composed of 13 members representing a variety of disciplines and skills

- *Sample design*
- *Data collection and technology*
- *Survey measurement and cognitive design*
- *Economics and Public Policy*
- *Past direct involvement with the CE's design and implementation*
- *Users of the CE*

The CE Panel on Redesigning the BLS Consumer Expenditure Surveys

- *Don Dillman, Department of Sociology, Washington State University (Chair)*
- *David Betson, College of Arts and Letters, University of Notre Dame*
- *Mick Couper, Institute for Social Research, University of Michigan*
- *Robert Gillingham, Independent Consultant, Potomac Falls, VA*
- *Michael Link, The Nielsen Company, Marietta, GA*
- *Bruce Meyer, Harris School of Public Policy Studies, University of Chicago*
- *Sarah Nusser, Department of Statistics, Iowa State University*
- *Andy Peytchev, RTI International, Research Triangle Park, NC*
- *Mark Pierzchala, Independent Consultant, Rockville, MD*
- *Robert Santos, The Urban Institute, Washington, DC*
- *Michael Schober, New School for Social Research, New York City, NY*
- *Melvin Stephens, Jr. Department of Economics, University of Michigan*
- *Clyde Tucker, Independent Consultant, Vienna, VA*
- *Carol House, Study Director, Committee on National Statistics*

Bias and Conflict of Interest

- *Once a year, Panel Members are required to complete 'Conflict of Interest Statements'*
- *At the initial meeting of the Panel, a discussion is held where Panel Members discuss any potential biases and conflicts of interest they may affect how their views on the Panel's work. At this meeting, the Panel decides whether additional Panel Members may be needed.*
- *While travel expenses are paid, the Panel Members volunteer their time and are otherwise not compensated*

Gathering Information and Evidence

- *Open Meetings – Presentations and Discussion (February 8, 2011)*
- *Workshops or Conferences*
  - *Household Survey Producer Workshop (June 1 and 2, 2011)*
  - *Redesign Options Workshop (October 26 and 27, 2011)*
- *Panel Activities*
  - *Survey of Existing Literature*
  - *Data Collection and Analysis*
  - *Outreach (attended CE User's Workshop)*
  - *Gaining Experience (attempt to answer the CE survey)*

Arriving at a Consensus Report

Closed Meetings, Emails, Web Conferences

First Meeting – February 9, 2011
Second Meeting after first workshop – June 3, 2011
Third Meeting after second workshop – October 29, 2011

A lot of emails and phone calls, drafting of report

Fourth Meeting – January 25 and 26, 2012
Web Conference Call – March 2, 2012

Report sent to Review – June 8, 2012

What is a Consensus Report?

The set of conclusions and recommendations that address the questions posed in the Panel's charge that each Panel Member can 'live with' but don't necessarily represent their most preferred set of conclusions and recommendations.

'Live with' in the context of a NAS Panel means that in the Panel's expert judgment, the Panel's recommendations are supported by the best available evidence.

Report Review

- *NAS appoints a Review Monitor and a Review Coordinator to act as 'editors' in the review process*
- *Reviewers are selected by NAS to review the report to insure that the report is responsive to the Panel's Charge and that all conclusions and recommendations are supported by evidence*
- *The Study Director and Chair with the assistance of the Panel prepare a response to review where reviewers comments are addressed by making changes to the text or providing an explanation of why no change is needed*
- *The Review Monitor and Coordinator must sign off on the adequacy of the Response to Review prepared by the Panel*

Release of Report

- *Sponsors are provided a prepublication version of the report and are briefed by Panel Chair and Staff Director*
- *Public Release of Report*
  - *NAS Web Site*
  - *Public Briefing*

Issues with The Current Consumer Expenditure Surveys

For

Measuring What We Spend: Toward a New Consumer Expenditure Survey

Workshop Discussion of NRC Report

October 16, 2012

Don A. Dillman, Chair

National Academies Panel on Redesign Options for the Consumer Expenditure (CE) Surveys

My Purpose

- Discuss measurement and non-response issues associated with the current CE Surveys.
- Provide background for understanding recommendations of the National Academies Panel on Redesigning the BLS Consumer Expenditure Surveys.

But, first some greatly deserved thank-you's

- To the Bureau of Labor Statistics for:
  - The Gemini Project, which had already done part of what a NA Panel might otherwise be asked to do.
  - Assistance provided regularly throughout the Panel's work.
- Presenters and discussants at Panel sponsored meetings and workshops.
- Contractors for two proposed redesigns of the CE surveys.
- The CE Panel members who brought survey design expertise from a variety of scientific perspectives.

Three considerations that made coming to grips with issues facing the CE surveys challenging

- Diverse data requirements drive the identification of CE design issues.
  - Consumer Price Index.
  - Administration of federal/state government programs.
  - Policy analyses and economic research.
- Different CE design issues are linked to different data requirements
- Needed survey design expertise is held by different organizations and people.

Diverse Expertise

- Data collectors—organizations and field representatives
- Respondents who must be willing and able to answer the questions.
- Agency staff who summarize, analyze and report CE data.
- Professionals outside government who analyze data, often from various disciplinal perspectives.
- The scientific disciplines involved in constructing valid and reliable measures of consumer expenditures and the survey methods used to collect them.

Why the CE Surveys need to be redesigned (See Conclusions 5-1 and 5-2)

- Underreporting of expenditures for both the Diary and Interview are a concern.
- Underreporting varies across categories and appears to be less for interview.
- The questions are different in each mode as is the context for asking them and that contributes to differences in measurement between modes. .
- Both the diary and interview can and need to be improved.

Sources of Underreporting

(Conclusions 5-3 to 5-14)

- I will begin with the interview and then move on to the diary survey.
- Our goal here was to identify specific issues that might be focused on with hope of getting improvement through redesign efforts.
- Panel conclusions stemmed from a variety of sources: published literature, disciplinal perspectives, and the experiences of panel members who completed the CE Interview and/or diary.

Synopsis of the interview task

- 5 in-person interviews, ~65 minutes each, are conducted with each consumer unit (household) three months apart.
- Interview #1 asks demographic information, major durable goods, and a one month recall of expenditures.
- Interviews. #2-5 are quite repetitive asking respondents to report details of expenditures for the last three months.
- A respondent "booklet" includes 36 pages, each with 7-70 items on it to help identify specific expenditures.
- Detailed assets and changes from earlier time periods are generally asked in interview. #2 and #5.

Some additional aspects of response situation (from field representative debriefing surveys)

- Interview is usually completed with one person in the consumer unit who needs to report for all persons.
- Records less likely to be extensively used in interview (31%) than NOT used (39%).
- Respondent has to learn rules for assigning dates of purchase, delivery and payment, which may differ.

Detail requested for the last 3 months (interviews 2-5) can be daunting, e.g.

- Did you purchase any “pants, jeans, or shorts” (in the last three months)?
  - Please describe the item
  - Was this purchased for someone inside or outside of your household?
  - For whom was this purchased? (Enter name, age, sex).
  - When did you purchase it?
  - How much did it cost?
  - Did this include sales tax?

A difficult asset question from the fifth interview

- On the last day of last month, what was the total balance or market value (including interest earned) of checking accounts, brokerage accounts and other similar accounts?
- How does the amount your household had on the last day of last month compare with the amount your household had on the last day of last month one year ago?

Two major barriers to accurate interview reporting

- Knowledge and recall of expenditure amounts and details.
  - R may not know this information.
  - R may have known, but cannot remember.
- Motivation to report amounts and details
  - R may be unwilling or reluctant to share what is known.
  - R may be unwilling to make effort to determine amounts and details of expenditures.

Why is <u>knowledge and recall </u>a barrier to accurate reporting?

- Only one person likely to do the interview.
- Records are infrequently used.
- Even if records are used, they may not correspond to rules for reporting date of expenditure and details.
- Purchasing and paying for products and services has become much more complicated over the last 30 years.

Societal Changes in how consumer expenditures are made makes recall more difficult

- Greater range of consumer expenditures.
- Retail outlets more likely to sell unrelated items in single purchase (e.g. food + luggage + motor oil +socks)
- On-line purchases are growing.
- Multiple payment methods may be used by one person—cash, debit-card, credit card, and checks.
- Automatic deductions {paycheck(s), credit card(s), bank account(s)} may not be registered in memory .
- Transactions often made without cognitive attention to amount; the card is simply “swiped” and for smaller purchases no signature required.

Societal Changes in how consumer expenditures are made makes recall more difficult(2)

- List prices set to allow discounts on discounts.
- Store credits and discounts sometimes applied to total amount of all purchases, making knowledge of purchase price less obvious.
- Enormous variation in receipt structures, including abbreviations, makes knowledge of cost more difficult.
- Day item obtained may not be date payment made.
- **Observation: Purchasing behavior is much more complex than when CE interviewing (and Diary) methods were developed, and these changes work against the recall of specific purchases and cost.**

Don A. Dillman, NRC Workshop on CE Survey Redesign , October 16, 2012

15

Societal changes in household structure add to the recall problem

- More households have multiple wage earners.
- Households are more likely to include unmarried partners.
- Household members more likely to have expenditure obligations to other households (e.g. children of divorced parents).
- Partners/spouses less likely to inform each other of all expenses.
  - Partner agreements, “I pay for food, you pay for rent,” may compound the knowledge problem.

Don A. Dillman, NRC Workshop on CE Survey Redesign , October 16, 2012

16

For example: How much did this necktie cost?

- Sign in store: "Neck wear: assorted vendors. Buy 1 (59.95). Get 1 free" (Sign was on multiple tables)
- Bought two ties (and a third item)
- Clerk: Use of store credit card "today" gives additional 20% off of purchase.
- Time of year Charity donation today would give 10% off current purchase and 10% of all expenditures would come as gift certificate after December 31st. I declined.
- Clerk offers to send receipt by email instead of my taking receipt with me.

Outcome: Necktie purchase

- Price was 59.95 on tie, but receipt said 41.65 was "today's price".
- Clerk gave me 20% off of 41.65 for total of 33.32/tie.
- Two for one did not apply because I provided store credit card and clerk took the 20%.
- Checked online billing the next day to see what was on credit card—30 lines of single spaced code for bill.
- Very difficult to know how much this tie cost.

Adding up the challenge

Purchases from more sources . . .
+ less association of product with store
+ fewer purchases at known list price
+ more payment streams (methods)
+ automatic deductions for some expenditures
+ less absorption of final price at purchase time
+ less sharing of detail within households

= less certainty on specific expenditures

The knowledge/recall problem in methodological perspective

- One cannot report what one does not know, e.g. expenditures never "encoded" or partner expenditures not shared.
- The interview cannot rely on establishing recall of details by connecting expenditures to a particular payment method (check book), physical location of purchase (shopping trip), or routine behavior of a member of the consumer unit.
- Cognitive methods that could be used to enhance memories (e.g. calendar of events over last 3 months, or personalizing recall methods to individuals) are likely to make an already long interview more burdensome.

Motivation has multiple aspects

- Belief by some that answering detailed expenditure and asset questions are an invasion of privacy or beyond the need of government to know.
- The interview situation discourages obtaining information from other members of the household unless they are present.
  - Partner/teenager expenses
  - Searching for receipts and perusing records

Don A. Dillman, NRC Workshop on CE Survey Redesign , October 16, 2012 21

Motivation in Methodological Perspective

- Some requests for information (alcohol purchases, gambling costs) reduce reporting because of social undesirability.
- Seeking great detail that requires significant effort to calculate or find , encourages satisficing: "This "number" is good enough".
- Seeking information respondent cannot easily provide encourages estimation.
- Estimation often results in telescoping, ie. reporting significant expenditures for more than three months.

Don A. Dillman, NRC Workshop on CE Survey Redesign , October 16, 2012 22

Motivation in Methodological Perspective (2)

- Answering "yes" to questions, e.g. Did you take an overnight trip or did you purchase any other clothing, typically leads to additional questions.
- All five interviews can include dozens of potential screener questions, and respondents soon learn that a "no" or "not that I remember" results in fewer questions and a shorter interview.
- The inability of respondents to give accurate answers and a belief that the answer they can give will not be useful, may help justify saying "no".

Mutual interests of respondent and the interviewer

- The respondent typically wants interview to be shorter and less work.
- The Interviewer also wants interview to be shorter so the respondent will agree to another interview.
- Thus, interviewers may not encourage respondents to look for receipts, go through records, divide receipts, or use other recall methods.

Effects of switching data collection to the telephone

- 38% of CE interviews are completed by telephone, especially the later ones.
- Recall visual aid (the notebook)not used.
- Telephone interviews obtain fewer “yes” answers to screener questions.
- Receipts and other records are less likely to be used.

In summary, interview data quality may be negatively affected by…

- Interviewing only one person in multi-person households, i.e. proxy reporting.
- The long (three month) recall period.
- Lack of knowledge respondent has for providing extreme detail requested.
- Lack of motivation for making effort needed to provide accurate answers.
- The repetitive nature and length of the interviews.
- The easy avoidance of additional questions by saying “No, I didn’t purchase any of that.”

The Diary completion task is also demanding

- Households are asked to report everything they purchase for two weeks.
- Week 1 diary; in-person delivery with selected questions and instructions.
- Week 2 diary; in-person delivery when Week 1 diary is picked up.
- Interviewer returns a third time to pick up Week 2 diary.
- Respondent is encouraged to make diary entries each day purchases are made.

Learning to complete the diary can be confusing

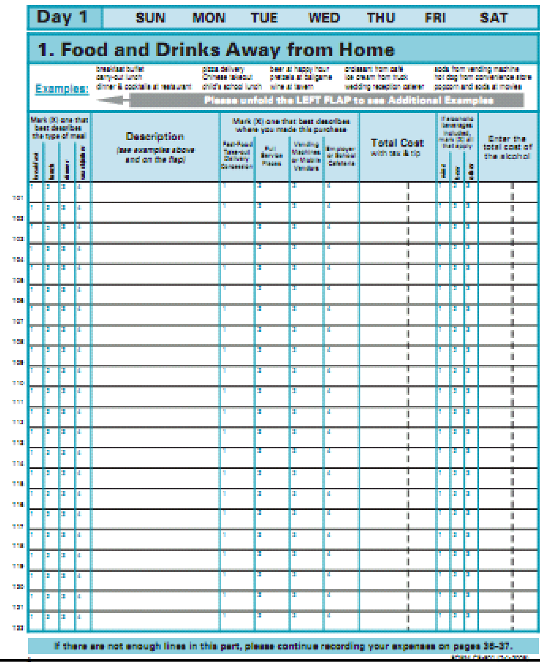
- 44 numbered pages + covers + flaps = 52 surfaces of information
- Respondent asked to write in names of others in household for whom they are reporting.
- 15 of pages provided instructions, 28 pages are laid out by “day” and have labeled tables for:
  - 1 Food and Drinks Away From Home
  - 2 Food and Drinks for Home Consumption
  - 3 Clothing, Shoes, Jewelry and Accessories
  - 4 All Other Products, Services and Expenses
- Respondent asked to provide up to six pieces of information for each entry.

Understanding requires round-about processing, e.g.

- P. 9 asks for Day 1 Food and Drinks for Home.
- Right flap of back cover provides answers to frequently asked questions.
- P.3 has instructions for "How to fill out your diary" including request to record expenses each day.
- P.2 has general instructions and section on what not to record.
- Examples for Food and Drinks are on P. 5, e.g. specify white, wheat or rye bread, and instant vs. ground coffee.
- Subsequent Daily pages for Food and Drinks are on Pp.13, 17, 21, 25, 29 and 33; "additional pages". for overflow are Pp. 38-41

Graphical navigation guidance is quite limited

- All pages mix blue, gray and black positive and negative print; full color limited to outside covers.
- Numbers are assigned to Day (1-7) as well as expenditure categories (1-4) and pages (1-44), so may not be an effective guide.
- If diary is not filled out each day expenses are made, flipping of non continuous pages is required.



31

How to Fill Out Your Diary

The diary is divided into 7 days and each day is divided into 4 parts. Enter each item in the appropriate part for each day.

<u>These are the 4 parts within each day of the diary:</u>

1. **Food and Drinks Away from Home**
   - Mark one of the four choices that best describes the type of meal and describe briefly.
   - Mark one of the four choices that best describes where you made the purchase.
   - Enter the total cost with tax and tip.
   - If alcohol was part of the purchase, check whether it was wine, beer, and/or other alcohol and enter the total cost of the alcohol.
2. **Food and Drinks for Home Consumption**
   - Describe the item.
   - Mark whether the item was fresh, frozen, bottled/canned, or other.
   - Enter the cost without tax and deduct any discounts or coupons.
   - Mark the last column if the item was purchased for someone not on your list (e.g. gifts).
3. **Clothing, Shoes, Jewelry, and Accessories**
   - Describe the item and enter the cost without tax.
   - Mark the appropriate sex and age range of the person for whom the item was bought.
   - Mark the last column if the item was purchased for someone not on your list (e.g. gifts).
4. **All Other Products, Services, and Expenses**
   - Describe the item and enter the total cost without tax.
   - Mark the last column if the item was purchased for someone not on your list (e.g. gifts).

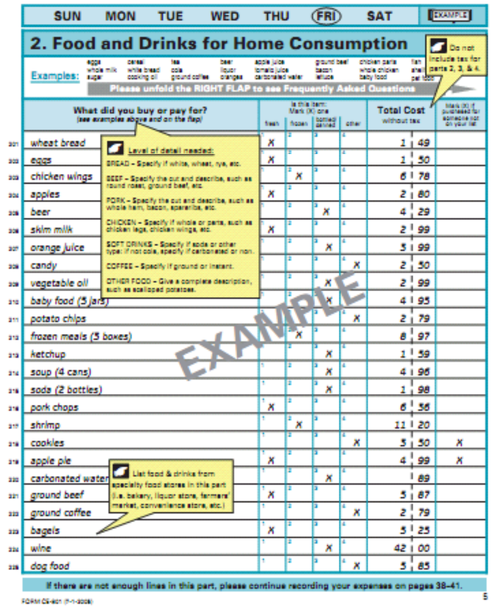
There is an "Additional Pages" section on pages 36–44 in case you run out of lines on any particular day.

Look on the next 4 pages for examples and tips on how to record your purchases.

***Please Note: If you are unsure about whether to include an item or where to record an item, write it down wherever it seems best or make a note and ask your field representative.**

32

Redesign , October 16, 2012



SUN MON TUE WED THU FRI SAT EXAMPLE

2. Food and Drinks for Home Consumption

Do not include tax for parts 2, 3, & 4.

Examples: eggs, whole milk, sugar, cereal, white bread, cooking oil, tea, cola, ground coffee, beer, liquor, oranges, apple juice, tomato juice, carbonated water, ground beef, bacon, lettuce, chicken parts, whole chicken, baby food, fish, shell..., pet food

Please unfold the RIGHT FLAP to see Frequently Asked Questions

What did you buy or pay for? (see examples above and on the flap)	Is this item: Mark (X) one — fresh	frozen	bottled/canned	other	Total Cost without tax	Mark (X) if purchased for someone not on your list
wheat bread	X				1 49	
eggs	X				1 50	
chicken wings		X			6 78	
apples	X				2 80	
beer			X		4 29	
skim milk	X				2 99	
orange juice			X		5 99	
candy				X	2 50	
vegetable oil			X		2 99	
baby food (5 jars)			X		4 95	
potato chips				X	2 79	
frozen meals (5 boxes)		X			8 97	
ketchup			X		1 59	
soup (4 cans)			X		4 96	
soda (2 bottles)			X		1 98	
pork chops	X				6 56	
shrimp		X			11 20	
cookies				X	5 50	X
apple pie	X				4 99	X
carbonated water			X		89	
ground beef	X				5 87	
ground coffee				X	2 79	
bagels	X				5 25	
wine			X		42 00	
dog food				X	5 85	

Level of detail needed:
BREAD – Specify if white, wheat, rye, etc.
BEEF – Specify the cut and describe, such as round roast, ground beef, etc.
PORK – Specify the cut and describe, such as whole ham, bacon, spareribs, etc.
CHICKEN – Specify if whole or parts, such as chicken legs, chicken wings, etc.
SOFT DRINKS – Specify if soda or other type: if not cola, specify if carbonated or non.
COFFEE – Specify if ground or instant.
OTHER FOOD – Give a complete description, such as scalloped potatoes.

List food & drinks from specialty food stores in this part (i.e. bakery, liquor store, farmers' market, convenience store, etc.)

EXAMPLE

If there are not enough lines in this part, please continue recording your expenses on pages 38–41.

FORM CE-801 (7-1-2008) 5

33

Knowledge as a barrier to Completing the Diaries

- In principle less of a problem unless
  - Respondent doesn't ask other unit members for expenses.
  - Receipts not kept for some purchases.
  - Receipts make it difficult to identify items (abbreviations)
- Knowing and applying rules may produce inaccurate recordings that are difficult to catch.
- Information may not be available in a timely way, e.g. auto deductions from bank account.
- Society changes (see interview discussion) also apply here.

Motivation as a barrier

- Recording requested detail on daily basis may be seen as a considerable burden.
- Long itemized receipts may be difficult to enter, e.g. if discount applied to an entire bill and abbreviations used.
- Abbreviations may be impossible to interpret.
- Respondents may delay shopping trip to avoid the "need" to record.

In Sum, Diary responses are negatively affected by…

- …both knowledge (e.g. proxy reporting) and motivation.
- In addition, design and layout procedures present problems—some respondents decide only to collect receipts because of how difficult the task seems to be.
- An outcome. Some diaries (especially second week) do not get completed.

Is one of these modes (interview vs. diary) fundamentally better than the other?

- A difficult question to answer.
- The original goal was to collect different kinds of information—big and recurring expenses from interview and small detailed ones from diary.
- Each now asks some of the same topics, but in different ways, e.g. food for home consumption.
  - Interview- "What has been your or your household usual WEEKLY expense for grocery shopping?"
  - Diary—A listing of each item that asks, e.g. white vs. wheat bread, types of meat, and if purchase was for someone else.
- These are quite different questions, subject to different kinds of errors associated with "estimation" vs. "reporting of details".

A cognitive perspective on the search for improved accuracy

- Typical approaches to getting better measurement include:
  - Ask more detailed questions.
  - Create a memory context, e.g. daily activities and event history of last three months.
  - Personalize questions to how people are most likely to recall expenditures.
- **Observation: These kinds of efforts make a demanding survey interview/diary even longer and more demanding?**

Possible Synopsis of Major problems of Interview vs. Diary as now implemented

While the <u>interview</u> asks people to report detail they never knew or can't recall (primarily a knowledge problem)…

The <u>diary</u> asks people to do something they won't do (primarily a motivation problem).

A related question: Is the extreme detail necessary?

- The CE surveys serve multiple purposes:
  - CPI Index
  - Administration of government programs
  - Policy analyses
- Are less detailed estimation questions adequate for some of these purposes, but not others?
- Is the extreme detail necessary for any of these purposes (e.g. white vs. wheat bread)?
- Would recall and reduction of proxy reporting in self-administered reporting produce better data than on-the-spot-estimates of expenditures.

Other demand issues (Conclusions 5-15 to 5-16).

- Response rates for diary and interview have both been declining and are in the low 70's.
  - Interview quality seems to go down with the push Field Representatives face from respondents to use the telephone.
  - Diary quality seems to go down with the second week of collection.
- Administrative data may provide a potential to "fill-in" data.

Conclusion

- The issues described today are only some of the issues that involved extensive discussion among panel members, and with most panels there were many viewpoints expressed.
- These discussions provided background for developing recommendations that will be discussed by Carol House, National Academies Study Director for the CE Panel and Andy Peytchev, a member of the Panel.

Measuring What We Spend: Toward a New Consumer Expenditure Survey

Panel on Redesigning the BLS Consumer Expenditure Surveys
Don A. Dillman and Carol C. House, *Editors*

Committee on National Statistics
Division of Behavioral and Social Sciences and Education

Andy Peytchev
October 16, 1012

CHAPTER 6: PATHWAY

Recommendation 6-4: A broader set of nonexpenditure items on the CE that are synchronized with expenditures will greatly improve the quality of data for research purposes, as well as the range of important issues that can be investigated with the data. The BLS should pay close attention to these issues in the redesign of the survey.

CHAPTER 6: PATHWAY

Recommendation 6-5: A tablet computer should be utilized as a tool in supported self-administration.

However, a paper option should continue to be available for respondents who cannot or will not use a tablet computer. Visual design principles should be applied to redesigning the paper instrument in a way that improves the ease of self-administration and is aligned with the tablet modules.

CHAPTER 6: PATHWAY

Recommendation 6-6: BLS should develop a preliminary roadmap for redesign of the CE within six months. This preliminary roadmap would include a prioritization of the uses of the CE, an articulation of the basic CE design alternative that is envisioned with the redesign, and a listing of decision points and highest priority research efforts that would inform those decisions.

CHAPTER 6: PATHWAY

Recommendation 6-7: A critical element of any CE redesign should be the use of incentives. The incentive structure should be developed, and tested, based on careful consideration of the form, value, and frequency of incentives. Serious consideration should be given to the use of differential incentives based on different levels of burden and/or differential response propensities.

CHAPTER 6: PATHWAY

Recommendation 6-8: BLS should pursue a long-term research agenda that integrates new technology and administrative data sources as part of continuous process improvement. The introduction of these elements should create reductions in data collection and processing costs, measurement error, and/or the statistical variance and complexity of the CPI estimate. The agenda should address the robustness of new technology and a cost/quality/risk tradeoff of using external data.

CHAPTER 6: PATHWAY

Recommendation 6-9: BLS should increase the size and capability of its research staff to be able to effectively respond to changes in the contextual landscape for conducting national surveys and maintain (or improve) the quality of survey data and estimates. Of particular importance is to facilitate ongoing development of novel survey and statistical methods, to build the capacity for newer model-assisted and model-based estimation strategies required for today's more complex survey designs and non-sampling error problems, and to build better bridges between researchers, operations staff, and experts in other organizations that face similar problems.

Recommendation 6-10: BLS should seek to engage outside experts and organizations with experience in combining the development of tablet computer applications along with appropriate survey methods in developing such applications.

CHAPTER 6: PATHWAY

Recommendation 6-11: BLS should engage in a program of targeted research on the topics listed in this report that will inform the specific redesign of the CE.

CHAPTER 6: PATHWAY

Research Topics for Designs

- Use of a tablet device.
- How people keep financial records.
- Collecting data on a reduced set of 96 expenditure categories.
- Use of incentives.
- Instrument development.
- Privacy vs. open access.
- Evaluate potential impact from reducing proxy reporting of expenditures.
- Experiment with imputation methods and other statistical approaches.
- Evaluate the effectiveness of using more intensive methods.

Compatible Research

- Experiment with other technologies to record and extract data.
- Split questionnaire design.
- Evaluate the utility and the ability to obtain data from additional sources.
- Augmentation of sample with wealthy households.
- Identify and evaluate sources of auxiliary data (e.g., retailer data).

CHAPTER 6: PATHWAY

Research Topics for Specific Designs

Design A

Develop models that would estimate quarterly and annual expenditures and income at the household level from the four weeks of reported detailed data plus the data reported on larger and routine expenditures.

CHAPTER 6: PATHWAY

Research Topics for Specific Designs

Design B

- Investigate the assumption that a “bounding” interview is unnecessary to avoid telescoping and other issues.
- Investigate the accuracy and completeness of aggregated expenditures for periods up to six months and for estimates of averages (i.e., average monthly spending gasoline) used in this prototype to construct a full set of microdata for the entire six-month period.
- Develop appropriate models to “disaggregate” aggregated expenses using data from the one-week supported journal.
- Develop methodology for a successful component that will use an intensive interview and process based on prior collation of records and financial software to achieve a budget balance for the year at the household level as described below. Extend existing research done by Fricker, Kopp, and To (2011) to fully evaluate its potential and limitations.

CHAPTER 6: PATHWAY

Research Topics for Specific Designs

Design C

- Research and develop models for estimation using the base survey and two components of data collection.
- Research and develop models for imputing at the household level “smaller expense items” collected on the *Detailed Expenditure Component* and not on the *Household Profile Component* into the household-level dataset to complete the overall household expense profile.

CHAPTER 6: PATHWAY

Recommendation 6-12: BLS should fund a “methods panel” (a sample of at least 500 households) as part of the CE base, which can be used for continued testing of methods and technologies. Thus the CE is never again in the position of maintaining a static design with evidence of decreasing quality for 40 years.

Consumer Expenditure Survey Research & Plans for the Future

Adam Safir
Branch Chief
Research and Program Development
Consumer Expenditure Survey

October 16, 2012

BUREAU OF LABOR STATISTICS
U.S. DEPARTMENT OF LABOR

Presentation Overview

1. Background
2. Gemini Project Redesign Objectives
3. Focus of Recent CE Research
4. Gemini Project Design Team Plans
5. Use of CNSTAT's Report

BLS

Continuous Survey Improvement Process

- Biennial – CEQ instrument revisions
- 2013 – Inc Tax Estimation w/NBER Calculator
- 2009 – Telephone thresholds (CEQ)
- 2005 – Diary keying and auto-coding system
- 2005 – User friendly diary form
- 2005 – Contact History Instrument (CHI)
- 2004 – Income imputation
- 2004 – CAPI (CED) demographics and income
- 2003 – CAPI (CEQ)

BLS

Redesign Objectives

- Reduce measurement error, and in particular, underreporting
- Reduce burden
- Hold neutral, or reduce, costs
- Monitor redesign results

- Research agenda

BLS

Focus of Recent CE Research: Reduce Measurement Error

- Reduce number of interviews
- Reduce interview length
- Reduce reference period length
- Reduce proxy reporting
- Maximize record use & minimize recall
- Incorporate new technology
- Incorporate multi-mode interviewing

BLS

Focus of Recent CE Research: Reduce Burden

- Reduce number of interviews
- Reduce interview length
  - Streamlining the questionnaire
  - Split questionnaire research
  - Global questions
    - Diary to interview imputation
    - Backcasting
    - Within quarter imputation
    - Auxiliary source data

BLS

Focus of Recent CE Research: Reduce Costs

- Reduce interview length
- Reduce number of interviews
  - For example, eliminate the bounding interview

BLS

Focus of Recent CE Research: Monitor Redesign Results

Data Quality Profile

- A consistent, well-defined set of metrics can be used to establish baselines for monitoring trends in the quality of survey activities and evaluate the impact of survey design options

Measurement Error Analysis

- Determine sources of measurement error on expenditure reporting and develop a methodology for tracking and evaluating changes in measurement error due to design changes

Burden Index

- Measure the effect of alternative design options

In-Progress & Planned Research

In-Progress & FY2012

- Records Information & Feasibility
- Web Diary Test, Individual Diary Test

FY2013

- Redesign Monitoring
- Redesign Feasibility Testing

FY2014

- Additional Redesign Feasibility Testing
- Additional Analyses of Completed Studies

BLS

Gemini Project Design Team

- Initiated in July, 2012
- Joint team with the Bureau of the Census

Objective

- Create a single proposal for a redesigned CE that expects to result in a verifiable reduction in measurement error

Scope

- Synthesize inputs from 2009-12 Gemini Project activities, as well as from independent design proposals (Abt-SRBI, CNSTAT, University of Nebraska, University of Wisconsin, Westat)

BLS

Gemini Project Design Team: Timeline

- Jul 2012 – Kick-off
- Oct 2012 – Data collection recommendations, by quex group
- Nov 2012 – Redesign proposal outline (i.e., table shell)
- Nov 2012 – Decisions regarding independent design proposals
- Dec 2012 – Summary of FR, Census, BLS staff redesign ideas
- Dec 2012 – High level decisions re: survey redesign elements
- Feb 2013 – Present update at CE Annual Meeting
- Feb 2013 – Detailed overview of proposed redesign elements
- Mar 2013 – Draft redesign proposal report
- Jun 2013 – Final redesign proposal report

BLS

Gemini Project Design Team: Challenges

- Synthesizing research results into a comprehensive redesign plan
- Budget for research, testing, evaluation, and implementation
- Sample size requirements
- Timing of research findings
- Pace of social & technological change

BLS

CNSTAT Report: Value

High-level thinking, justification, and guidance from nationally-recognized experts in economics, statistics, and survey methodology on factors affecting the quality and cost of CE data, and improved options for collecting those data, culminating in:

- Concurrence on Issue Areas
- Broad Recommendations
- Specific Design Proposals
- Cost Estimates for Maintaining Budget Neutrality
- Advocacy for Additional Resources

BLS

CNSTAT Report: Promising Design Features

- One sample design
- Flexible recall periods & interview structure
- Increased use of technology, e.g., tablets
- Use of tech to encourage ‘in the moment’ reporting
- Increased reliance on self-administration
- Increased use of records
- Reduce proxy reporting
- Mixed mode data collection
- Large incentives
- Modular design, with a core survey

BLS

CNSTAT Report: Panel's Recommendations

6-01 Prioritize CE Data Uses for Redesign Trade-offs

6-02 Implement a Major Redesign ($)

6-03 Fund Several Major Feasibility Studies ($)

6-04 Sync Reference Periods for Exp & Non-Exp Items

6-05 Use Tablet for Self-Administration (w/Paper) ($)

6-06 Develop a Redesign Roadmap within 6 Months

6-07 Use Incentives ($)

6-08 Pursue a Long-Term Research Agenda

6-09 Increase Size & Capability of Research Staff ($)

6-10 Engage Outside Experts in App Dev ($)

6-11 Target Research on CNSTAT Recommended Topics

6-12 Fund a Methods Panel (or Research Sample) ($)

BLS

Thank You,
from CE to CNSTAT!

Contact Information

Adam Safir

safir.adam@bls.gov

BLS

Design A Evaluation

Sarah Nusser
Department of Statistics
Center for Survey Statistics and Methodology
Iowa State University

Questions

- What are the design's strengths and weaknesses?
- Which users is this design optimal for?
- What will the output of data collection look like?
- What statistical procedures will need to be applied?

Users for Design

- Target users: CPI
- Federal and state programs that use information at the level of the CPI or in aggregation
- Economic research, but not in the ideal form

Strengths / priorities unique to A

- Accurate reporting of concurrent expenses using records to the extent possible
  - Relatively short, but intensive collection to minimize respondent fatigue
  - Focus on relatively easy modes of recording items in natural units and in real time to minimize recall bias

Strengths / priorities unique to A

- Keep sample design and data collection simple
  - Focus on reducing measurement error and increasing response rates
  - Keep costs low
  - Simplify logistical operations of a new survey
  - Simplify post-data collection statistical processing

Questions

- Is 2 weeks too long for accurate reporting for all household members?
  - When does respondent fatigue start affecting quality?
  - What is the statistical gain of the 2nd week in a single reporting period?
  - Some models to consider: Dietary intake surveys, maybe some media surveys that require recording prospectively (radio listening survey)

Questions

- What does “recall” mean for the larger or recurring items?
  - Many of these items often have records associated with them
  - Develop methods to encourage record use here
- Is there sufficient sample size for special CPI calculations?
  - Survey has tons of data for modeling
  - Small area estimation might mitigate the need for direct estimates in smaller domains

Structural constraints

- Does not provide a direct and complete picture of household spending and income over the time periods of interest to economic researchers
  - With CPI as a goal, effort is devoted to expense detail, compromising time span of data collection
  - Statistical modeling can be used to bridge this gap
  - Survey has lots of data to support longitudinal modeling of aggregate categories of interest to researchers

Data for 2-wk data collection event

- Household or person-level
  - Demographics, life events (person)
  - Income, assets, labor force participation (person)
  - Recurring expenses (household, maybe person)
  - Larger purchases (household, maybe person)
- Expenditure-level by person
  - Individual purchases for items by person
- All expense and income data are in natural units for accurate reporting amounts

Design A 9

Data structure over time

Qtr	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	….
1	2 wks	2 wks							
1		2 wks	2 wks						
1			2 wks	2 wks					
2				2 wks	2 wks				
2					2 wks	2 wks			
2						2 wks	2 wks		
3							2 wks	2 wks	
…									…

Design A 10

Minimum statistical processing

- Translate the entered data into units that are comparable across individuals and households
  - Could be done by survey instrument
  - Could offer this as a check to respondents
- Diagnostics, editing for microdata
- Sample and nonresponse weighting, variance estimation variables
- Calculate CPI budget shares
  - Need a model to estimate quarter, annual params

Design A 11

Some additional opportunities

- More detailed nonresponse adjustment
- Small area estimation for specific CPI domains
- Measurement error modeling
  - Strong interest in seasonality → repeated observations
  - Evaluate quality of information (possibly for comparison within an experimental setting)
  - Potential bridge for uses beyond the CPI and state/federal programs

Design A 12

Example: dietary intake surveys

- Focus is on parameters from the "usual intake" distribution across individuals
  - Distn of indiv means: annual mean daily consumption of a dietary component for an indiv
- Food frequency: estimate of typical daily intake as a direct measure of usual intake
  - Shown to have considerable measurement error
- 24-hr dietary intake record or recall: more accurate information on a short period of time
  - Better quality, wrong concept
  - Can be related to usual intake via a model

Example: dietary intake surveys

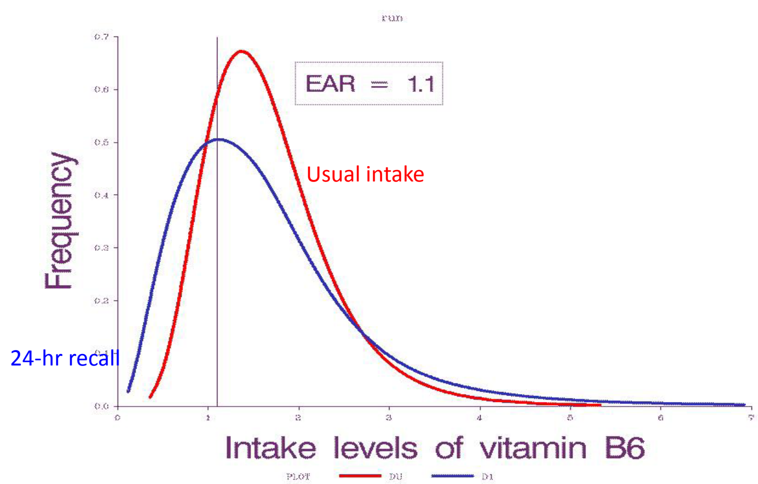
24hr recall = usual intake + error

$$D_{ij} = U_i + e_{ij}$$

D_{ij}	24 hr recall for indiv i on day j (NHANES: 2 days)
U_i	usual intake for indiv i (indiv mean over days)
e_{ij}	error in 24hr recall as an est of usual intake

- Interest is in parameters for distn of usual intake → estimate this distribution

Vitamin B_6 (mg/d) , women 19-50 yrs



Example: dietary intake surveys

- Lots of different nutrients, foods, etc.
  - Goal of methodology development was to serve the full range of behaviors
- Food intakes, for example, have many 0 values
  - Is this a structural 0 (not a consumer of the food) or a 0 on that day (consumer)
  - Mixture model to allow for a parameter that
- Ratios
  - Dietary component in relation to calorie intake
- Policy analysis
  - Food fortification alternatives

Talking Points: Michael Schober, October 16, 2012
Workshop Discussin of NRC Report.

SUPPORTED JOURNAL: a new mode. Not implemented in the same way in the three designs, and we as a panel were not able to fully hash out exactly what each design means by this. So let me give my take, and others can chime in, dispute, etc.

NOT simply a diary, and not simply self administered. Not quite an interview either. Promoting: Use of records. Done at the R's convenience!!! Potentially asynchronous--reduce time pressure so that we aren't sending the signal that estimation is desirable. Extra motivation, staff support.

All the motivational and cognitive issues raised this morning are serious, and need to be addressed. We believe they can't be addressed through minor revisions of current modes; it's not clear that we know that errors of reporting in current modes aren't systematic in a pernicious pattern.

The notion of a supported journal seems to us a reasonable compromise between the various legitimate needs at a price point that is plausible: large enough sample, more intensive data collection that supports record checking and responding in the ways and when convenient to the respondent but that doesn't require a too-expensive stable of interviewers or sample, potential for real-time data entry and monitoring and subsequent intervention, etc.

Premise: in order to get the higher quality record checking interviews we are looking for, there will need to be sufficient incentive for Rs to really be willing, to take the time, and feel this was worth it. The aim is to create a quite different experience than respondents now have.

I propose getting this right will require the insights of experience designers (not just HCI, human factors, or app programmers): what makes people want to, be willing, feel safe and non-suspicious, to provide their data.

A point of clarification on terminology:
The issue of recall is complex and we are not all using the term to mean the same thing. Need to make further distinctions. Supported journal allows for lots of different options:

Different dimensions/options for interviewing that can be deployed/implemented independently, and that can be mixed and matched in a supported journal:

- Retrospective vs concurrent reporting: this mode can be used for both.
- Use of records vs. memory. (Different from estimation vs. precise recall)
- Time pressure: synchronous vs. at R's pace.
- Self-administered vs. interviewer-administered -- we are blurring this distinction here.

Interviewer support could be face to face, via voice, via text, via automated help...

Different combinations of these options will have different cost structures, and presumably different effects on the respondent's experience, and ultimately the data quality.

We do not have good evidence on this for tablets. Initial evidence supports the plausibility that asynchronous reporting of numerical values on a mobile device via text (Schober, Conrad et al. 2012) can lead to less rounding/estimation and greater disclosure of sensitive behavioral information than synchronous reporting via voice, whether to an interviewer or automated interviewing system. But obviously this is not evidence on tablets with consumer expenditure questions in a CE sample....

Responsive monitoring: to set this up right will require thinking from a respondent-centered perspective!
--We need not simply a finger-wagging compliance center, but a truly supportive staff on every front--both responding to what R needs AND "pushing"--contacting R, motivating R, coming in person, etc.   Think of Apple tech support--easily scheduling a call for help.

Designing this right will require serious rethinking of how we are conceiving of interface or app design, to include the entire process and experience, facilitating getting records, making the respondent experience seamless and intuitive, based on how respondents (rather than researchers) are thinking of their expenditures (natural units). This also allows customization and alternate pathways for different respondents.

Staffing: Needs a new kind of field representative/responsive monitoring staff, whose job is to train Rs, collaboratively, on what we are asking of them, and on what kind of accuracy we need.

Measuring What We Spend:
Toward a New Consumer Expenditure Survey

The BLS
Consumer Expenditure Tablet

Presentation by
Mark M Pierzchala

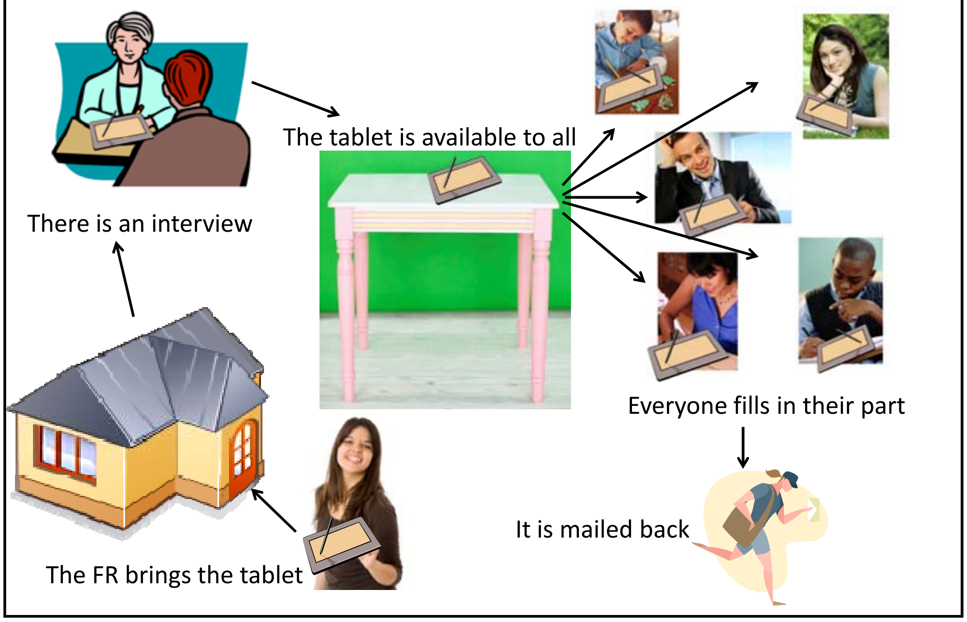
A BLS Tablet is

- An expenditure recorder
- A (self-)interviewer
- A communicator
- A case manager
- Secure
- Under BLS control
- A government device
- Light, fast, & spiffy
- Much bigger than a cell
- The same for all HHs

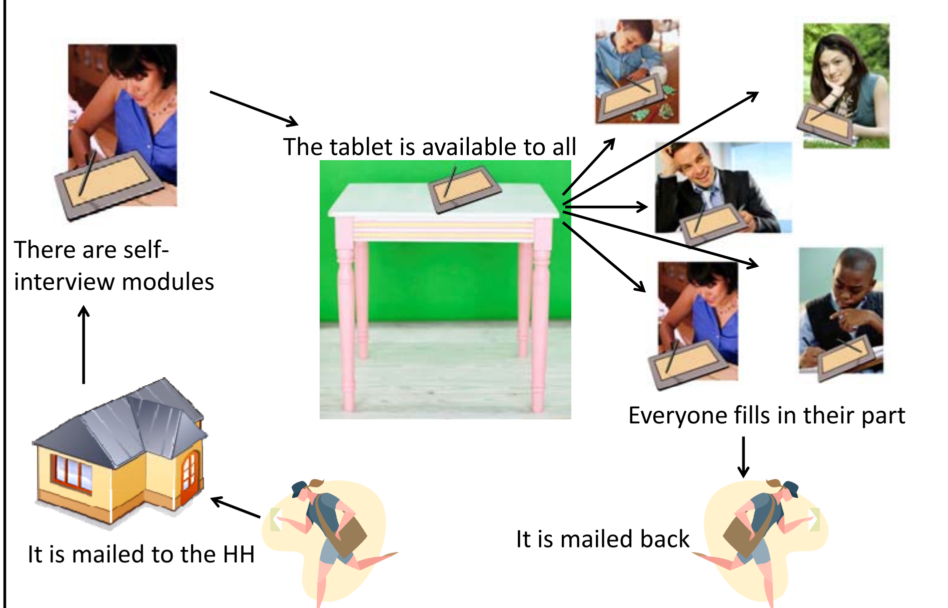
- Instant-on
- Private to the person
- On demand
- Easily used
- Supportable
- Available in the HH

= A compelling device

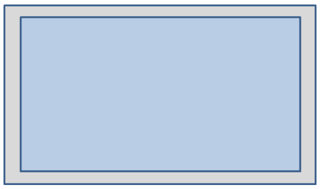
The Tablet in the Household (1)



The Tablet in the Household (2 .. n)

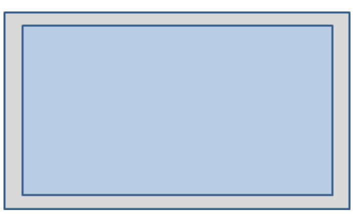


The Tablet as a Surveyor



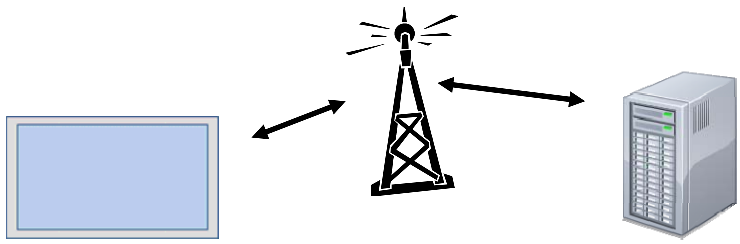
- Sets up the HH
- Combines the diary and the interview (the 'Journal')
  - Records expenditures
  - Collects related data
- Communicates with the home office
- Keeps things straight

The Tablet as a User Device



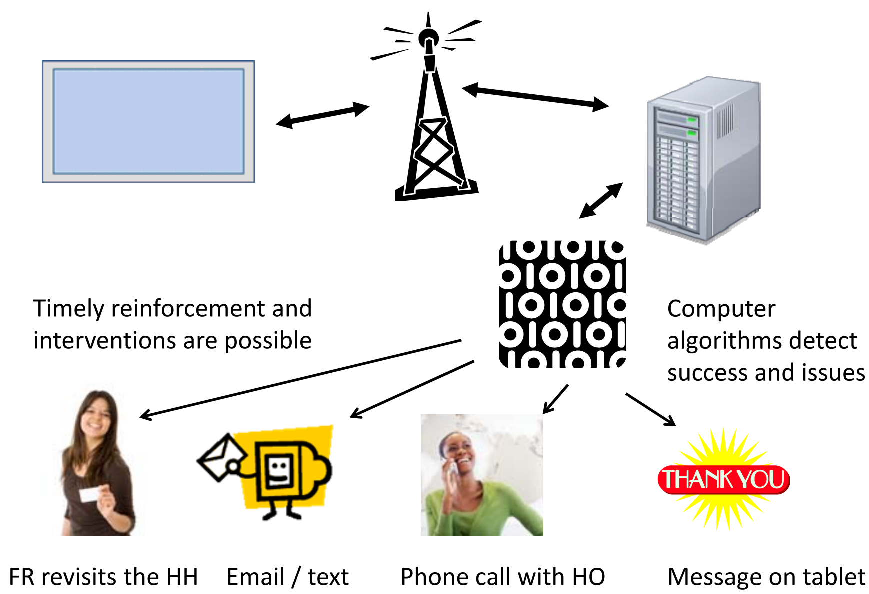
- Instant-on
- Fast, fast, fast . . .
- Easy, easy, easy . . .
- Self-evident
- Private individual entries
- The 'app' is pre-loaded
- The tablet is 'tied down'
- Can be used with or without a pen

The Tablet as a Communicator



- As data are entered, they are sent to the home-office server
  - Using phone-system transmission (no Internet access assumed)
  - Only data go back and forth
- No data are persisted on the tablet
  - During a session, data may be held in RAM on the table
- Data are encrypted and compressed
- The server has everything real-time
  - This enables powerful case-management possibilities

Case Management / Support



Technology: Further comments on the choice of tablets

Sarah Nusser
Department of Statistics
Center for Survey Statistics and Methodology
Iowa State University

Questions

- Why tablets?
- What other technologies were considered?
- Why not use the respondents' technology?
- Why focus on a device rather than an application?
- What is the impact of tablets on proxy reporting?

National Resources Inventory

- Remote sensing data collection
  - Every 5 years from 1982-1997
  - Annually 2000-2012 and forward
- Sources
  - Field visits → mylar overlays on borrowed 35 mm slides → contracted high quality aerial photos for sample segments
- Data collection tools
  - Paper forms → Newton handheld e-forms → integrated system with GIS tools for data collection and progress/quality monitoring
- Sample design
  - Simple repeated panel design → Panel observed every year + supplemental panels each year
  - Estimation changes accordingly
- Field studies: rugged handheld with GPS, iPads with GIS, GPS

Philosophical perspective

- A survey organization today must consider technology a moving target
  - We can't predict very accurately what's next
  - The leaps that are hardest to predict are paradigm shifts
  - Must constantly be working incrementally to advance systems that are used for production

Philosophical perspective

- A longitudinal survey needs a stable technology base and consistent software to support data collection
  - Leverage training that field staff and respondents receive for data collection to improve operational efficiency and data quality
  - Avoid changes in meaning of variables associated with new methods
  - Need to minimize risk of failure

How does this work?

- Develop a simple form of a system for initial production
- Evolve the system incrementally in ways that avoid issues with changing the data
- Anticipate the next change, do some research and prototyping, work out approach
- Design a study or overlap panel that can be used to evaluate and estimate any differences in old and new
- Use estimation to adjust as needed (avoid if possible)
- Full implementation of new system

Form factor

- Device choice is less about which device today vs 5 years from now, and more about whether the device will *support low burden, high data quality self-administered data collection, while containing data collection costs*

Form factor

- Key issue: need to facilitate easy entry of a multidimensional household data vector that gets entered by numerous people, whose reporting behavior varies in unpredictable ways
- Survey instrument software application will be complicated

Form factor

- Tablet form factor is far easier when the application gets complicated
  - Screen real estate
  - Ease of interacting with interface
  - Good models for creating usable software for complex applications

Why a dedicated device?

- Control over the appearance of the survey software
  - More consistency in data quality likely
- Coverage includes individuals who don't have a suitable personal device
- Lower cost
  - Simpler software and backend system
  - Simpler training for field staff
  - Cheaper support – 1 software, controlled device

Using a respondent’s device

- Smart phone, or more broadly, web-based survey instrument
- Developing and supporting software for a handful of platforms (device or web browser) is expensive
  - Software must work well in each unique condition requires → complex development
  - Lack of control over device and network issues → technical support is challenging
  - Lack of control over the consistency over display of survey instrument and its usability
  - Form factor aside, this would be costly and potentially reduce data quality

Using a respondent’s device

- Coverage
  - Do all members of the household have a suitable device?
  - Expect larger fraction of non-coverage relative to a dedicated device
- Smart phone
  - Application or web screen real estate is inadequate given complexity of survey instrument

Other technologies

- Scanners
  - Backend processing and quality checking would be expensive to get this to work for every situation
  - Should continue to be investigated
- Voice recognition
  - Not really considered
  - Would need to be investigated

Proxy reporting

- Design allows multiple respondents per household
- Anticipate better control of member's recording with a trained primary household respondent
- System offers a way to monitor compliance of members and provide feedback
- Fewer, but still some concerns about proxy reporting
  - Quality and cost improvements associated with a dedicated single device outweighed them
  - Warrants some investigation

Measuring What We Spend: Toward a New Consumer Expenditure Survey

Incentives Discussion

Panel on Redesigning the BLS Consumer Expenditure Surveys
Don A. Dillman and Carol C. House, *Editors*

Committee on National Statistics
Division of Behavioral and Social Sciences and Education

Andy Peytchev
October 16, 2012

Outline

- Short summary on incentives
- Three arbitrary examples to illustrate different incentive structure
- Three questions from BLS

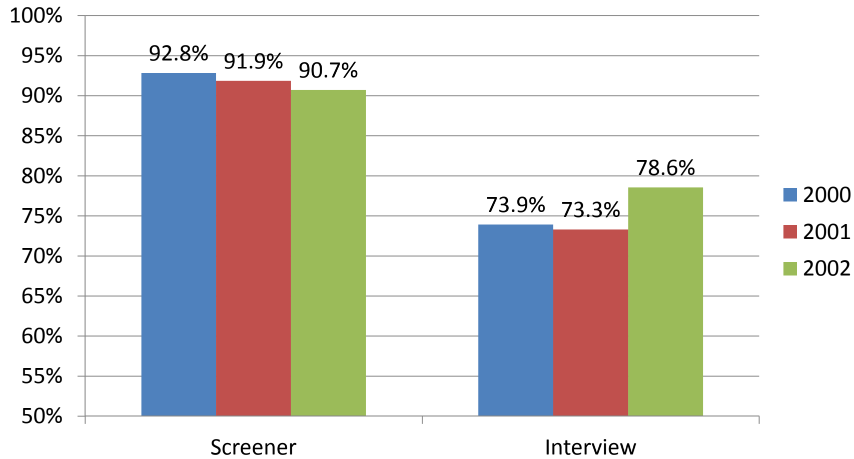
General Considerations for Incentives

- Cash
- Prepaid (not for diary completion)
- Most effective at the correct point in the survey process
- For surveys with longitudinal components, should be used throughout the process
- Differential incentives should be considered
- Larger (e.g., $100+) amounts should be considered, especially when intrusive or sustained
- Investigate use of intrinsic incentives

Example 1: Single Incentive to All Sample Members

- The National Survey of Drug Use and Health introduced a $30 promised incentive in 2002, for completing an interview with the selected adult
- Experimentally tested $0, $20, and $40
  - Substantially higher response rates, highest in $40
  - Lower cost per interview in both incentive conditions

NHSDA/NSDUH



Data source: Kennet et a., 2005

Example 2: Differential Incentives across Sample Members

- National Intimate Partner and Sexual Violence Survey
- Phase 1: $10 promised incentive
- Phase 2 (subsample of nonrespondents): $40 incentive

Example 3: Differential Incentives across Survey Components and Sample Members

- National Survey of Family Growth, cycle 6
- Phase 1 and 2:
  - Screener incentive
  - Interview incentive
- Phase 3:
  - Screener incentive
  - Interview incentive

The CE Redesign Options

- Recruitment and interview
- Self administration of the journal with detailed expenditure data, over a sustained period of time
- Repeated journal administration

Panel Recommendation (6-7)

A critical element of any CE redesign should be the use of incentives. The incentive structure should be developed, and tested, based on careful consideration of the form, value, and frequency of incentives. Serious consideration should be given to the use of differential incentives based on different levels of burden and/or differential response propensities.

Questions

- 1. How dependent are the designs on the incentives?
  - a. We are concerned about getting approval for such a large incentive. Does the panel feel that the designs would work without an incentive? With an approved incentive?
- 2. Did the panel consider the impact of the incentive on the spending behavior during the reference period?
- 3. Does the panel have any recommendations about how to test incentives? Can the amounts or procedures be tested independently of other design features, or would we need to test them in conjunction?

1. How dependent are the designs on the incentives?

- The panel finds incentives to be an essential component of the proposed designs
- This is not a typical survey interview task
  - Journal for continuous tracking of expenditures
  - Numerous expenditures being recorded on a daily basis

2. Did the panel consider the impact of the incentive on the spending behavior during the reference period?

- The impact on expenditures was discussed, and the form and timing of the incentives need to be carefully considered and tested

3. Does the panel have any recommendations about how to test incentives? Can the amounts or procedures be tested independently of other design features, or … in conjunction?

- The incentive structure, timing, and amounts should be tailored to the particular design

- The panel has provided some general differences in the incentive amounts and structure across the different design options

Survey Design Robust to Nonresponse Bias and Variance: a NSDUH Example