

OBJECTIVES

- › **Target** nonresponding cases after an initial period of data collection.
- › **Prioritize** targeted cases in the interviewers' workload.
- › **Aim** to
 - improve response rates,
 - reduce nonresponse bias,
 - minimize variations in response rates among subgroups,
 - not increase the cost.

RESEARCH QUESTIONS

How to effectively target cases?

- › Tourangeau et al. (2017): the expected value of a case, $V_i = \hat{\rho}_i W_i \Delta_i$
 - $\hat{\rho}_i$: estimated response propensity
 - W_i : design weight
 - Δ_i : the effect on the sample balance
- › Can this be applied to a panel survey, in which cases were
 - recruited at different times, and
 - designed with differential probabilities of selection

CONSIDERATIONS IN PRACTICE

$\hat{\rho}_i$: estimated response propensity

- › $\hat{\rho}_i$ for sample recruited in year 1 is greater than $\hat{\rho}_i$ for sample recruited in year 1+x.
- › Need to maintain the share of samples recruited at different years.

W_i : design weight

- › W_i reflects sample design
- › W_i for oversampled subgroups is smaller than W_i for the general group.
- › Low W_i is also important

Δ_i : the effect on the sample balance

- › Δ_i can be measured by sample imbalance, the distance away from a balanced set.
- › A balanced set is constructed by a set of variables reducing nonresponse bias.

A method to consider $\hat{\rho}_i$, W_i , Δ_i altogether

- › Well controlled for all three factors
- › No cancel-out effect

METHODS

A repeated stratified selection, allocated by sample imbalance

- ...balancing refers to a second phase of a selection, the first phase being the drawing of a sample from the population. ••
— Särndal (2011)

Step 1. Construct fixed strata

- › Variables for consideration
 - sampling domains
 - variables known for reducing nonresponse bias
 - analysis variables of interest

Step 2. Allocate targeted%, dynamically

- During data collection, at each run time,
 - › Overall targeted% is fixed
 - › Sample imbalance for each stratum, $d_{stratum j}$, changes over time
 - › Targeted% for each stratum varies by $d_{stratum j}$

Step 3. Select high $\hat{\rho}_i$ cases, dynamically

- During data collection, at each run time,
 - › Predict response propensity for case i , $\hat{\rho}_i$
 - › Sort cases by $\hat{\rho}_i$ in each stratum
 - › Select the top cases till targeted% is reached

APPLICATION (A Westat conducted panel survey)

Step 1. Construct fixed strata

- > Form initial strata by cross-classifications of the variables constitute
 - sampling domains
 - weighting cells for nonresponse adjustment in the prior round
- > Collapse initial strata with small stratum size
- > Stratum size changes at each time

Stratum j	Stratum size
j=1	
j=2	
...	

Step 2. Allocate targeted%, dynamically

During data collection, at each run time,

- > Calculate sample imbalance for stratum j, $d_{stratum j}$,

$$d_{stratum j} = 1 - \left(\frac{\# \text{ complete to date, stratum } j}{\text{panel sample size, stratum } j} \right)^2$$

- > Adjust targeted% by $d_{stratum j}$
- > The larger $d_{stratum j}$ stratum gets larger targeted% allocated

Stratum j	Stratum size	$d_{stratum j}$	targeted%
j=1		↑	↑
j=2		↓	↓
...			
Fixed overall targeted			15%

Step 3. Select high $\hat{\rho}_i$ cases, dynamically

During data collection, at each run time,

- > Predict response propensity for case i, $\hat{\rho}_i$:

$$\hat{\rho}_i = \frac{1}{1 + e^{-(\beta_0 + \sum_i \beta_i \times X_i)}}$$

- > β_i were trained by logistic regression with historical data

X_i	
level of field effort to date	time into data collection
	# contact attempts
	# appointments made
	# refusals
	# days since 1st contact attempt
current response status	
prior round response status	

RESULTS

- > Comparisons between groups: targeted vs. not targeted

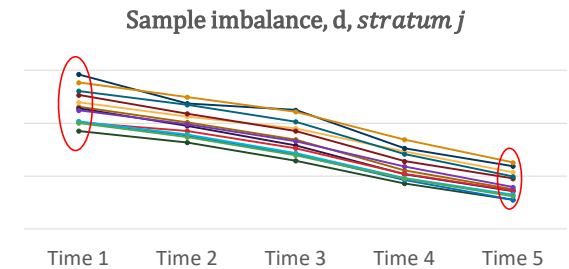
- 26 strata were formed in total

Number of strata, in which	
response rate of targeted > not targeted	20
contact attempts made for targeted > not targeted	23

- Overall response rate difference: targeted higher than not targeted

Response rate difference, Targeted - Not targeted	95% CI
11%	[7%, 15%]

- > Dispersion among $d_{stratum j}$ decreased over time, due to dynamic allocation



26 operational strata were combined into 13 analysis strata.

Special thanks to Pengyu Huang, Westat, who conducted data analysis.