

Adaptive Curtailment of Survey Followup Based on Contact History Data

by Eric Slud^{1,2} and Chandra Erdman¹

¹US Census Bureau

¹Center for Statistical Research & Methodology

²Mathematics Department, University of Maryland

Disclaimer

This paper describes research and analysis of the authors, and is released to inform interested parties and encourage discussion. Results and conclusions are the authors' and have not been endorsed by the Census Bureau.

4-Step Strategy for Adaptive Analysis Applied to American Community Survey

- (1) Partition population into groups by cumulative counts of contact-history events to separate interview rates.
 - (2) Use each group's contact history before each contact attempt to calculate time-specific interview rate curves.
 - (3) Define respondent burden increments over time in terms of the same groups.
-
- (4) Analyze alternative policies defined by restricting contact-history variables by recharacterizing lost interviews.

American Community Survey Background

- Households sampled in monthly panels, followed for 2 months
- 1st month: for mailable addresses, multiple mailings and self-response
- 2nd month: for addresses with telephone #'s, multiple call attempts from CATI call centers
- 3rd month, CAPI: subsample unmailable addresses and CATI nonresponders for personal visits

ACS Data, June 2011-Feb.2012 Panels

<u>1,079,985 HU's</u>	<u>total CATI and/or CAPI</u>
600,203	CATI-only
284,887	CATI-eligible
307,054	CAPI-only
274,772	CAPI-eligible
190,728	both CATI & CAPI
183,399	non-Type C

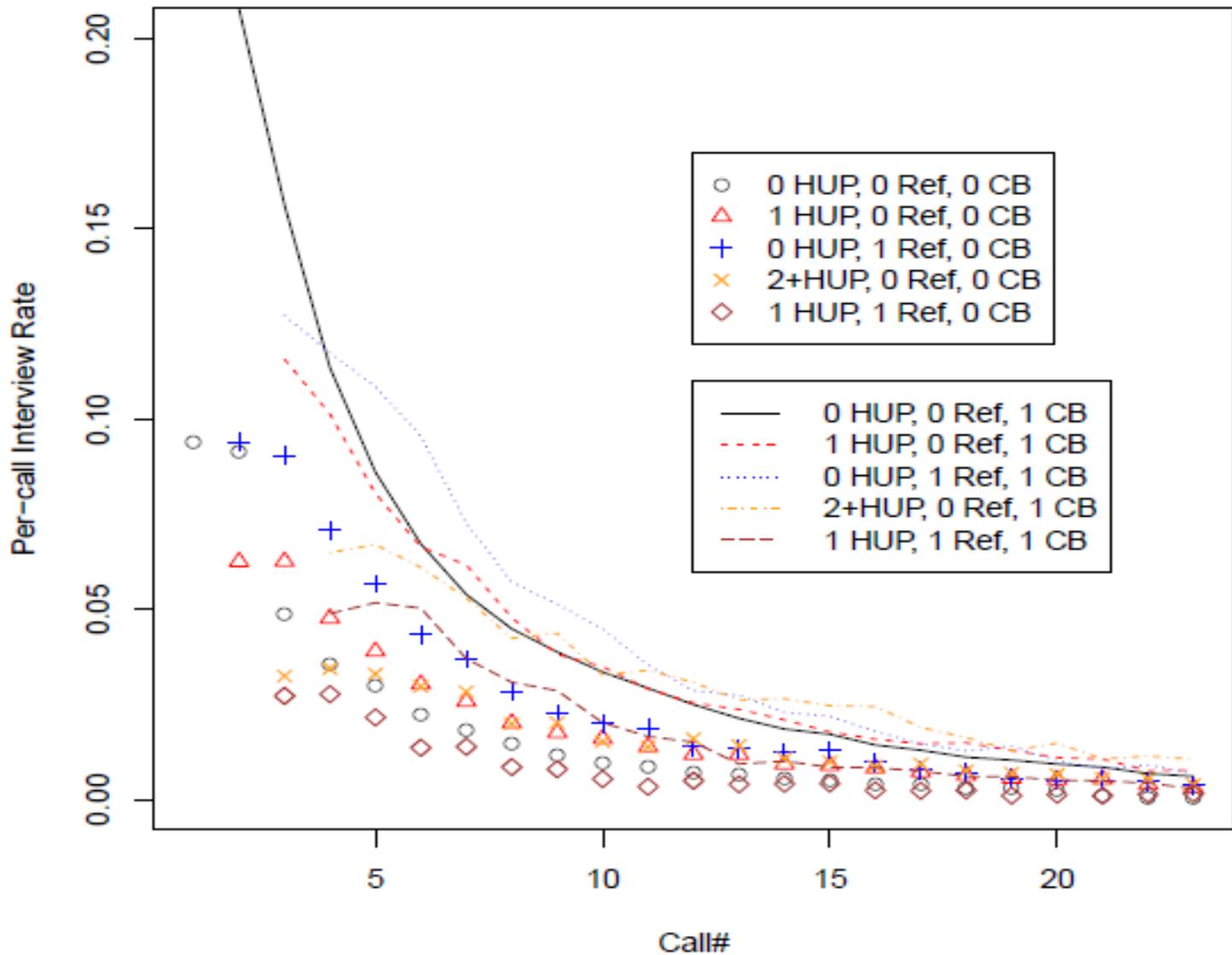
CATI Methodology

- i. Partition cases into disjoint groups based on CATI history of Refusals (REF), Hangups (HU), and Callback requests (CB)
- ii. Calculate interview yield over time by group
- iii. Define perceived Burden as a cost element, increasing at different rate with changing cumulative (REF, HU, CB) counts.

Methodology, cont'd

- Yield per contact-attempt is calculated as a function over t (= call attempt) of fractions
(# interviews at t) / (# eligible & *at risk* at t)
discrete hazard function
- An adaptive approach to scheduling would look for population subgroups for which this yield curve is particularly high or particularly low.

Interview Rate in CATI, by Group Points no-Callback, Lines with-Callback



Additional Steps in CATI

- Define Respondent Burden increments for each call t by call history (REF, HU, CB)
- Griffin & Hughes (2013) simplified this, calculated Burden as total calls and total post-reluctance calls
- Compared 14 alternative policies curtailing followup by total calls, total unanswered, and (REFusal, HangUp) counts

Impact of CATI results

2012 CATI parameters allowed: 25 total calls,
20 unanswered calls, 2 refusals & 3 hang-ups

- Tested 14 alternative control policies varying these parameters.
- Changed in April 2013 to curtail followup using maximum parameters of : 15 total, 12 unanswered, 2 Refusals, 2 Hangups.
- Monitored performance since then (Griffin 2013): savings and response rate projections realized.

Steps Implemented in CAPI

- Partitioned CAPI population by counts of interviewer coded noninterview types on the CHI instrument
 - Hard refusals
 - Reluctance (Noncoded refusals)
 - categorized Soft refusals
- Formed 5 groups to separate overall interview rates; separated further by being in CATI (including Ineligibles) vs. CAPI-only
- Discrete hazard-rate for interviews in time-varying groups plotted as function of visit t

CAPI Groups Defined using CHI and in-CATI Variables

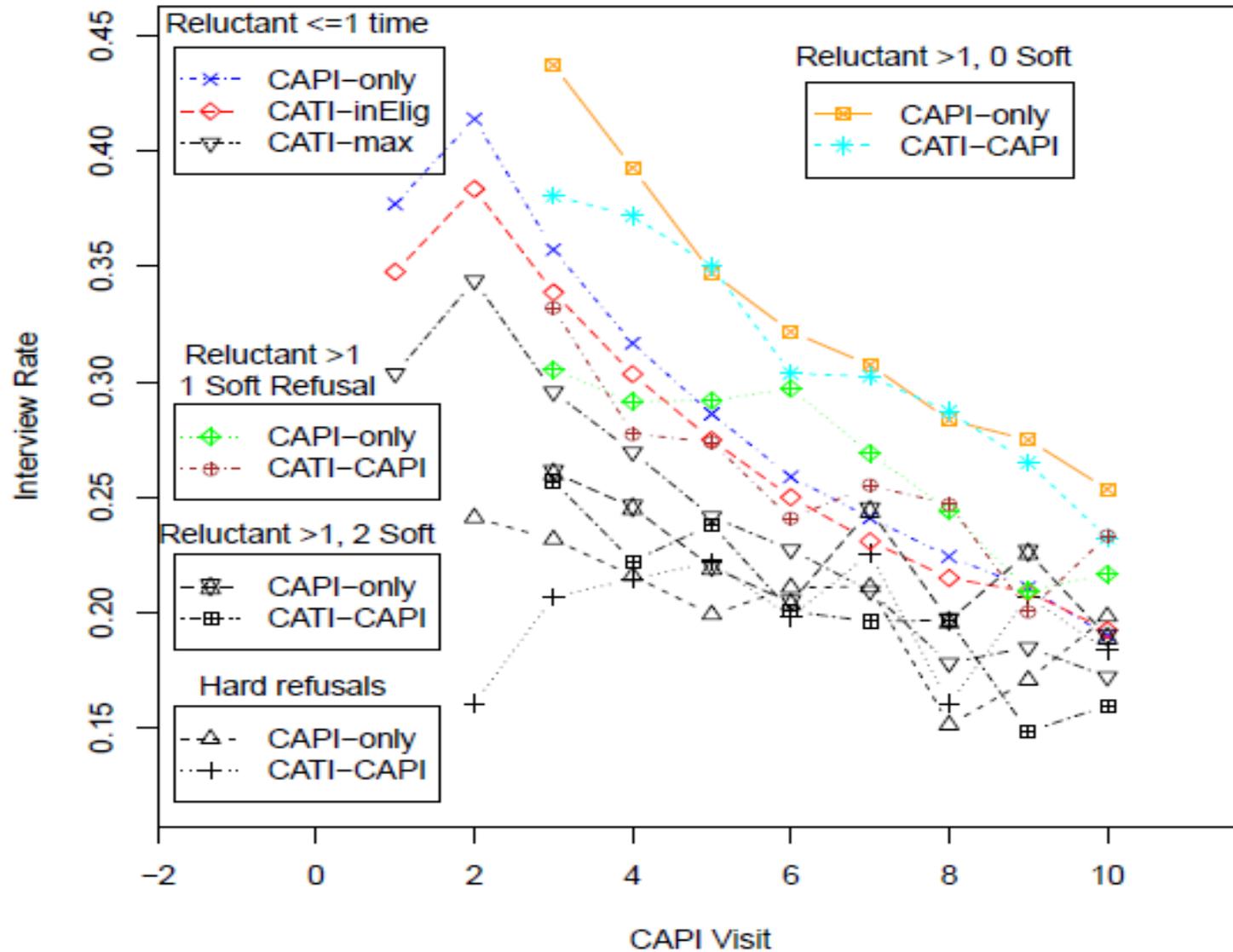
CAPI HISTORY	# of CAPI INTERVIEWS	Total ELIGIBLE CAPI CASES	CAPI Interview RATE
HARD REFUSALS --- NOT IN CATI	1,162	4,090	0.284
--- IN CATI	1,134	4,189	0.271
RELUCTANT CASES			
0 or 1 Visits coded RELUCTANT --- NOT IN CATI	248,584	255,693	0.972
in CATI – but INELIGIBLE for CATI	126,951	130,728	0.971
in CATI – MAX REF’s or Calls	33,654	35,231	0.955
2+ RELUCTANT Visits, 0 SOFT REFUSALS - NOT IN CATI	9,427	9,898	0.952
-- in CATI	7,979	8,388	0.951
2+ RELUCTANT Visits, 1 SOFT REFUSAL - NOT IN CATI	2,787	3,151	0.884
--- in CATI	2,603	3,032	0.859
2+ RELUCTANT Visits, 2+ SOFT REFUSALS- NOT IN CATI	1,597	1,940	0.823
--- in CATI	1,463	1,831	0.799
TOTAL for CAPI	437,611	458,171	0.955

Remarks on CAPI Groups

- Groups apart from cases Reluctant at 0-1 visits are small, less than 10% of the CAPI total
- Rates do not separate dramatically (except for Hard refusals), but are consistent in trend
- Similar CAPI-only and CATI-CAPI group sizes, with similar rates except that CAPI cases with Soft Refusals which were in CATI have slightly smaller rates

But the time-dynamic rate curves are fairly consistent over time – leaving room for group-specific adaptive designs to make a difference !

Visit-Specific CAPI Interview Completion Rate by Groups Defined Dynamically by History Paradata



Discussion

- Methods here primarily data-analytic
- “*Model-free*” time-varying cell-based response propensity modeling in terms of Contact History Variables
- Time-specific interview rates and increments in perceived burden are linked, via dynamic groups
- Adaptive design in this setting consists of choice of group-specific optimal control parameters including targeting or curtailment of followup.

References

- C. Erdman & D. Morris (2013) NHIS propensity model
- D. Griffin & T. Hughes (2013) – ACS Memo series
- D. Griffin (2013) – ACS Memo series
- R. Groves (2012 JOS & 2013 POQ) papers on [response propensity models with time-varying covariates](#)

For time-dynamic models, `discrete hazards' as in many biostatistical papers and in textbook

Klein, J. and Moeschberger, M. (2003) **Survival Analysis.**

Thank you !

Contact info

Eric.V.Slud@census.gov

Chandra.Erdman@census.gov