Managing (SIPP) with Adaptive and Responsive Design (Lessons learned from 2018/2019 SIPP data collection)

Jason Fields

Adaptive Survey Design Workshop
November 4-5, 2019

U.S. Census Bureau

This presentation is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

Outline

- Background about SIPP

- SIPP Typical Production Schedule

- Adaptive Survey Design (ASD) Timeline and Experiments

- 2018 SIPP (New sample, new strategy)

- 2019 SIPP (Many things - all at once)

- Challenges

- Lessons Learned (and still learning)

Survey of Income and Program Participation (SIPP)

- ## National longitudinal panel survey
  - 53,000 households in new 2018 Sample
  - Overlapping panel introduced in 2019

- ## Key Estimates
  - Labor Force Dynamics
  - Health Insurance Coverage
  - Temporary Assistance for Needy Families (TANF)
  - Women, Infants, and Children (WIC)
  - Income and Poverty
  - Wealth and Assets
  - Supplemental Security Income (SSI)
  - General Assistance (GA)
  - Supplemental Nutrition Assistance Program (SNAP)
  - Living Arrangements and Family Dynamics

Survey of Income and Program Participation (SIPP)

- Originally designed to compensate for the limitations of the Current Population Survey (CPS) and complement those estimates
    - CPS ASEC (March Supplement) uses an annual recall
    - Primarily focuses on status at interview and count of weeks worked

- Originally, SIPP was designed to have a shorter recall period
    - Staggered 4-month interviews
    - Overlapping panel design to boost cross-sectional estimates
    - SIPP focuses on dynamics, inter-relationship of topics, and components of income and assistance

- Offers the most detailed income and comprehensive program participation variables of the major nationally representative surveys

Survey of Income and Program Participation (SIPP)

- Organized around short longitudinal panels
    - First panel began in 1984

    - 1984-1993 were primarily 2½-year panels
    - 1996-2004 were 4-year panels
    - 2008 was extended to a 5th year
    - 2014 and later expected to continue to be 4-year panels

    - 4-month recall period / 3 interviews per year (1984 – 2008 panels)
    - 12- to 18-month recall / 1 interview per year (2014 to current)

    - 1984-1993 were overlapping panels
    - 1996-2014 are abutting 'end-to-end' panels
    - 2018/2019 resuming overlapping panels to ensure high quality calendar-year estimates

Survey of Income and Program Participation (SIPP)

- Adults (age 15+) interviewed in Wave 1
  - Data collected for all people
  - Proxy interviews for children under 15

- Need to ensure good representation of underrepresented and dynamic households
  - Renters and other mobile populations (marital transitions, multigenerational households, etc.)
  - Program participants
  - Poverty population (oversampled)

- Follows all Wave 1 interviewed adults in subsequent waves
  - Interview all household members at each address with original Wave 1 adult
  - Locating and interviewing movers is a significant effort

United States **Census** Bureau™

U.S. Department of Commerce
Economics and Statistics Administration
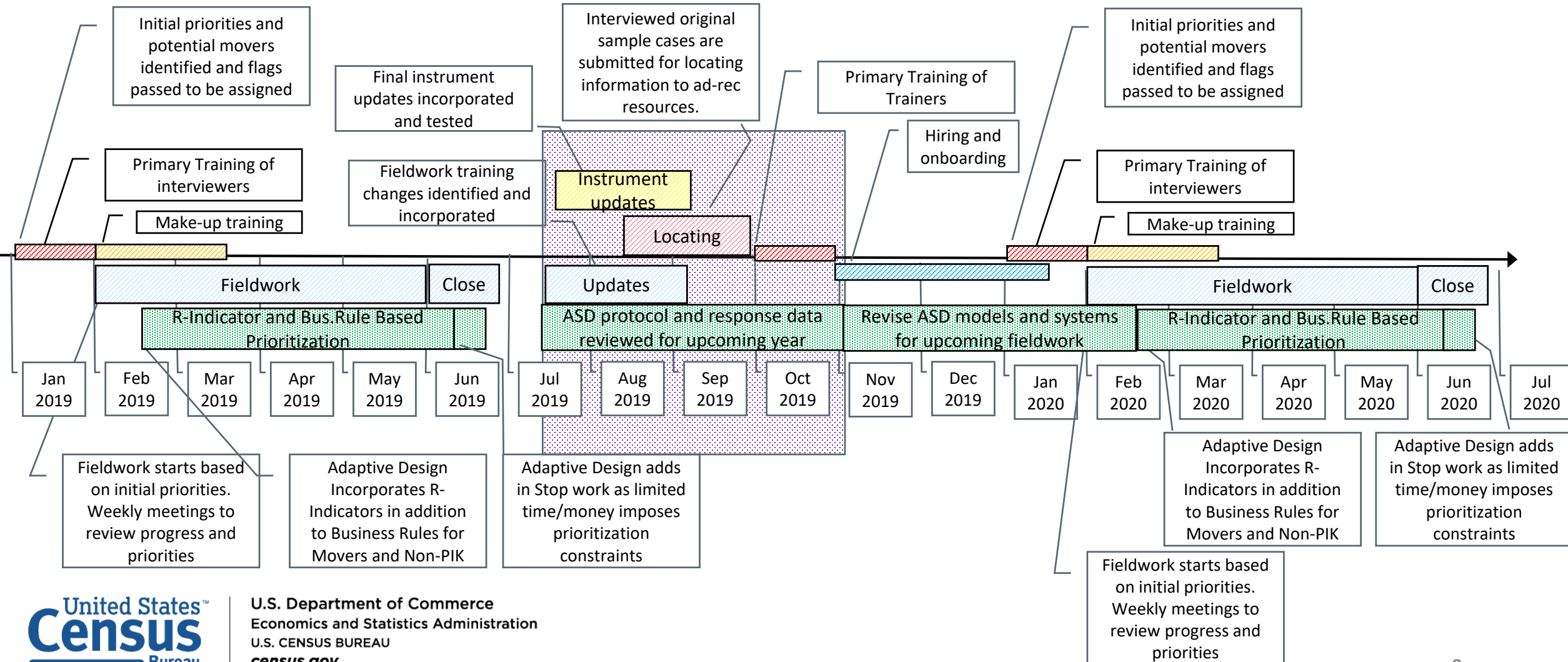U.S. CENSUS BUREAU
census.gov

Typical Production Cycle

Annual administration with 4-5 month collection window

- Annual hiring (starts in late summer and extends to interviewing)
- Full 2-day generic for new hires plus 4.5-day survey training for all interviewers (December-January)

File management

- Collection from prior year ends in June/July – collected sample reviewed in July and August
- Splits and movers identified and prepared for input to the next year
- Administrative data linkages occur in the summer
- Locating sample against administrative sources occurs in late summer – prospective addresses identified
- Adaptive survey design rules reviewed and initial prioritization schema set in late summer
- Sample delivered to field division in the early fall
- Collection and monitoring throughout February – July

- Data processing (in an ideal cycle)
  - Collected data begins review in July – August
  - Data processing utilizing administrative data informed imputation from August through February
  - Final public use preparation and release March-May

Typical Production Timeline

Initial priorities and potential movers identified and flags passed to be assigned

Final instrument updates incorporated and tested

Interviewed original sample cases are submitted for locating information to ad-rec resources.

Primary Training of Trainers

Initial priorities and potential movers identified and flags passed to be assigned

Primary Training of interviewers

Fieldwork training changes identified and incorporated

Hiring and onboarding

Primary Training of interviewers

Make-up training

Instrument updates

Make-up training

Locating

Fieldwork

Close

Updates

Fieldwork

Close

R-Indicator and Bus.Rule Based Prioritization

ASD protocol and response data reviewed for upcoming year

Revise ASD models and systems for upcoming fieldwork

R-Indicator and Bus.Rule Based Prioritization

Jan 2019 | Feb 2019 | Mar 2019 | Apr 2019 | May 2019 | Jun 2019 | Jul 2019 | Aug 2019 | Sep 2019 | Oct 2019 | Nov 2019 | Dec 2019 | Jan 2020 | Feb 2020 | Mar 2020 | Apr 2020 | May 2020 | Jun 2020 | Jul 2020

Fieldwork starts based on initial priorities. Weekly meetings to review progress and priorities

Adaptive Design Incorporates R-Indicators in addition to Business Rules for Movers and Non-PIK

Adaptive Design adds in Stop work as limited time/money imposes prioritization constraints

Adaptive Design Incorporates R-Indicators in addition to Business Rules for Movers and Non-PIK

Adaptive Design adds in Stop work as limited time/money imposes prioritization constraints

Fieldwork starts based on initial priorities. Weekly meetings to review progress and priorities
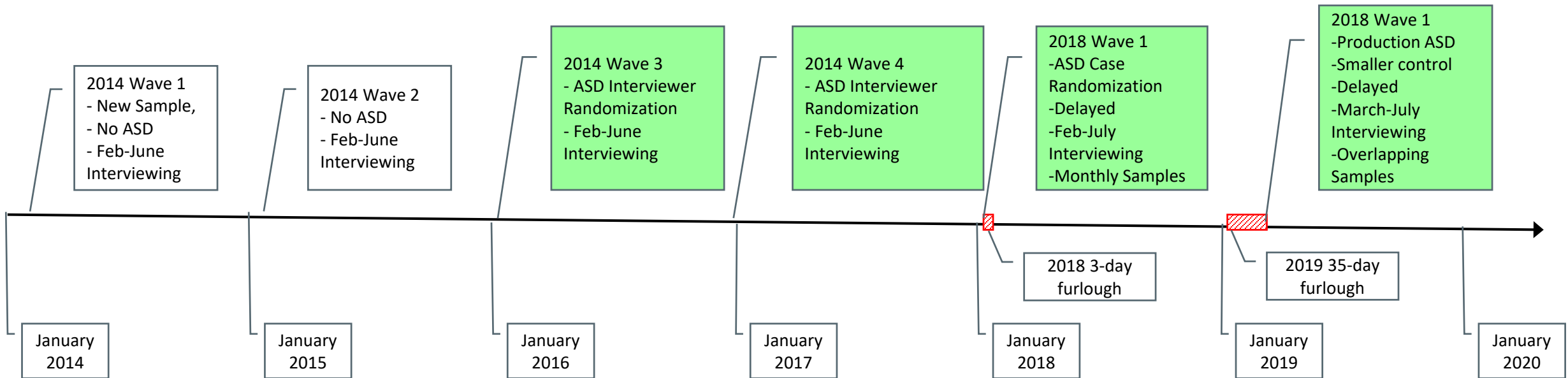
Motivation for Adaptive Survey Design

- Declining response rates

- Data quality as survey priority

- Difficulty hiring, on-boarding, training, and retaining field interviewers

- Competing priorities with multi-survey interviewers

- Need for agility based on challenges with the federal budget cycle

- Balancing cost, quality, and time

- Infrastructure in place to allow testing and production development

2014 – 2018 Adaptive Survey Design Timeline

2014 Wave 1
- New Sample,
- No ASD
- Feb-June Interviewing

2014 Wave 2
- No ASD
- Feb-June Interviewing

2014 Wave 3
- ASD Interviewer Randomization
- Feb-June Interviewing

2014 Wave 4
- ASD Interviewer Randomization
- Feb-June Interviewing

2018 Wave 1
- ASD Case Randomization
- Delayed
- Feb-July Interviewing
- Monthly Samples

2018 Wave 1
- Production ASD
- Smaller control
- Delayed
- March-July Interviewing
- Overlapping Samples

2018 3-day furlough

2019 35-day furlough

January 2014

January 2015

January 2016

January 2017

January 2018

January 2019

January 2020

2014 – 2018 Adaptive Survey Design

In 2016 and 2017 SIPP interviewers were randomly assigned to experimental treatment and control groups

- Some eligible to see actual priority status next to each case on their laptop
- Control interviewers only see medium priority for all cases
- Both 2016 and 2017 data collections consisted of all returning sample, initially interviewed in 2014. The goals of the ASD were to prioritize cases that
  - Will be representative of 2014 Wave 1 sampled households
  - Have moved at some point between interviews
  - Contain at least one household member who did not have a good link to administrative data and protected identity key (PIK)
- 2017 Tested controls against business rule only and business rule + r-indicator treatments

2014 – 2018 Adaptive Survey Design

In 2018 and 2019 all SIPP interviewers are eligible to see priority status next to their cases on their laptops

- Experimental design based on cases not interviewers

- With 2018 consisting of entirely of new sample and 2019 partially consisting of new sample, the goals of the ASD for new sample were
    - To obtain a large respondent pool that is representative of the frame for returning sample in future waves

- The goals of the ASD in 2019 for returning sample to obtain cases that
    - Will be representative of 2018 Wave 1
    - Have moved at some point between interviews
    - Contain at least one household member who did not have a good link to administrative data and protected identity key (PIK)

2014 – 2018 Adaptive Survey Design

- **High (H):** High priority cases should get first attention each day you work. Contact attempts should always be made within a week of a case being marked High Priority. You are encouraged to work High Priority cases as often as necessary to complete them quickly

- **Medium (M):** Work Medium Priority cases in the same manner as you usually would, completing each within a reasonable amount of time and achieving adequate progress throughout the interviewing period. Follow the interviewing procedures outlined in the training materials and official memoranda

- **Low (L):** These cases are usually "on temporary hold," meaning the sponsor does not want these cases to be worked in the field temporarily

2018 Adaptive Survey Design

Sampled five independent representative of about 10,600 cases across five different months instead of one continuous span of 53,000 cases

- Why:
  - Planned as a contingency for continuing resolutions, lapses, or insufficient funds
  - Would allow shorter collection period to have representative collected samples rather than easiest cases
  - Requested by field management to make SIPP consistent with other monthly samples and have monthly closeouts

Adaptive design plans

- Focus on developing initial priorities for the new sample – using the Planning Database (PDB)
- Having one month to obtain responses instead of four affects the ASD intervention schedule
- First week of each month targeted households in areas with mean number of people > 3
- Subsequent weeks used R-indicator and/or contact history response models to prioritize cases
- There were no low priority cases at any point of data collection

2018 Adaptive Survey Design

- Goal: to obtain a large respondent pool that is representative of the sample frame for returning sample in future waves

- Experimental Household level Block Design
  - Approximately one half of households received a response treatment
    - Targeted households in areas with mean number of persons per household > 3
    - Targeted households most likely to respond based on contact history after first 10 days of month
  - Approximately one half of households received a representative treatment
    - Targeted households in areas that were under-represented after first 10 days of month
  - Approximately one quarter of household had high priorities randomly assigned

- This experiment resulted in largely null findings
  - The response treatment resulted in less than a percentage point higher than their counterparts
  - The representative treatment resulted in a +0.02 difference in R-indicator

2018 Adaptive Survey Design

- This experiment was stopped after the third month, so the survey could focus on other production related challenges

- The fourth and fifth month only used R-indicators for prioritization. It did not prioritize cases according to the response treatment

- Approximately 8,000 cases that were not sufficiently worked in the earlier months, were given additional time and effort at the end of data collection

2019 Adaptive Survey Design

- Monthly samples discontinued after 2018
  - Successful at providing flexibility for funding uncertainty, but became operationally difficult
  - Field staff found monthly close-out difficult to complete reluctant or complex households
  - There were many cases that were not adequately worked at the end of each month

- In 2019, SIPP reintroduced an overlapping sample.
  - This is the first time case prioritization is in effect while there is both Wave 1 and Wave 2 sample
  - Target Wave 1 households that will yield a large respondent pool and reduce nonresponse bias for future waves
  - Target Wave 2 households that will combat attrition and attrition bias

- Adaptive survey design goals for 2019 – target cases that
  - Will benefit the most from early starts at the beginning of data collection
  - Are underworked and may contribute to lack of representation during middle of data collection
  - That will yield the highest possible quality at the end of data collection

2019 Adaptive Survey Design

- The 2016, 2017, and 2018 ASD's decisions were made at the macro-level, missing interviewer impacts at the micro-level
  - Each interviewer's percent of high priorities on workload were drastically different
    - Having too many or too few priorities on their workload makes it so that the interviewer isn't able to prioritize
    - Some interviewer's high priorities cases were scattered geographically, leading to inefficient routing; contradicting instructions given to them by their supervisors
    - Some interviewers have said that prioritization is difficult to manage

- There is no perfect solution for the problem of high priorities disproportionately affecting interviewers, for 2019 we established new rules for when a high priority case should be displayed
  - Cap the percent of displayed high priorities at 35 percent
  - Geographic focus
    - If percent of high priorities > 35 at start of data collection, then we find the geographic center of those high priority cases, making the furthest high priorities cases display at medium until enough of the other high priorities cases are worked sufficiently
  - Overwork threshold
    - If percent of high priorities > 35 past week 8 of data collection, then we make the high priority worked the most display as medium priority before any priority changes

- Assign only one new high priority per interviewer every two-weeks, beginning week 8
  - Choose new high priority case based on R-indicator and level of effort exerted

United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

2019 Adaptive Survey Design

- Furlough
  - Lost the month before interviewing
    - Prospective interviewers in the queue didn't all wait
    - All trainings in January were lost
    - Month delay in start-work with very high count of cases unreleased due to no staff

- Clearances for interviewers
  - Added to the delays staffing, impacted all surveys
  - Bureau-wide Lack of on boarded interviewer's in 2019, there were an unprecedented number of cases per interviewer and geography with no staff available

- Making the best of the situation
  - Beginning in 7/1/2019, focus was directed to returning sample rather than new sample
  - Maintain the investment in the cases with multiple waves of data

- Adaptive design allowed focused decisions
  - Propensity to respond
  - Contact attempts
  - R-Indicator
  - Prioritize based on geographic coverage

Challenges in general

- Difficult hiring environment (yes – always a problem in some areas)
- Clearances and delays in onboarding
- Interviewers often work multiple Census Bureau surveys
- Furlough and continuing resolutions
- Budget level uncertainty
- Respondents increasingly reluctant, distrustful
- Response rates continue to slip even before respondents experience the survey
- SIPP remains a long, hard, longitudinal survey
- Survey design doesn't currently allow year-round interviewing

United States™ **Census** Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Challenges for Adaptive Survey Design

- Communicating and monitoring compliance or issues with decentralized workforce

- Tailoring for the realities of interviewer's work

- Need enough time to direct work and realize impact

- Integrated information inputs needed

- Metrics to prioritize and generate improved quality need to be continuously developed

- Adapting to externalities

Lessons Learned

- Communicating and monitoring compliance or issues with decentralized workforce
  - Adjust communication
    - As opposed to a single memorandum providing interviewers with priority instructions. The priority protocols are now featured in their onboarding trainings
    - Additionally, a secondary memorandum was sent to all interviewers trying to instill intrinsic value for the case prioritization methodology
    - Text messages are sent to interviewers phones to either remind them that priority status has changed or to alert them with feedback on how they are did the following intervention period

- Adaptive Design moves the quality needle slowly, but has other significant value as a tool for survey management
  - Redirect work or stop work based on data and management priorities

- Monitoring with control cases randomly assigned to priority is one way to maintain informed decision-making
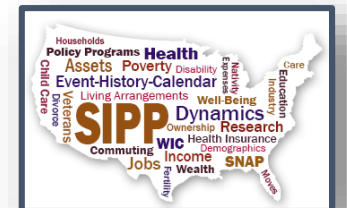
Lessons Learned

- Case level prioritization was a necessary improvement, but workload level management is also necessary
  - The nature of sample coverage and non-response leads to concentrations of priorities
  - When every case is prioritized, there is no advantage

- Forced to make hard choices, continue to use data quality improvement as the primary driver when possible

- The intersection of the survey questionnaire and field procedures requires tools like adaptive design to provide flexibility in the face of operational challenges

- In 2018 and 2019 the budget and operational situation was an exceptional set of circumstances, adaptive design investment and experimentation during the prior years was certainly valuable

United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Thank You!!

Jason Fields

Jason.M.Fields@Census.Gov

Combatting Non-response Bias for Wave 1 Sample

NR Variable from SIPP NR Report	Levels	Variable Used in ASD	Levels
Gender of Householder	Male/Female	Percent of Fem HH with No Hus*	> 20%, ≤ 20%
Num. of HH Member	1,2,3-4,5+	Mean Num. of HH Members*	< 1, 2-4, >4
Race of Householder	Black, non-black	Percent of Household Black*	>25% Black, ≤ 25% Black
Tenure	Owner, Renter	Percent of Households Rented*	> 75%, (25-75), ≤ 25% Rent
Urban/Rural	Urban, Rural	(Urban/Rural)	Urban, Rural
Region	MidWest, North, South, West	Region	MidWest, North, South, West
CBSA Area	Central city, MSA but not central city, MSA not census place Not census place	CBSA Area	Central city, MSA but not central city, MSA not census place Not census place
Within PSU Strata	Low-income High-income	Percent of Households in Poverty	> 20% Poverty, ≤ 20% Poverty

*-PDB Substitute

Overlapping Intervention Schedule

Month	Wave 1	Wave 2+
1	H – Households in areas with mean number of people greater than +3.5, Households estimated to need an early start M- Otherwise	H – Likely Movers, Non-PIKs, Movers Identified M- Otherwise
2	H – Households in areas with mean number of people greater than +3.5, Households estimated to need an early start M- Otherwise	H – Likely Movers, Non-PIKs, Movers Identified, Under-represented/Likely Respondent L – Over-represented/Unlikely Respondent M – Otherwise
3	H –Under-represented/Likely Respondent L – Over-represented/Unlikely Respondent M – Otherwise	H –Non-PIKs, Movers Identified, Under-represented/Likely Respondent L – Over-represented/Unlikely Respondent M – Otherwise
4	H – Under-represented/Likely Respondent L – Over-represented/Unlikely Respondent M – Otherwise	H –Non-PIKs, Movers Identified, Under-represented/Likely Respondent L – Over-represented/Unlikely Respondent S – very unlikely to respond M – Otherwise