A stableness of resistance model for nonresponse adjustment with callback data

Wang Miao

Department of Probability and Statistics
Peking University

mwfy@pku.edu.cn

May 26, 2022

https://arxiv.org/abs/2112.02822

https://www.math.pku.edu.cn/teachers/mwfy

Joint work with Xinyu Li and Baoluo Sun.

Outline

1. Paradata, callback data, and missing data
2. Identification with callback data
3. Semiparametric inference
4. Simulations and applications to CES, Card data, and ANES NRFU
5. Discussion

Outline

Paradata

Paradata are the records tracking the collection process of survey data (Couper, 1998).

Table 1: Data structure of a sampling survey

Frame/Questionnaire data				Paradata
ID	X_1	$\cdots$	X_p	Process of data collection
1	x_{11}	$\cdots$	x_{1p}	date, time, length,
2	x_{21}	$\cdots$	x_{2p}	mode of communication,
:	:	:	:	attitude of the respondent,
:	:	:	:	record of contacts,
n	x_{n1}	$\cdots$	x_{np}	gender of interviewer
Survey data				

Paradata are widely available in modern surveys, e.g., U.S. National Health Interview Survey (NHIS), British Survey of Social Attitudes (BSSA), and the European Social Survey (ESS).

Missing data and callback data

Callback data: a traditional form of paradata.

The frame data are often prone to missingness.

In many surveys interviewers continue to contact nonrespondents and the contact attempts are recorded. Sometimes called level-of-effort data (Biemer et al., 2013).

Table 2: Data structure of a sampling survey with callbacks

Frame/Questionnaire data				Contact attempts			
ID	X	Y	R	R_1	R_2	...	R_K
1	x_1	y_1	1	1	1	$\cdots$	1
2	x_2	y_2	1	0	1	$\cdots$	1
:	:	NA	0	0	0	$\cdots$	0
:	:	NA	0	0	0	$\cdots$	0
n	x_n	y_n	1	0	0	$\cdots$	1

Use of callback data in social sciences

Appropriate use of callback data can improve survey quality and compensate for deficiencies in surveys.

- Callbacks have routinely been used to monitor response rates and to study how design features affect contact and cooperation in the course of data collection (Bates, 2003; Groves & Couper, 1998); e.g., calls made during weekday evenings and on weekends are more likely to be responded.
- Kreuter (2013) provides a comprehensive literature review on the use of paradata in analyses of survey data.
- Olson (2013) reviewed categories of paradata and challenges and opportunities in using paradata for nonresponse adjustment.

Missing data analysis

Missingness mechanisms (Rubin, 1976; Little & Rubin, 2002)

- Missing at random (MAR) $R \perp\!\!\!\perp Y \mid X$;
- Missing not at random (MNAR) $R \not\perp\!\!\!\perp Y \mid X$;

A large body of the missing data analysis literature rests on missingness at random (MAR) or ignorability. Parametric approaches: Likelihood-based inference; Imputation. Semiparametric approaches: Regression-based estimation (REG); Inverse probability weighting (IPW); Doubly robust estimation (DR).

The biggest challenge for MNAR is identification: the joint distribution is not uniquely determined from the observed data distribution.

Missing data analysis

Strategies to achieve identification:

- restrictive parametric models, e.g., Heckman (1979)'s selection model; counterexamples, the normal-logistic model (Wang et al., 2014; Miao et al., 2016);
- instrumental variable approach (Manski, 1985; Newey, 2009; Das et al., 2003; Sun et al., 2018; Liu et al., 2020; Tchetgen Tchetgen & Wirth, 2017),
- shadow variable approach (Miao & Tchetgen Tchetgen, 2016; D'Haultfœuille, 2010; Wang et al., 2014; Zhao & Shao, 2015; Kott, 2014).

- Sensitivity analysis (Robins et al., 2000) and graph models (Fay, 1986; Sadinle & Reiter, 2017; Malinsky et al., 2020).

Missing data analysis

Researchers have traditionally sought auxiliary variables from the sampling frame, but it is surprisingly difficult for practitioners and the difficulty is amplified in multipurpose studies where multiple survey variables are concerned and multiple auxiliary variables are necessitated.

Moreover, instrumental and shadow variable approaches invoke additional no interaction or completeness conditions that further limit their use.

Lastly, they break down if the auxiliary variables also have missing values, e.g., due to failure of contact in surveys.

In contrast, callback data offer an important source of auxiliary information for nonresponse adjustment. Follow ups are commonly made in many surveys to increase the response rate, the contact process are recorded by interviewers and are often kept for all units.

Using callback data for missing data analysis

Callback data had not been used widely in statistical analysis until recently.

- The early idea of Politz & Simmons (1949): use the number of nights that a respondent had been at home during the past week to account for the "not at homes"by weighting;
- The "continuum of resistance" model: nonrespondents are more similar to delayed respondents than they are to early respondents, so that the most reluctant respondents can be used to approximate the nonrespondents. (Lin & Schaeffer, 1995; Groves & Couper, 1998; Kreuter et al., 2010; Little, 1982).
- Model the joint likelihood of the callbacks and the frame variables; require untenable assumptions to achieve identification, e.g., assume the response probabilities are equal across different attempts or levels of the frame variables (Biemer et al., 2013; Drew & Fuller, 1980).
- Chen et al. (2018); Zhang et al. (2018) generalize the Heckman (1979) Selection model to incorporate callbacks to improve efficiency.
- Daniels et al. (2015) advocate the use of pattern mixture models for sensitivity analysis.

Using callback data for missing data analysis

Most notably, Alho (1990); Kim & Im (2014); Qin & Follmann (2014); Guan et al. (2018) employ propensity score models to make nonresponse adjustment with callbacks and propose inverse probability weighted and empirical likelihood-based estimators.

Their model:

$$\text{logit } f(R_k = 1 \mid R_{k-1} = 0, X, Y) = \alpha_{k0} + \alpha_{k1} X + \gamma_k Y$$
$$\text{with } \alpha_{k1} = \alpha_1, \quad \gamma_k = \gamma \text{ for } k = 1, 2,$$

So far, identification of semiparametric and nonparametric propensity score models with callbacks is not available.

Using callback data for missing data analysis

In contrast, we consider a fundamentally nonparametric identification strategy:

- we propose an identifying assumption that allows for nonparametric and nonlinear propensity score models,
- establish the semiparametric theory
- and propose a suite of semiparametric estimators including doubly robust ones.

Outline

Identification

Notation: Frame variables (X, Y), X fully observed covariates, Y the outcome prone to missing values, $\mu = E(Y)$.

Callback data $R_k, k = 1, 2$ the response state of Y with $R_k = 1$ if Y is available in the kth call and $R_k = 0$ otherwise. $\Rightarrow R_2 \geq R_1$.

Define the odds ratio functions for the response propensity in the first and second calls as follows,

$$\Gamma_1(X, Y) = \log \frac{f(R_1 = 1 \mid X, Y) f(R_1 = 0 \mid X, Y = 0)}{f(R_1 = 0 \mid X, Y) f(R_1 = 1 \mid X, Y = 0)},$$

$$\Gamma_2(X, Y) = \log \frac{f(R_2 = 1 \mid R_1 = 0, X, Y) f(R_2 = 0 \mid R_1 = 0, X, Y = 0)}{f(R_2 = 0 \mid R_1 = 0, X, Y) f(R_2 = 1 \mid R_1 = 0, X, Y = 0)}.$$

Γ_k measures the impact of the missing outcome on the response propensity, the degree of nonignorable missingness, the resistance to respond caused by the outcome.

Identification

Assumption 1.

(i) Callback: $R_2 \geq R_1$;

(ii) Positivity: $0 < f(R_1 = 1 \mid X, Y) < 1$ and $0 < f(R_2 = 1 \mid R_1 = 0, X, Y) < 1$ for all (X, Y);

(iii) Stableness of resistance: $\Gamma_1(X, Y) = \Gamma_2(X, Y) = \Gamma(X, Y)$.

Theorem 1. Under Assumption 1, $f(X, Y, R_1, R_2)$ is identified.

No parametric models for the propensity scores or restrictions on the effects of covariates are imposed.

Identification

An immediate application to the linear logistic model.

Proposition 1. Assuming that $\text{logit}\ \pi_k(X,Y) = \text{logit}\ f(R_k = 1 \mid R_{k-1} = 0, X, Y) = \alpha_{k0} + \alpha_{k1}X + \gamma Y$, then α_{k0}, α_{k1}, and γ are identified.

Alho (1990); Kim & Im (2014); Guan et al. (2018) have to assume that $\alpha_{k1} = \alpha_1$.

Outline

Semiparametric efficiency theory

The stableness of resistance assumption defines the model

$$\mathcal{M}_{\text{npr}} = \left\{ f(X, Y, R_1, R_2) : \begin{array}{l} \text{Assumption 1 holds;} \\ A_1, A_2, \Gamma, f_2 \text{ are unrestricted.} \end{array} \right\}$$

Theorem 2. The efficient influence function for μ in the nonparametric model $\mathcal{M}_{\text{npr}}$ is

$$\begin{aligned} \text{IF}(O;\mu) \quad = \quad & \left\{ \frac{R_1}{\pi_1} - \frac{R_1}{\pi_2}\frac{1-\pi_1}{\pi_1} + \frac{R_2 - R_1}{\pi_2^2} \right\} Y \\ & - \left\{ \frac{R_1}{\pi_1} - \frac{R_1}{\pi_2}\frac{1-\pi_1}{\pi_1} + \frac{R_2 - R_1}{\pi_2^2} - 1 \right\} \frac{E(Y/\pi_2 \mid X, R_2 = 0)}{E(1/\pi_2 \mid X, R_2 = 0)} \\ & - \mu. \end{aligned}$$

Parameterization

We focus on estimation of the outcome mean $\mu = E(Y)$.

$$\begin{aligned} f(Y, R_1, R_2 \mid X) &= c_1(X) \cdot f(R_1 \mid X, Y = 0) \cdot \exp\{(R_1 - 1)\Gamma(X, Y)\} \\ &\quad \cdot f(Y \mid R_2 = 1, R_1 = 0, X) \\ &\quad \cdot f(R_2 \mid R_1 = 0, X, Y)^{1-R_1} f(R_2 = 0 \mid R_1 = 0, X, Y)^{-1} \end{aligned}$$

The baseline propensity scores
$A_1(X) = f(R_1 = 1 \mid X, Y = 0)$,
$A_2(X) = f(R_2 = 1 \mid R_1 = 0, X, Y = 0)$,
the odds ratio function $\Gamma(X, Y)$,
and the second-call outcome distribution
$f(Y \mid R_2 = 1, R_1 = 0, X)$.

Can be parameterized separately without hindering congeniality.

Hereafter, let $\pi_1(X, Y) = f(R_1 = 1 \mid X, Y)$, $\pi_2(X, Y) = f(R_2 = 1 \mid R_1 = 0, X, Y)$, $f_2(Y \mid X) = f(Y \mid R_2 = 1, R_1 = 0, X)$.

Doubly robust estimation

An estimator of μ motivated by the efficient influence function is

$$\begin{aligned}\hat{\mu}_{\mathrm{dr}} &= \hat{E}\left[\left\{\frac{R_1}{\hat{\pi}_1} - \frac{R_1}{\hat{\pi}_2}\frac{1-\hat{\pi}_1}{\hat{\pi}_1} + \frac{R_2 - R_1}{\hat{\pi}_2^2}\right\} Y\right] \\ &\quad -\hat{E}\left[\left\{\frac{R_1}{\hat{\pi}_1} - \frac{R_1}{\hat{\pi}_2}\frac{1-\hat{\pi}_1}{\hat{\pi}_1} + \frac{R_2 - R_1}{\hat{\pi}_2^2} - 1\right\} \frac{E(Y/\hat{\pi}_2 \mid X, R_2 = 0; \hat{\beta}_{\mathrm{dr}}, \gamma_{\mathrm{dr}})}{E(1/\hat{\pi}_2 \mid X, R_2 = 0; \hat{\beta}_{\mathrm{dr}}, \gamma_{\mathrm{dr}})}\right].\end{aligned}$$

where the nuisance parameters $(\hat{\alpha}_{1,\mathrm{dr}}, \hat{\alpha}_{2,\mathrm{dr}}, \hat{\beta}, \hat{\gamma}_{\mathrm{dr}})$ are obtained by solving

$$\begin{aligned}0 &= \hat{E}\left\{R_2(1-R_1)\cdot\frac{\partial \log f_2(Y \mid X;\beta)}{\partial\beta}\right\}, \\ 0 &= \hat{E}\left[\left\{\frac{R_1}{\pi_1(\alpha_1,\gamma)} - 1\right\}\cdot V_1(X)\right], \\ 0 &= \hat{E}\left[\left\{\frac{R_2 - R_1}{\pi_2(\alpha_2,\gamma)} - (1-R_1)\right\}\cdot V_2(X)\right], \\ 0 &= \hat{E}\left[\left\{\frac{R_2 - R_1}{\pi_2(\alpha_2,\gamma)} - \frac{1-\pi_1(\alpha_1,\gamma)}{\pi_1(\alpha_1,\gamma)}R_1\right\}\cdot\{U(X,Y) - E(U(X,Y) \mid X, R_2 = 0; \beta, \gamma)\}\right],\end{aligned}$$

Doubly robust estimation

Theorem 3. Under Assumption 1 and certain regularity conditions, $(\hat{\alpha}_{1,\mathrm{dr}}, \hat{\gamma}_{\mathrm{dr}}, \hat{\mu}_{\mathrm{dr}})$ are consistent and asymptotically normal provided one of the following conditions holds:

- $A_1(X;\alpha_1), \Gamma(X,Y;\gamma)$ and $A_2(X;\alpha_2)$ are correctly specified; or
- $A_1(X;\alpha_1), \Gamma(X,Y;\gamma)$ and $f_2(Y \mid X;\beta)$ are correctly specified.

Furthermore, $\hat{\mu}_{\mathrm{dr}}$ attains the semiparametric efficiency bound for the nonparametric model $\mathcal{M}_{\mathrm{npr}}$ when all models $\{A_1(X;\alpha_1), A_2(X;\alpha_2), \Gamma(X,Y;\gamma), f_2(Y \mid X;\beta)\}$ are correct.

$(\hat{\alpha}_{1,\mathrm{dr}}, \hat{\gamma}_{\mathrm{dr}}, \hat{\mu}_{\mathrm{dr}})$ are doubly robust against misspecification of $A_2(X;\alpha_2)$ and $f_2(Y \mid X;\beta)$, provided that the first-call propensity score $\pi_1(X,Y;\alpha_1,\gamma)$ (i.e., $A_1(X;\alpha_1), \Gamma(X,Y;\gamma)$) is correctly specified.

Estimation of a general smooth functional

Consider estimation of θ defined by the solution to a given estimating equation $E\{m(X, Y; \theta)\} = 0$. Assuming $\partial E\{m(\theta)\}/\partial\theta$ is non-singular, IPW, outcome regression-based, and doubly robust estimation of θ can be obtained simply by replacing $Y - \mu$ with m in the corresponding estimating equations of μ.

The efficient influence function for δ in the nonprametric model $\mathcal{M}$ is $\mathrm{IF}(O; \theta) = -[\partial E\{m(\theta)\}/\partial\theta]^{-1}\phi(O)$, where

$$\begin{aligned}\phi(O) &= \left\{\frac{R_1}{\pi_1} - \frac{R_1}{\pi_2}\frac{1-\pi_1}{\pi_1} + \frac{R_2 - R_1}{\pi_2^2} - 1\right\}\left\{\frac{E(m/\pi_2 \mid X, R_2 = 0)}{E(1/\pi_2 \mid X, R_2 = 0)} - m\right\} \\ &\quad + m.\end{aligned}$$

Takeaway points

- We propose the stableness of resistance assumption for identification, which is so far the most parsimonious condition characterizing the most flexible model for nonresponse adjustment with callbacks;
- we establish identification and develop IPW, outcome-regression based, and doubly robust estimation methods;
- we establish the semiparametric efficiency theory for using callbacks.

Outline

Simulations

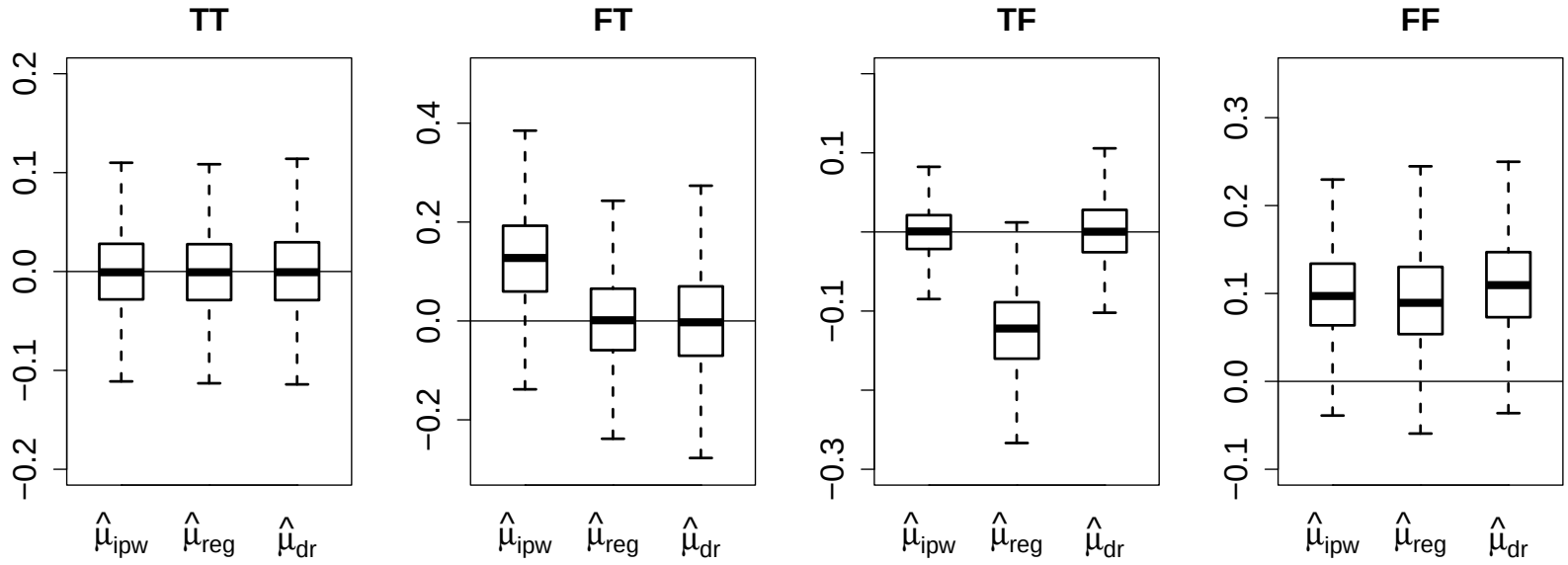


Figure 1: Bias for estimators of μ.

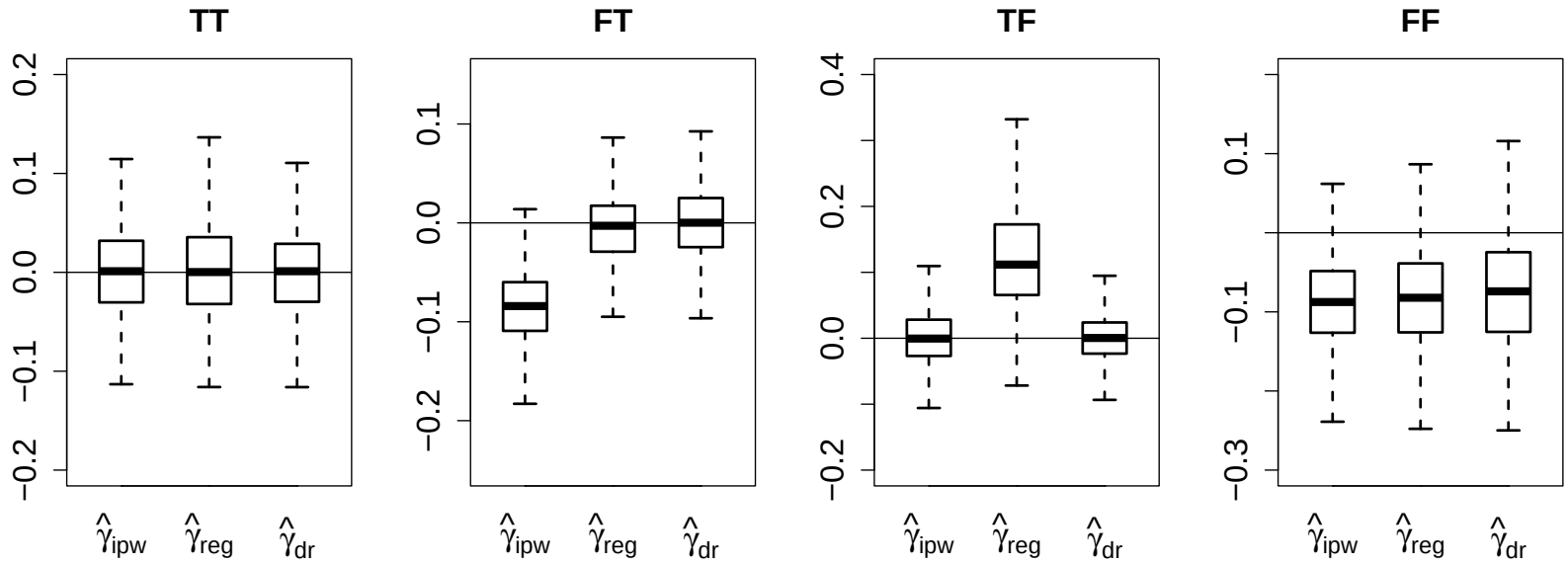


Figure 2: Bias for estimators of γ.

Simulations

Table 3: Coverage rate of 95% confidence interval

	μ			γ		
Scenarios	IPW	REG	DR	IPW	REG	DR
TT	0.953	0.949	0.948	0.963	0.962	0.948
FT	0.705	0.957	0.950	0.349	0.959	0.954
TF	0.954	0.341	0.949	0.947	0.744	0.954
FF	0.528	0.631	0.479	0.682	0.722	0.768

Application to the Consumer Expenditure Surveys

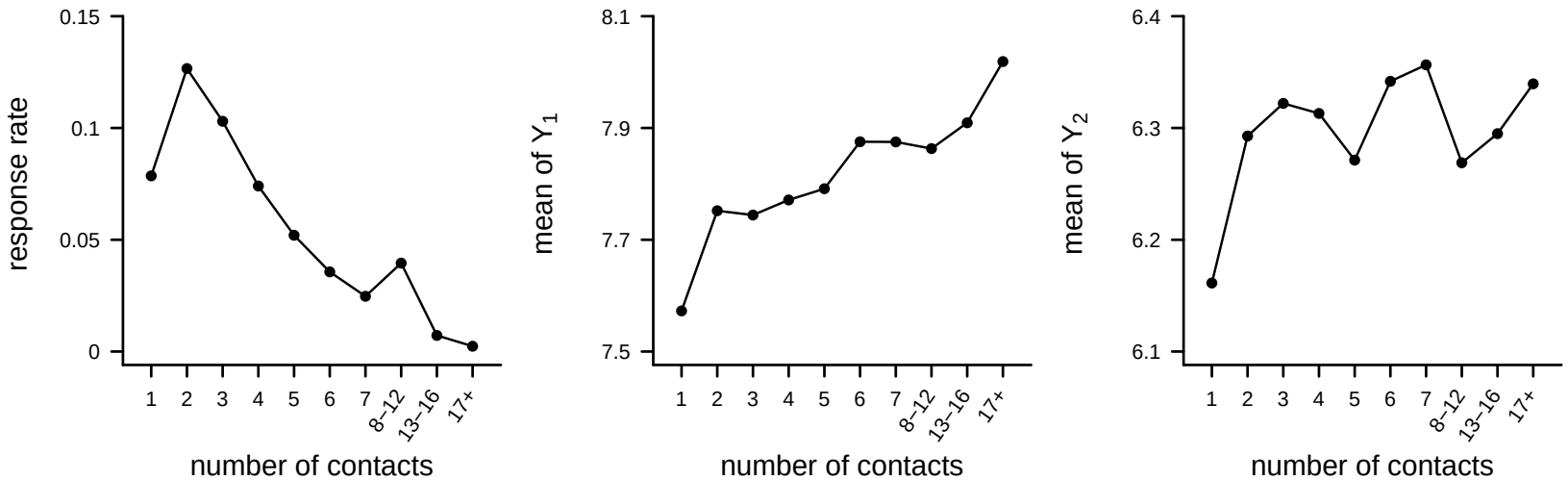
The Consumer Expenditure Survey (CES; National Research Council, 2013) is a nationwide survey conducted by the U.S. Bureau of Labor Statistics to find out how American households make and spend money. The survey data are annually released since 1980, which contain detailed paradata including the callback history.

We use the public-use microdata collected in the fourth quarter of 2018 for illustration.

Outcomes: Y_1 and Y_2 are the log of last quarter's expenditures on housing and on utilities, fuels, and public services, respectively.

Application to the Consumer Expenditure Surveys

9709 households with 27 contacts, 1992 responded in the early contact (1–2 calls), 3287 late contact (3+ calls), and 4430 never responded.



(a) Response rates (b) Outcomes mean (c) Outcomes mean

Figure 3: Response rates and outcomes mean for respondents for the CES application.

Application to the Consumer Expenditure Surveys

We use logistic propensity score models and a bivariate normal baseline outcome model.

Table 4: Point estimate, 95% confidence interval, and p-value for the CE application

	μ_1				γ_1		
	IPW	REG	DR	CC	IPW	REG	DR
Estimate	7.850	7.859	7.842	7.756	-0.265	-0.252	-0.238
CI or p-value	(7.769, 7.931)	(7.808, 7.910)	(7.800, 7.884)	(7.734, 7.778)	0.008	0.002	0
	μ_2				γ_2		
	IPW	REG	DR	CC	IPW	REG	DR
Estimate	6.339	6.352	6.345	6.285	-0.030	-0.048	-0.056
CI or p-value	(6.257, 6.422)	(6.299, 6.405)	(6.296, 6.393)	(6.264, 6.306)	0.806	0.615	0.521

The IPW and REG produce estimates close to DR; however, the CC estimate of the outcomes mean, in particular of μ_1, is well below the DR estimate. The odds ratios estimates obtained by IPW, REG, and DR are all negative, suggesting that high-spending people are more reluctant to the survey or more difficult to contact. The expenditure on housing play a more important role in the response process, this may be because the survey takes personal home visit as one of the main modes of interview and people with high expenditure on housing are more difficult to reach.

Reanalysis of Card (1995)'s dataset

3613 observations on the log of hourly wage in 1976 and 1978, years of schooling (*educ*), and a vector of fully-observed covariates including *exper*, *exper*^{2}, *black*, *south*, *smsa*, *nearc4*.

The wages for only 3010 units were recorded in 1976, and previous authors (e.g., Card, 1995; Okui et al., 2012; Wang & Tchetgen Tchetgen, 2018) have used this subsample to estimate the effect of education on wage.

Among the 603 nonrespondents to the call in 1976, 201 units were recorded in the follow up in 1978, which can be viewed as a second-call sample.

Let Y denote the log of hourly wage in 1976, $W = (\textit{educ}, C^{\mathrm{T}})$, $Z = (\textit{nearc4}, C^{\mathrm{T}})^{\mathrm{T}}$, and $X = (\textit{educ}, \textit{nearc4}, C^{\mathrm{T}})^{\mathrm{T}}$ with $C = (1, \textit{exper}, \textit{exper}^2, \textit{black}, \textit{south}, \textit{smsa})^{\mathrm{T}}$.

The instrumental variable estimand defined by $E\{Z(Y - \theta^{\mathrm{T}} W)\} = 0$, particularly the coefficient of *educ* (θ_{educ}), is of interest for evaluation of the causal effect of education on wage.

Reanalysis of Card (1995)'s dataset

Table 5: Point estimate, 95% confidence interval, and p-value for the NLSYM application

θ_{educ}

	IPW	REG	DR	CC
Estimate	0.111	0.107	0.102	0.132
CI	(0.006, 0.216)	(-0.015, 0.228)	(-0.016, 0.220)	(0.033, 0.231)

	γ			$\alpha_{1,\text{educ}}$		
	IPW	REG	DR	IPW	REG	DR
Estimate	1.651	2.249	2.199	-0.097	-0.107	-0.108
P-value	0	0.001	0	0.010	0.044	0.003

Reanalysis of Card (1995)'s dataset

The complete-case estimate is quite close to previous analysis results (Card, 1995; Okui et al., 2012) and suggests a significant return to schooling among respondents to the call in 1976.

Adjustment for nonresponse shows a significant return to schooling among the population consisting of both the respondents and the nonrespondents. but attenuates the estimates of θ_{educ}.

This result shows potential heterogeneity of the effects of education on wage: the effect in the nonrespondents to the call in 1976 is smaller than in the respondents.

This is supported by the estimation results for the propensity score model. Wage and education both have significant impacts (γ and $\alpha_{1,\mathrm{educ}}$) on nonresponse to the call in 1976.

Besides, the results suggest that men with higher wage and lower education are more likely to respond, and therefore, the missingness mechanism is likely nonignorable.

ANES NRFU Survey on 2020 U.S. presidential election

For decades, overestimation of turnout has been an issue in election surveys, and researchers have struggled with how to adjust for turnout bias (Enamorado & Imai, 2019). Since its beginning in the late 1940s, the American National Election Study (ANES) estimates of voter turnout are well known to be substantially higher than official turnouts.

The NRFU study uses mailed questionnaires to gather self-report data:

- began on January 28, 2021 with a randomized advance postcard;
- the first class invitation was mailed on February 1;
- followed by replacement questionnaires on March 2 and March 30;
- Completed questionnaires 3,779, the response rate 48.3%.
- The voting-eligible population (VEP) turnout for the 2020 presidential election is 66.2%. The weighted voter turnout based on the respondents is over 85%, indicating a severe overestimation bias.

Application to the ANES NRFU Survey

Table 6: Estimates of voter turnout

	turnout rate		**odds ratio parameter γ**	
method	estimate	95% confidence interval	estimate	p-value
CC	0.870	(0.853,0.886)	—	—
MAR	0.809	(0.783, 0.835)	—	—
COR	0.857	(0.837,0.876)	—	—
IPW	0.662	(0.544,0.780)	1.561	0.0086
REG	0.623	(0.428,0.817)	1.866	0.0497
DR	0.666	(0.523,0.808)	1.486	0.0124

- The DR estimate of the turnout is close to 0.662—the VEP voter turnout.
- Significant selection bias。
- Who did not vote is more hesitant to respond or more difficult to contact.

Application to the ANES NRFU Survey

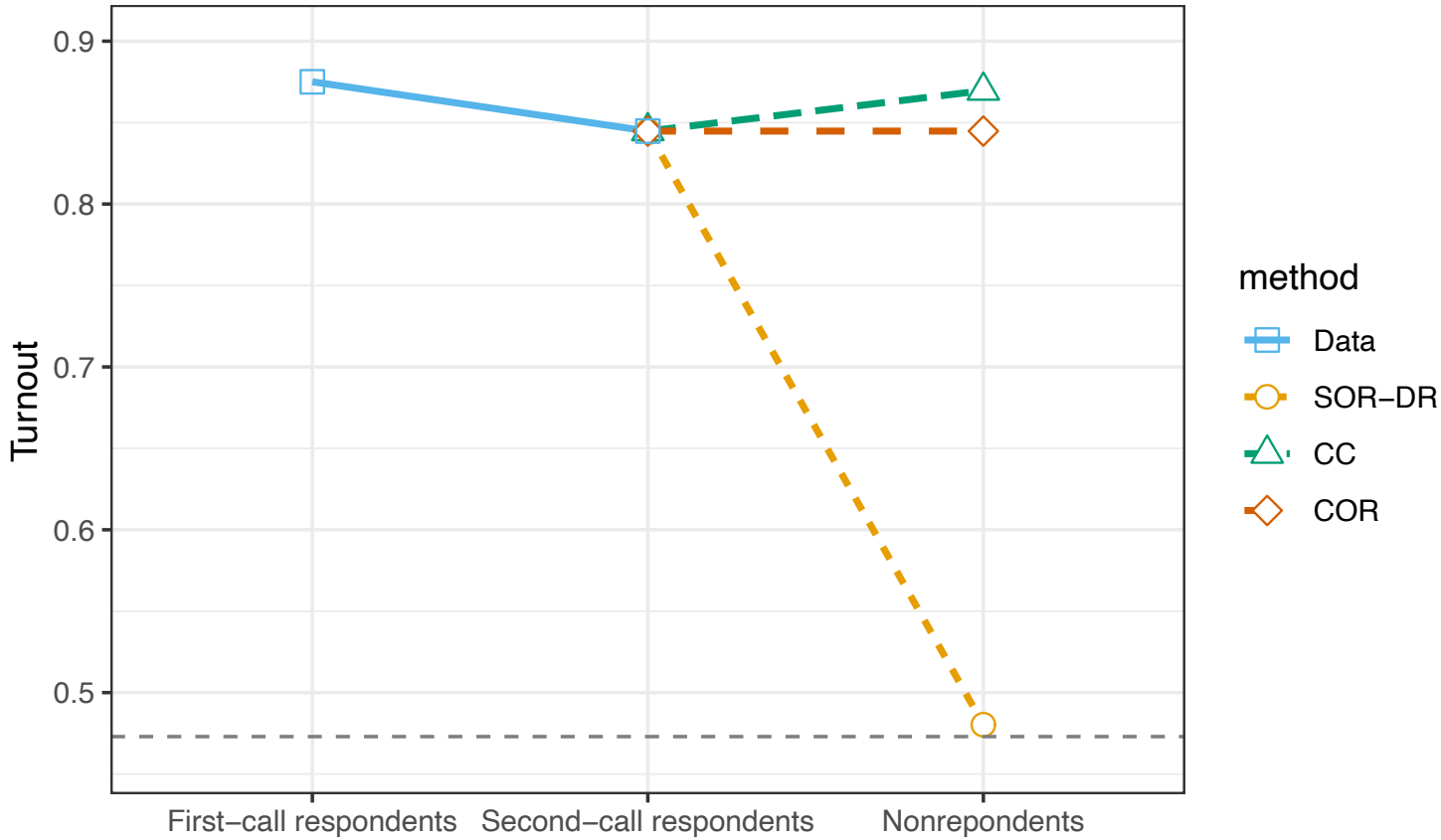


Figure 4: Turnout at each response stage.

Significant difference in the turnout between nonrespondents and respondents.

Application to the ANES NRFU Survey

Table 7: The coefficients of covariates in propensity score models

	The first contact stage		**The second contact stage**	
variable	estimate	p-value	estimate	p-value
intercept	-2.356	<0.0001**	-2.563	<0.0001**
m1sent: advance postcard	-0.200	0.1771	-0.281	0.3137
version: on page 1	0.197	0.1779	0.227	0.3434
title: long	-0.130	0.3817	-0.392	0.1709
incvis: visible	0.105	0.4700	0.178	0.4788
race: black	-0.296	0.1734	-0.130	0.6891
gender: male	-0.398	0.0055**	-0.606	0.0306**
age2: 30-59	0.671	0. 0012**	0.574	0.0521*
age3: 60+	1.487	<0.0001**	1.797	<0.0001**
education: some college	0.318	0.2850	0.387	0.3654

- Short title, political content on the first page, and a visible cash incentive may yield a higher response rate.
- Mailed advance postcards could be skipped, as it has no substantial response promotion. This is also discovered by DeBell (2022).
- Female and senior people are more likely to respond in both contacts.
- Educated people tend to respond.

Outline

Some extensions

With the assist of callback data one can test MAR because it is a special case of our stableness of resistance assumption.

We focused on nonresponse adjustment, but callback data are also useful to inform the design and organization of surveys, e.g., allocation of time and staff resources.

We proposed doubly robust estimation, and it is of interest to construct multiply robust estimation in the sense of (Vansteelandt et al., 2007).

Explore the idea of the stableness resistance in causal inference, case-control studies, and longitudinal studies.

Thanks!

References

Alho, J. M. (1990). Adjusting for nonresponse bias using logistic regression. *Biometrika* **77**, 617–624.

Bates, N. (2003). Contact histories in personal visit surveys: The survey of income and program participation (sipp) methods panel. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association*.

Biemer, P. P., Chen, P. & Wang, K. (2013). Using level-of-effort paradata in non-response adjustments with application to field surveys. *Journal of the Royal Statistical Society: Series A* **176**, 147–168.

Card, D. (1995). Using geographic variation in college proximity to estimate the return to schooling. In *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*, L. Christofides, E. Grant & R. Swidinsky, eds. Toronto: University of Toronto Press.

Chen, B., Li, P. & Qin, J. (2018). Generalization of Heckman selection model to nonignorable nonresponse using call-back information. *Statistica Sinica* **28**, 1761–1785.

References

Couper, M. (1998). Measuring survey quality in a casic environment. In *Proceedings of the Survey Research Methods Section of the ASA at JSM1998*. pp. 41–49.

Daniels, M. J., Jackson, D., Feng, W. & White, I. R. (2015). Pattern mixture models for the analysis of repeated attempt designs. *Biometrics* **71**, 1160–1167.

Das, M., Newey, W. K. & Vella, F. (2003). Nonparametric estimation of sample selection models. *The Review of Economic Studies* **70**, 33–58.

DeBell, M. (2022). Experimental effects of advance postcards, survey title, questionnaire length, and questionnaire content on response rates and incentive costs in a mail non-response follow-up survey. *Survey Practice* , 33168.

D'Haultfœuille, X. (2010). A new instrumental method for dealing with endogenous selection. *Journal of Econometrics* **154**, 1–15.

Drew, J. & Fuller, W. A. (1980). Modeling nonresponse in surveys with callbacks. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association*.

Enamorado, T. & Imai, K. (2019). Validating self-reported turnout by linking public opinion surveys with administrative records. *Public Opinion Quarterly* **83**, 723–748.

References

Fay, R. E. (1986). Causal models for patterns of nonresponse. *Journal of the American Statistical Association* **81**, 354–365.

Groves, R. M. & Couper, M. P. (1998). *Nonresponse in Household Interview Surveys*. John Wiley & Sons.

Guan, Z., Leung, D. H. & Qin, J. (2018). Semiparametric maximum likelihood inference for nonignorable nonresponse with callbacks. *Scandinavian Journal of Statistics* **45**, 962–984.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica* **47**, 153–161.

Kim, J. K. & Im, J. (2014). Propensity score adjustment with several follow-ups. *Biometrika* **101**, 439–448.

Kott, P. S. (2014). Calibration weighting when model and calibration variables can differ. In *Contributions to Sampling Statistics*, F. Mecatti, L. P. Conti & G. M. Ranalli, eds. Cham: Springer, pp. 1–18.

Kreuter, F. (2013). *Improving surveys with paradata*. New Jersey, Hoboken: John Wiley & Sons.

References

Kreuter, F., Müller, G. & Trappmann, M. (2010). Nonresponse and measurement error in employment research: making use of administrative data. *Public Opinion Quarterly* **74**, 880–906.

Lin, I.-F. & Schaeffer, N. C. (1995). Using survey participants to estimate the impact of nonparticipation. *Public Opinion Quarterly* **59**, 236–258.

Little, R. J. (1982). Models for nonresponse in sample surveys. *Journal of the American statistical Association* **77**, 237–250.

Little, R. J. & Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. Wiley: New York.

Liu, L., Miao, W., Sun, B., Robins, J. & Tchetgen Tchetgen, E. (2020). Identification and inference for marginal average treatment effect on the treated with an instrumental variable. *Statistica Sinica* **30**, 1517–1541.

Malinsky, D., Shpitser, I. & Tchetgen Tchetgen, E. J. (2020). Semiparametric inference for non-monotone missing-not-at-random data: the no self-censoring model. *Journal of the American Statistical Association* .

Manski, C. F. (1985). Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimator. *Journal of Econometrics* **27**, 313–333.

References

Miao, W., Ding, P. & Geng, Z. (2016). Identifiability of normal and normal mixture models with nonignorable missing data. *Journal of the American Statistical Association* **111**, 1673–1683.

Miao, W. & Tchetgen Tchetgen, E. (2016). On varieties of doubly robust estimators under missingness not at random with a shadow variable. *Biometrika* **103**, 475–482.

National Research Council (2013). Measuring what we spend: Toward a new consumer expenditure survey. National Academies Press.

Newey, W. K. (2009). Two-step series estimation of sample selection models. *The Econometrics Journal* **12**, S217–S229.

Okui, R., Small, D. S., Tan, Z. & Robins, J. M. (2012). Doubly robust instrumental variable regression. *Statistica Sinica* **22**, 173–205.

Olson, K. (2013). Paradata for nonresponse adjustment. *The Annals of the American Academy of Political and Social Science* **645**, 142–170.

Politz, A. & Simmons, W. (1949). An attempt to get the "not at homes"into the sample without callbacks. *Journal of the American Statistical Association* **44**, 9–16.

References

Qin, J. & Follmann, D. A. (2014). Semiparametric maximum likelihood inference by using failed contact attempts to adjust for nonignorable nonresponse. *Biometrika* **101**, 985–991.

Robins, J. M., Rotnitzky, A. & Scharfstein, D. O. (2000). Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*. Springer, pp. 1–94.

Rubin, D. B. (1976). Inference and missing data (with discussion). *Biometrika* **63**, 581–592.

Sadinle, M. & Reiter, J. P. (2017). Itemwise conditionally independent nonresponse modelling for incomplete multivariate data. *Biometrika* **104**, 207–220.

Sun, B., Liu, L., Miao, W., Wirth, K., Robins, J. & Tchetgen Tchetgen, E. (2018). Semiparametric estimation with data missing not at random using an instrumental variable. *Statistica Sinica* **28**, 1965–1983.

Tchetgen Tchetgen, E. J. & Wirth, K. E. (2017). A general instrumental variable framework for regression analysis with outcome missing not at random. *Biometrics* **73**, 1123–1131.

References

Vansteelandt, S., Rotnitzky, A. & Robins, J. (2007). Estimation of regression models for the mean of repeated outcomes under nonignorable nonmonotone nonresponse. *Biometrika* **94**, 841–860.

Wang, L. & Tchetgen Tchetgen, E. (2018). Bounded, efficient and multiply robust estimation of average treatment effects using instrumental variables. *Journal of the Royal Statistical Society: Series B* **80**, 531–550.

Wang, S., Shao, J. & Kim, J. K. (2014). An instrumental variable approach for identification and estimation with nonignorable nonresponse. *Statistica Sinica* **24**, 1097–1116.

Zhang, Y., Chen, H. & Zhang, N. (2018). Bayesian inference for nonresponse two-phase sampling. *Statistica Sinica* **28**, 2167–2187.

Zhao, J. & Shao, J. (2015). Semiparametric pseudo-likelihoods in generalized linear models with nonignorable missing data. *Journal of the American Statistical Association* **110**, 1577–1590.