Survey Treatments and Response Modes: Bayesian Survival Analysis with Competing Risks

Hiroaki Minato

U. S. Energy Information Administration
1000 Independence Ave., SW, Washington, DC 20585

Abstract

As part of research on Web survey method for Residential Energy Consumption Survey (RECS), the U.S. Energy Information Administration (EIA) sponsored the 2014 students' survey practicum in the University of Maryland's Joint Program in Survey Methodology (JPSM). Under a self-administered survey called the Home Energy Use Survey (HEUS), the practicum conducted several experiments on participatory incentives and questionnaire instruments. Here, we integrate both experimental treatments and contact procedures as survey treatments. A survey treatment is a composite treatment, for example, consisting only of advance letter and first invitation letter for a Web survey. A composite treatment takes place over time and composite treatments build up over time, depending on when responses come in, if at all, and what modes (Web or paper) the respondents choose. Possible response modes also depend on a survey treatment—in the example of survey treatment above, a response by the Web form is the only possible response mode. In this paper, we examine the relationships between survey treatments and response modes in terms of counts and rates of response over survey time period as well as over survey treatments.

Key Words: Bayesian survival analysis, competing risks, Residential Energy Consumption Survey (RECS), self-administered questionnaire, survey response modes, Web survey

1. Introduction

In surveys, a contact protocol (CP) is believed to make a difference in the survey cooperation or participation rate. For example, in an in-person at-household survey, an interviewer does not suddenly show up at the doorstep of selected sample households. There usually are planned procedures before interviewers' visits such as mailing of advance letters to notify the sample households of scheduled interview dates and times as well as to request their cooperation for the survey. In case of a self-administered questionnaire (SAQ) survey, the advance letter might provide the date when a paper questionnaire is expected to arrive at the household or the date when a survey Web site opens.

Anticipating that for some sample cases an initial in-person contact is not established or an SAQ is not returned, a survey researcher often specifies in the CP some follow-up rules and methods, which could include a phone call to make an appointment for interviewing visit, a letter to remind or re-request the sample case to fill in and return the SAQ, an offer for an alternative participatory or response mode, and so on.

Thus, a typical CP is applied over time interactively with sample cases—with a sequence of respondent-contact rules coming into effect conditionally on respondent-response statuses. Those sample cases that participated in an in-person at-household survey at initial interviewer visits would not normally get any further contacts. In an SAQ survey, second-time reminder postcards would be mailed only to those sample cases that received first-time reminder postcards but have not participated in the survey.

A CP generates various steps of contact procedure, which are dependent on sample case status over time. In other words, assignment of contact procedure steps is not random but implicitly self-selected by sample cases. It is a part of meta-response and interacts with the other parts of meta-response—response mode and time to response. As a result, special care is needed in comparing outcomes between sample cases receiving different contact procedure levels.

For the medical and behavioral science fields, Susan Murphy[1] (2012, 2015) has been developing the methodology called "adaptive interventions" (or "adaptive treatment strategies" or "dynamic treatment regimens"), where treatments are individually tailored:

> "Formally, an adaptive intervention is a sequence of decision rules that specify how the intensity or type of treatment should change depending on the patient's needs. We are developing sequential, multiple assignment, randomized trials (SMART) to enable scientists to build adaptive interventions."

There, the scientists have control in deciding treatments to the patients, while in our sample-case "care" problem, treatments are determined by potential survey participants, not survey researchers, which complicates our data structure and analysis. Nevertheless, we can learn a lot from the adaptive interventions methodology.

2. Example

As part of research on Web survey method for Residential Energy Consumption Survey (RECS), the U.S. Energy Information Administration (EIA) sponsored the 2014 students' survey practicum in the University of Maryland's Joint Program in Survey Methodology (JPSM). Under a self-administered survey called the Home Energy Use Survey (HEUS), the practicum conducted several experiments on participatory incentives and questionnaire instruments. Our objective was to assess the methodological feasibility of collecting residential consumption and characteristics data by a Web SAQ instrument, supplemented by a paper SAQ instrument if necessary. RECS has recently been conducted by in-person at-household interviewing.

A simple random sample of 10,000 residential single-family addresses was drawn from the Computerized Delivery Sequence File constructed by Marketing Systems Group. In order to control the variability among housing types, the address-based sampling (ABS) sample was limited to single-family addresses. Also, to increase the incidence rate and possibly the response rate, the ABS sample excluded addresses indicated as drop points, educational, P.O. Boxes, vacants, seasonal, or multi-dwelling.

The HEUS questionnaire is based on EIA's RECS questionnaire, including "questions about the respondents' housing unit characteristics, appliance usage, usage of electronics, heating and air conditioning, housing unit measurements, household characteristics and utility bills" (Peugh et al. 2014, p. 13). However, the practicum developed the HEUS questionnaire instrument more or less independently from EIA, trying to optimize its performance for each of the two possible survey response modes—Web SAQ and paper SAQ. An expected time length to complete the questionnaire was 10 to 15 minutes.

There were five experiments in this pilot study, three of which were concerned with the CP (the other two experiments were on some questionnaire wording and progress bar in the Web SAQ instrument):
❖ *Experiment 1* (Incentive) on survey participation incentives:

---

[1] Susan Murphy presented the Wald Memorial Lectures I, II, and III on this topic at the 2015 Joint Statistical Meetings.

> *Treatment 1* ($2-$0) = two dollars included in the first paper SAQ mailing and none in the second paper SAQ mailing
> *Treatment 2* ($1-$1) = one dollar included in the first paper SAQ mailing and one dollar in the second paper SAQ mailing

❖ *Experiment 2* (Web Reference) on reference to the Web response mode option in the first and second paper SAQ mailings:
> *Treatment 1* (Yes) = making the reference
> *Treatment 2* (No) = not making the reference

❖ *Experiment 5* ("Dated Materials") on "Dated Materials" wording on envelope:
> *Treatment 1* (Yes) = including the wording on envelope
> *Treatment 2* (No) = not including the wording on envelope

To simplify our current analysis, the treatments in Experiment 5 are ignored or marginalized.

The CP is given in Table 1 below.

Table 1: Contact Protocol—Mailing dates for invitation, survey, and reminder materials

Contact		Mailing Dates
1	Advance letter (survey contact and response modes unspecified)	05/21/14
2	1st invitation letter (for Web SAQ)	05/29/14
3	1st reminder postcard (for Web SAQ)	06/05/14
4	2nd invitation letter (for Web SAQ)	06/12/14
5	1st paper SAQ	06/24/14 – 06/26/14
6	2nd reminder postcard (for 1st paper SAQ)	07/07/14
7	2nd paper SAQ (identical to 1st paper SAQ)	07/17/14

Both Experiment 1 (Incentive) and Experiment 2 (Web Reference) are conducted over Contact 5 (1st paper SAQ) and Contact 7 (2nd paper SAQ) in Table 1 with a two-factor two-level full factorial design. The treatments were randomly pre-assigned to the sample cases before the survey. Note that each Treatment in each Experiment is a sequence of two sub-treatments and also that administration of the sub-treatments is conditional on whether Contact 5 and Contact 7 are made, where initiation of Contact 7 is conditional on the result of Contact 6, whose initiation is conditional on the result of Contact 5.

The survey period with the Web SAQ was from 05/30/14 to 08/29/14, the survey period with the first paper SAQ was from 06/24/14 to 08/29/14, and the survey period with the second paper SAQ was from 07/17/14 to 08/29/14.

Since there are time delays between sending and receiving of mailing materials (from the survey operator to the respondent and vice versa), applicable or in-effect contacts may be different from attempted contacts. For example, the 1st reminder postcard (for Web SAQ) might have been mailed to a respondent who has completed the Web SAQ after the mailing date but before receiving the postcard. We identify in-effect contacts, using the sending dates from the survey operator and the receiving dates at the survey operator.

Similarly, since Experiments 1 and 2 are conducted over 1st paper SAQ and 2nd paper SAQ, actual in-effect experimental treatments, opposed to planned experimental treatments, depend on when a sample case responds, if at all, to the survey. For example, if a sample case responded to the survey by the Web SAQ right after receiving the 1st invitation letter (for Web SAQ), then the sample case receives no paper SAQ's and thus no experimental treatments. If a sample case responded to the survey by the 1st paper SAQ right after receiving it, then the sample case does not receive the 2nd paper SAQ or its imbedded treatments. In-effect experimental treatments are implied in-effect contacts regarding Contacts 5 and 7.

(In our attempted or planned treatments, one might see some similarity to the intention-to-treat concept in randomized clinical trials; however, our treatments are adaptive to legitimate responses or non-responses by the sample cases, not to compliancy status or any other non-sampling errors.)

Combining in-effect contacts and in-effect experimental treatments, we notate in-effect survey treatments as follows in Table 2.

Table 2: In-Effect Survey Treatments

In-Effect Survey Treatment	In-Effect Contact	In-Effect Experiment 1 Treatment (Contact #)	In-Effect Experiment 2 Treatment (Contact #)
01_mail_12	1 2		
02_mail_123	1 2 3		
03_mail_1234	1 2 3 4		
04_mail_12345_nw1	1 2 3 4 5	$1 (5)	No Web reference (5)
04_mail_12345_nw2	1 2 3 4 5	$2 (5)	No Web reference (5)
04_mail_12345_w1	1 2 3 4 5	$1 (5)	Web reference (5)
04_mail_12345_w2	1 2 3 4 5	$2 (5)	Web reference (5)
05_mail_123456_nw1	1 2 3 4 5 6	$1 (5)	No Web reference (5)
05_mail_123456_nw2	1 2 3 4 5 6	$2 (5)	No Web reference (5)
05_mail_123456_w1	1 2 3 4 5 6	$1 (5)	Web reference (5)
05_mail_123456_w2	1 2 3 4 5 6	$2 (5)	Web reference (5)
06_mail_1234567_nw11	1 2 3 4 5 6 7	$1-$1 (5 7)	No Web reference (5 7)
06_mail_1234567_nw20	1 2 3 4 5 6 7	$2-$0 (5 7)	No Web reference (5 7)
06_mail_1234567_w11	1 2 3 4 5 6 7	$1-$1 (5 7)	Web reference (5 7)
06_mail_1234567_w20	1 2 3 4 5 6 7	$2-$0 (5 7)	Web reference (5 7)

3. Analysis

We examine the relationships between survey treatments and response modes in terms of counts and rates of response over survey time period as well as over survey treatments.
In the HEUS data, response and time to response may be defined in at least four different ways, depending on what we want to know or compare. (Time in terms of days is sufficiently precise for our analytic purpose, and that time unit is implicit throughout this paper.)

(1). If our concern is whether a sample case responded to the survey or not, then we can construct a binary response variable, indicating the event that a sample case responded to the survey (by the Web SAQ or one of the paper SAQ's), and measure the time from the beginning of the survey (30 May 2014) to the response event or to the end of the survey (29 August 2014) in case of non-response or right censoring. The survey treatment may be considered as one observed variable like any other observed variables of interest such as geographical location. But, all the caveats in an observational study apply (Rosenbaum 2002, 2010).

(2). If we are to compare the count or rate of response among the three response modes (Web SAQ, 1st paper SAQ, and 2nd paper SAQ), then we can treat them as competing events or risks. The time to each event may be calculated with a different beginning. For a Web SAQ response event, the beginning remains the same as the beginning of the survey (30 May 2014). But, for a 1st paper SAQ response event, the beginning may be the mailing of the 1st paper SAQ (24, 25, or 26 June 2014), and for a 2nd paper

SAQ response event, the beginning may be the mailing of the 2nd paper SAQ (17 July 2014). An assumption is that any effects by a given survey treatment on the response mode (Web SAQ, 1st paper SAQ, or 2nd paper SAQ) and time-to-event are independent. This is quite an unreasonable assumption to make; at the least, it should be hypothesized and tested.

(3). If we consider the binary response variable like the one in (1) and if we compute a response rate for each survey treatment group, then the average of the time-to-events derived in (2) could be used to compare among the survey treatment groups. This average may correct the potential bias in the average based on the time-to-events specified in (1), because the latter average may overestimate the time by implicitly assuming that the paper SAQ respondents could have responded by the Web SAQ, too.

(4). If we are to compare the count or rate of response among the three response modes (Web SAQ, 1st paper SAQ, and 2nd paper SAQ), then we can treat them as competing events or risks like in (2). And, as in (1), we measure the time from the beginning of the survey (30 May 2014) to the response event or to the end of the survey (29 August 2014) in case of non-response or right censoring. This is the standard approach of competing risks. In order to avoid the observational study pitfalls, a separate analysis could be conducted for each survey treatment group; however, if we use some observational variables in the survey-treatment-specific analysis, the same caveats in (1) would still be in order.

In this paper, we stay in the data analytic framework (4), considering the in-effect survey treatments as a categorical covariate (in fact, the only covariate), but we take the Bayesian approach, hierarchically modeling the uncertainty in the treatment parameters by some prior distributions.

3.1 Data Description

The response events, or the competing risks, are classified by response mode (Web SAQ, 1st paper SAQ, or 2nd paper SAQ). The observed count of each response event for each in-effect survey treatment is given in Table 3.

Table 3: Counts of Response Events for In-Effect Survey Treatments (n = 9,496)

In-Effect Survey Treatment	Response Mode				Censored or Non-Response	Total
	Web SAQ	1st Paper SAQ	2nd Paper SAQ	Sub-Total		
01_mail_12	285	-	-	285	50	335
02_mail_123	131	-	-	131	21	152
03_mail_1234	192	-	-	192	65	257
04_mail_12345_nw1	5	131	-	136	15	151
04_mail_12345_nw2	12	130	-	142	10	152
04_mail_12345_w1	13	139	-	152	12	164
04_mail_12345_w2	19	137	-	156	8	164
05_mail_123456_nw1	3	114	-	117	24	141
05_mail_123456_nw2	3	99	-	102	14	116
05_mail_123456_w1	5	65	-	70	28	98
05_mail_123456_w2	4	81	-	85	17	102
06_mail_1234567_nw11	0	45	82	127	1,746	1,873
06_mail_1234567_nw20	1	53	75	129	1,786	1,915
06_mail_1234567_w11	12	40	101	153	1,790	1,943
06_mail_1234567_w20	7	47	53	107	1,826	1,933

| Total | 692 | 1,081 | 311 | 2,084 | 7,412 | 9,496 |

We excluded some sample cases from our analysis because of operational or data quality problems:

    A. There are 37 cases, which were removed from the survey for some unspecified reasons.

    B. One case received some of the survey contacts incorrectly.

    C. 463 cases did not receive the minimum contacts (Contacts 1 and 2) required to measure the time to event or the time to censoring, because those contact mails were undeliverable.

    D. Three cases received incomplete or partial sequence of contacts due to some undeliverable contact mails, though they received the minimum contacts (Contacts 1 and 2).

Thus, the total sample size for our analysis is 9,496.

7,412 sample cases out of 9,496 did not respond to the survey by any response mode, and they are considered the (right) censored cases. For most of them, the censored date is 29 August 2014, on which the survey data collection was terminated. However, there were 217 censored cases, to which the minimum contacts (Contacts 1 and 2) were presumably established but some of the follow-up contacts could not be made due to undeliverable mails. Their censored dates are set to the dates when the first undeliverable mails were mailed to them plus one day of mail delivery time.

There are five sample cases whose response mode was recorded as the 2nd paper SAQ but whose interview dates (the dates the 2nd paper SAQ's were returned to the survey operation) had preceded the mailing date of the 2nd paper SAQ. This inconsistency was resolved by re-recording the response mode as the 1st paper SAQ.

282 cases in our analysis sample of 9,496 were identified as out-of-scope (OOS), while there were 743 OOS cases in the original sample of 10,000. Some of the 282 OOS cases responded to the survey by one of the response modes: 27 by the Web SAQ, 32 by the 1st paper SAQ, and 6 by the 2nd paper SAQ. (217 of the 282 did not respond to the survey, but they were identified as OOS after drawing the sample.) Since the sample eligibility criteria described in Section 2 were mainly for an arbitrary control of variation and quality, these OOS cases are not excluded from our analysis.

Further, the 9,496 analysis sample includes 21 partially survey-completed cases. In our response mode analysis, they are not distinguished from survey-completed cases.

Finally, our main analytic interest lies in the time to response event defined by response mode, i.e., the time to a competing risk. In-effect survey treatments classify the sample cases, as we compute the cause-specific hazard functions over the survey period. The data structure is rather complicated, but we analyze the data to identify and compare any effects of the in-effect survey treatments on the hazard rate, i.e., the response rate parameterized by time, and the response mode of respondents.

3.2 Modeling

Counting the response events for a given day in the survey period, we have the following overall distribution.
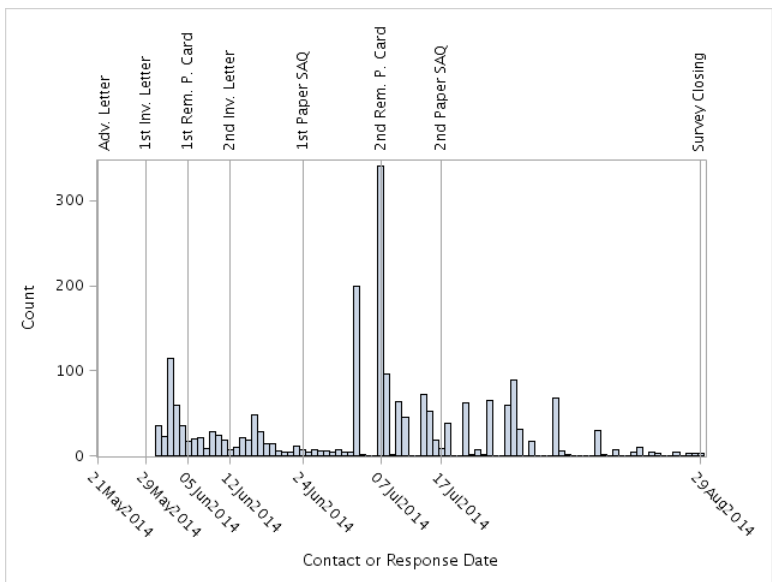
Figure 1: Overall Distribution of Response Date

Both of the 1st reminder postcard (for Web SAQ) and the 2nd reminder postcard (for 1st paper SAQ) may not have been very ineffective, although the reasons are unknown within the current study. The timing, the content, or any other factors could have caused the observations.

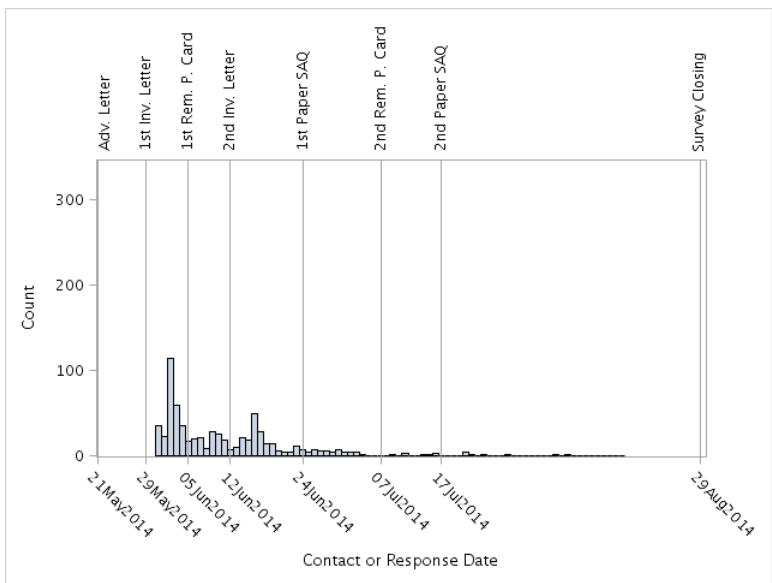Figures 2, 3, and 4 break up the overall distribution of response event date by the response mode.

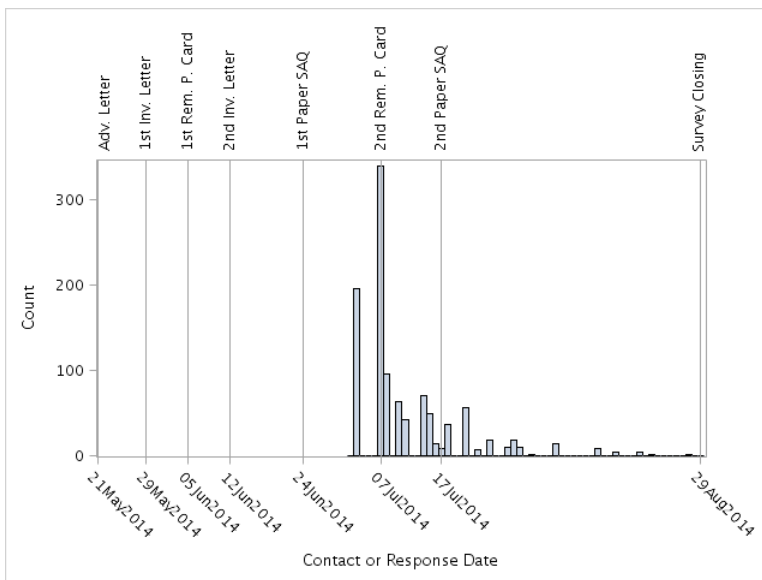**Figure 2:** Distribution of Web SAQ Response Date

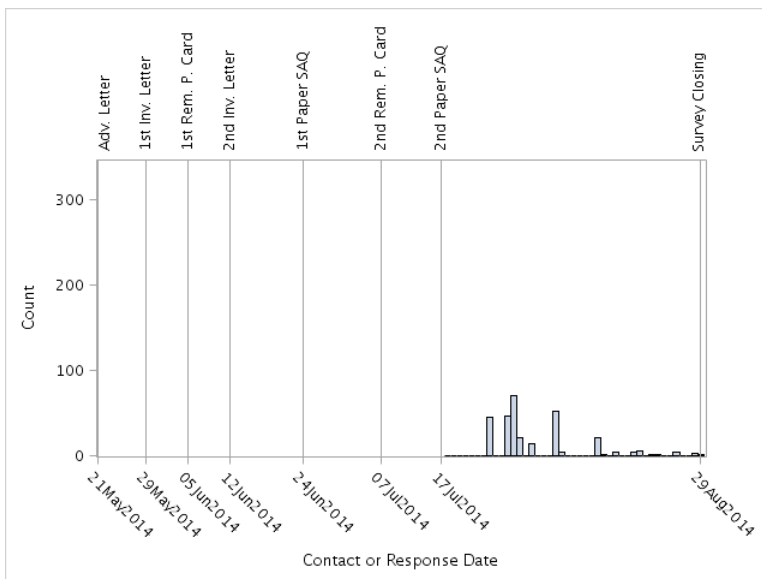**Figure 3:** Distribution of 1st Paper SAQ Response Date

Figure 4: Distribution of 2nd Paper SAQ Response Date

Note that, compared to the distribution of the Web SAQ response date, the distribution of the 1st paper SAQ response date is shifted to the right and that the distribution of the 2nd paper SAQ response date is shifted further to the right due to the delayed mailing dates of the 1st paper SAQ and the 2nd paper SAQ, respectively, by design. Recall, however, that the time to event, even when the event is a 1st paper SAQ response or a 2nd paper SAQ response, is measured from the beginning of the survey (30 May 2014).

The response date distribution in itself is interesting to look at, when the calendar dates are meaningful for analysis. There certainly are some Sunday or holiday effects at least on the delivery of the 1st and 2nd paper SAQ's. Here, we focus on the time-to-response distribution, which ignores the identity of calendar date but describes a time-length feature of sample response or non-response. Figure 5 depicts the cumulative response counts by response mode over the survey data collection period with 29 May 2014 as Day 0 and 29 August 2014 as Day 92. (The in-effect survey treatments are not distinguished.)
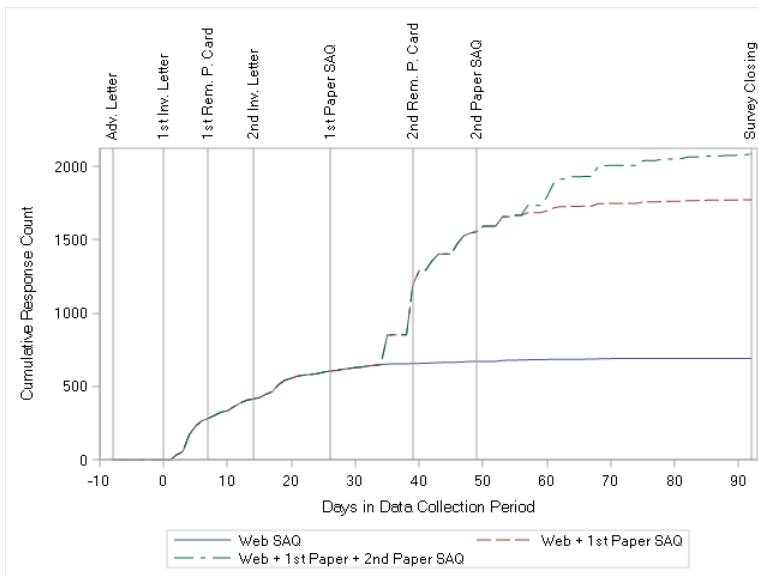
Figure 5: Cumulative Response Counts by Response Mode

For each in-effect survey treatment, the by-mode cumulative response counts can be plotted. As examples, Figure 6 shows the plot for the in-effect survey treatment of 06_mail_1234567_nw11 and Figure 7 show the plot for the in-effect survey treatment of 06_mail_1234567_n11. Recall that these two in-effect survey treatments differed only in the Experiment 2 (Web Reference) treatment assignment: 06_mail_1234567_nw11 referred the Web response mode option in the first and second paper SAQ mailings, while 06_mail_1234567_w11 did not make any such references. Both received the same Experiment 1 (Incentive) treatment: Treatment 2 ($1-$1) = one dollar included in the first paper SAQ mailing and one dollar in the second paper SAQ mailing.
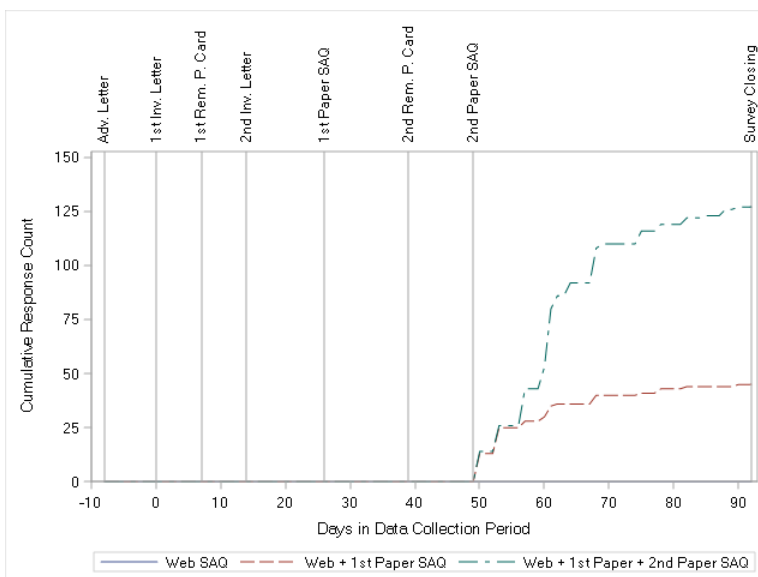

Figure 6: Cumulative Response Counts by Response Mode: 06_mail_1234567_nw11

Of the sample households that did not respond before the seventh in-effect contact with the nw11 experimental treatments, i.e., of the sample households that received all the in-effect contacts with the nw11 experimental treatments, Figure 6 shows effects, or responses by one of the response modes, of the in-effect survey treatment 06_mail_1234567_nw11after that treatment.

Figure 7: Cumulative Response Counts by Response Mode: 06_mail_1234567_w11

Similarly, of the sample households that did not respond before the seventh in-effect contact with the w11 experimental treatments, i.e., of the sample households that received all the in-effect contacts with the w11 experimental treatments, Figure 7 shows effects, or responses by one of the response modes, of the in-effec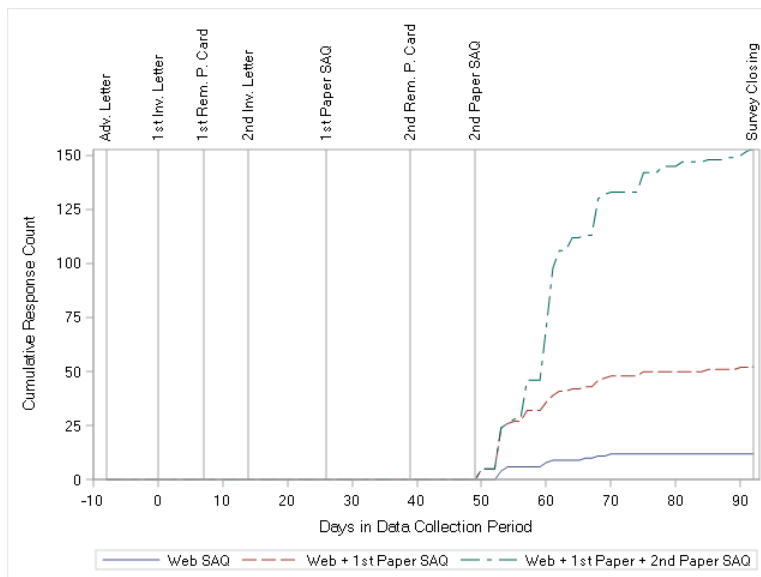t survey treatment 06_mail_1234567_w11 after that treatment. How do we estimate the difference in effect between, for example, 06_mail_1234567_nw11 and 06_mail_1234567_w11? And, how do we evaluate the estimated difference?

Variation in the time-to-response distribution is modelled by the in-effect survey treatment as a nominal categorical covariate (i.e., without any order constraints). Note that the four experimental treatments defined by Experiment 1 (Incentive) and Experiment 2 (Web Reference) were randomly assigned to the sample cases before the survey operation and that the assignment remained random at the end of Contact 4 (2nd invitation letter (for Web SAQ)) or at the end of In-Effect Survey Treatment 03_mail_1234. This is so, because there were no systematic interventions to disturb the random assignment—only the data cleaning steps in Section 3.1 slightly affected the sample balance of 2,500-2,500-2,500-2,500 in practically a random fashion.

Using the survival analysis terms, let us denote the failure time (the time to response) by T and the censoring (non-response or non-deliverable) time by G. If we have a complete set of K mutually exclusive competing risks (response modes), then our data can be written by the triplet (Y, C, δ), where $Y = \min(T, G)$, C is the censoring indicator variable, and δ is a random variable taking a value $k \in \{1, 2, \dots, K\}$. (A covariate or a vector of covariates x could be added.)

Here, we are interested in modeling the cause-specific hazard function, not necessarily the cumulative incidence function, though the latter is also described below for completeness. The hazard function is defined in the usual way as the instantaneous rate of failure conditional on non-failure before time t, but it is made specific to a given cause, say $\delta = k$:

$$h_k(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t, \ \delta = k \ |T \geq t)}{\Delta t}. \quad (1)$$

We introduce a vector of covariates, x, and (1) can be written as:

$$h_k(t, x) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t, \ \delta = k \ |T \geq t, x)}{\Delta t}. \quad (2)$$

(The extension to covariates is trivial, but it is kept in the notation in below.)

Since we have a complete set of *K* mutually exclusive competing risks, we have:

$$h(t,x) = \sum_{k=1}^{K} h_k(t,x). \quad (3)$$

We assume that competing risks are independent, but note that the assumption is untestable, although it may not be unreasonable particularly when the competing risks are conditioned on observed covariates (Kalbfleisch and Prentice 1980, Chapter 7).

Meanwhile, the cumulative incidence function (Kalbfleisch and Prentice 1980, Chapter 7) is the probability of failure before time *t*, and it is also specifically defined for a given cause, say $\delta = k$. (Figures 5, 6, and 7 empirically show the similar relations in terms of observed counts.) Continuing to let *x* be a vector of covariates, the (conditional on *x*) cumulative incidence function is the conditional probability:

$$F_k(t,x) = \Pr(T < t, \delta = k \mid x). \quad (4)$$

Fine and Gray (1999) call it the (conditional on *x*) subdistribution, because its limit, as *t* goes to infinity, is not necessarily one.

The (conditional on *x*) cumulative incidence function can be further represented as:

$$F_k(t,x) = \int_0^t S(u,x)h_k(u,x)\,du = \int_0^t S(u,x)dH_k(u,x), \quad (5)$$

where $S(u, x)$ is the marginal (over *k*) survival function, conditional on *x*, and $H_k(u,x)$ is the cause-specific cumulative hazard function, conditional on *x*:

$$H_k(u,x) = \int_0^u h_k(v,x)dv \quad (6)$$

and

$$S(u,x) = \Pr(T \geq u \mid x) = \exp\left(-\sum_{k=1}^{K} H_k(u,x)\right) = \exp(-H(u,x)). \quad (7)$$

Given the censored data, we can tie into, for example, the Kaplan–Meier estimator for the survival function or into, e.g., the Nelson-Aalen estimator for the cause-specific cumulative hazard function. We actually use the Breslow (1972) method to compute the survival function, exponentiating the negative empirical cumulative hazard function.

Note that one minus the Kaplan-Meier estimate of survival curve is the cumulative incidence function when there is only one competing risk: $S = 1 - F = \exp(-H)$. When there are more-than-one competing risks and when they are not independent, one usually looks at the cumulative incidence function (H_k), not the survival curve (S), because the former is specific about causes or competing events over time and thus is more informative (Kleinbaum and Klein 2005, pp. 412-420).

We model the cause-specific hazard function by the cause-specific base hazard function and survey treatment parameters, a special case of Cox regression model. (If one is interested in cumulative incidence functions rather than cause-specific hazard functions, one might follow Fine and Gray (1999) to regress the hazard of a given cumulative incidence function directly on covariates. However, this approach requires some caution in interpreting the hazard of the subdistribution.) Under the Bayesian hierarchical modeling approach, the survey treatment parameters are distributed under some prior probability model, representing our uncertainty in those parameters. It is possible to impose some order constraints on the survey treatment parameters, but we do not make any such model assumptions here.

Suppose we have *n* observations, indexed by *i*, each of which carries the following pieces of information:

$$\text{Data}_i = (y_i, c_i, \delta_i, x_i),$$

where y_i is the observed time, c_i is a censoring indicator (1 if censored and 0 otherwise), δ_i is a competing-risks index (1, …, K or undefined for a censored case), and x_i is a vector of covariates.

The likelihood function under an independent non-informative censoring mechanism is:

$$\prod_{i=1}^{n} [h_{\delta_i}(y_i, x_i)]^{(1-c_i)} S(y_i, x_i) = \prod_{i=1}^{n} [h_{\delta_i}(y_i, x_i)]^{(1-c_i)} \exp\left(-\sum_{k=1}^{K} H_k(y_i, x_i)\right)$$

$$= \prod_{i=1}^{n} \left\{ [h_{\delta_i}(y_i, x_i)]^{(1-c_i)} \prod_{k=1}^{K} \exp(-H_k(y_i, x_i)) \right\}$$

$$= \prod_{i=1}^{n} \prod_{k=1}^{K} [h_{\delta_i}(y_i, x_i)]^{(1-c_i)} \exp(-H_k(y_i, x_i)) .$$

Under Cox's proportional hazards regression model without any assumptions about the baseline hazards $\lambda_k(t)$, k=1, …, K, the partial likelihood function, assuming there are no tied failure times for simplicity, is:

$$L = \prod_{k=1}^{K} \prod_{i=1}^{n_k} \frac{e(x'_{k,i(k)} \beta_k)}{\sum_{j \in R(t_{k,i})} e(x'_{k,j} \beta_k)} ,$$

where n_k is the number of distinct failure times for k, $t_{k,i}$ is the i-th distinct failure time for k, $R(t_{k,i})$ is the risk set at time $t_{k,i}$ for k, and $i(k)$ is the index of the case that failed at time $t_{k,i}$ due to k.

The regression parameters β_k are given the uniform priors. The baseline hazards $\lambda_k(t)$ are unspecified here; thus, our model is semi-parametric. However, to be fully Bayesian instead of semi-Bayesian, one might consider a non-parametric mixture of piecewise exponential models, a non-parametric gamma or beta process, or a parametric Weibull model for the baseline hazard functions. Also note that some full model specification would be necessary if one is interested in estimating the cumulative incidence functions. At this moment, the most complete method in the context of Bayesian proportional hazards models is perhaps the Bayesian bootstrap method by Kim and Lee (2003).

3.3 Estimation Result

I consider the K risks as K strata in the Cox model and produce the mean estimates of the regression parameters over the K risks, but as a Bayesian hierarchical model. Alternatively, one can estimate the regression parameters for each risk. But, the latter approach could hit an estimation problem for those risks with small data after controlling for the covariates (see Table 3). (Again, cumulative incidence functions are not estimated in this paper.)

SAS PROC PHREG was used to estimate the posterior distribution by the Gibbs sampling with the MLE's as initial values. With the number of samples of 3,000 and the number of burn-in samples of 1,000 without thinning, the diagnostics looked all good—(a) autocorrelation, (b) the standard error of the posterior mean estimate, (c) the stationarity (the Heidelberger and Welch tests), (d) the convergence (the Gelman and Rubin diagnostics and the Geweke diagnostics), and (e) the accuracy of the estimated quantile of a chain (the Raftery and Lewis diagnostics). The estimates of the regression parameters or the in-effect survey treatment indicators are shown in Table 4.

Table 4: Posterior Distributions of In-Effect Survey Treatments

Posterior Distribution Summary

Parameter	N	Mean	Standard Deviation	Percentiles			95% HPD Credible Interval	
				25%	50%	75%	Lower Bound	Upper Bound
01_mail_12		0						
02_mail_123	3000	-0.97	0.19	-1.04	-0.97	-0.90	-1.19	-0.76
03_mail_1234	3000	-2.13	0.10	-2.20	-2.13	-2.06	-2.32	-1.93
04_mail_12345_nw1	3000	-2.75	0.11	-2.83	-2.76	-2.68	-2.98	-2.55
04_mail_12345_nw2	3000	-2.53	0.11	-2.61	-2.53	-2.46	-2.75	-2.31
04_mail_12345_w1	3000	-2.57	0.11	-2.65	-2.57	-2.49	-2.79	-2.36
04_mail_12345_w2	3000	-2.54	0.11	-2.61	-2.54	-2.46	-2.76	-2.33
05_mail_123456_nw1	3000	-4.00	0.12	-4.08	-4.00	-3.92	-4.24	-3.77
05_mail_123456_nw2	3000	-3.79	0.12	-3.87	-3.79	-3.72	-4.02	-3.54
05_mail_123456_w1	3000	-4.25	0.14	-4.35	-4.25	-4.16	-4.53	-3.98
05_mail_123456_w2	3000	-3.97	0.13	-4.05	-3.96	-3.88	-4.24	-3.73
06_mail_1234567_nw11	3000	-7.52	0.11	-7.59	-7.52	-7.44	-7.73	-7.30
06_mail_1234567_nw20	3000	-7.52	0.11	-7.60	-7.52	-7.44	-7.74	-7.30
06_mail_1234567_w11	3000	-7.36	0.11	-7.43	-7.36	-7.29	-7.58	-7.16
06_mail_1234567_w20	3000	-7.72	0.12	-7.80	-7.72	-7.64	-7.95	-7.50

Compared to the in-effect survey treatment group of 01_mail_12, the hazard rate (i.e., time-dependent response rate) of any other in-effect survey treatment group is decreased by the multiplicative factor of exp(parameter estimate), as all the parameter estimates are negative. The magnitude of the hazard rate reduction increases, as the number of in-effect contacts increases, which suggests a diminishing return of in-effect contacts. A hazard rate diminishment or a decrease in effectiveness of in-effect contact is sizable not only from the "12" group to the "1234567" group but also from the "123456" group to the "1234567" group. The decrease is about one from the "12" group to the "123" group, from the "123" group to the "1234" group, and from the "12345" group to the "123456" group, while the decrease is less than 0.5 from the "1234" group to the "12345" group, which means the "12345" group kept the similar level or rate of effectiveness as that of the "1234" group.

At each of the in-effect contacts where the two experiments were administered (i.e., in-effect contacts "12345", "123456", and "1234567"), the magnitudes among the four treatment groups are quite similar, suggesting none of the treatments are affecting the hazard rates differently. For example, of the sample households that survived the in-effect survey treatments "123456_" and received the in-effect survey treatments "1234567_", about 500 sample households showed some effects (Web, 1st paper, or 2nd paper SAQ responses) of those in-effect survey treatments, while the same in-effect survey treatments had no effects on more than 7,000 sample households. And, of the 500 or so sample households that responded, differential experimental treatments (nw11, nw20, w11, and w20) showed no differential effects on their hazard rates among those households.

4. Future is Now!

I believe a CP should and could be individualized or adapted for each sample case even during a given survey period. A big question is how. One approach would be to collect, analyze, and learn from paradata during a survey and dynamically to adapt its CP for each sample case. Paradata of a current survey could be augmented by past or external paradata, or a CP learning system of a current survey might incorporate some survey response theory constructed from or tested with those paradata. There are many interesting and applicable ideas in the machine learning field such as neural network with the Boltzmann machine (Wolchover 2013) as well as in the clinical trials field as I mentioned earlier. It is the time for us to learn and adapt!

Acknowledgements

References

Breslow, N. E. (1974), "Covariance Analysis of Censored Survival Data," *Biometrics*, 30: 89–99.

De Blasi, P. and Hjort, N. L. (2007), "Bayesian Survival Analysis in Proportional Hazard Models with Logistic Relative Risk," *Scandinavian Journal of Statistics*, 27: 229-257.

DeLong, D. M., Guirguis, G. H., and So, Y. C. (1994), "Efficient Computation of Subset Selection Probabilities with Application to Cox Regression," *Biometrika*, 81: 607–611.

Efron, B. (1977), "The Efficiency of Cox's Likelihood Function for Censored Data," *Journal of the American Statistical Association*, 72: 557–565.

Fine, J. P. and Gray, R. J. (1999), "A Proportional Hazards Model for the Subdistribution of a Competing Risk," *Journal of the American Statistical Association*, 94: 496-509.

Gasbarra, D., and Karia, S. R. (2000), "Analysis of Competing Risks by Using Bayesian Smoothing," *Scandinavian Journal of Statistics*, 27: 605-617.

Ge, M., and Chen, M.-H. (2012), "Bayesian Inference of the Fully Specified Subdistribution Model for Survival Data with Competing Risks," *Lifetime Data Anal*. [online], 18: 339-363, DOI: 10.1007/s10985-012-9221-9. Available at http://www.ncbi.nlm.nih.gov/pubmed/22484596.

Gumbel, E. J. (1960), "Bivariate Exponential Distributions," *Journal of the American Statistical Association*, 55: 698-707.

Hjort, N. L. (1990), "Nonparametric Bayes Estimators Based on Beta Processes in Models for Life History Data," *Annals of Statistics*, 18: 1259-1294.

Kalbfleisch, J. D. and Prentice, R. L. (1980), *The Statistical Analysis of Failure Time Data*, New York, NY: Wiley.

Kim, Y. and Lee, J. (2003), "Bayesian Bootstrap for Proportional Hazards Models," *Annals of Statistics*, 31: 1905-1922.

Kleinbaum, D. G., and Klein, M. (2005), *Survival Analysis*, (2nd ed.), New York, NY: Springer.

Lee, K. H., Haneuse, S., Schrag, D., and Dominici, F. (2015), "Bayesian Semiparametric Analysis of Semicompeting Risks Data: Investigating Hospital Readmission after a Pancreatic Cancer Diagnosis," *Journal of the Royal Statistical Society*: Series C (*Applied Statistics*), 64: 253–273, DOI: 10.1111/rssc.12078.

Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D., and Murphy, S. A. (2012), "A 'SMART' Design for Building Individualized Treatment Sequences," *Annual Review of Clinical Psychology* [online], 8: 21-48, DOI: 10.1146/annurev–clinpsy–032511–143152. Available in final edited form at http://www.annualreviews.org/doi/abs/10.1146/annurev-clinpsy-032511-143152. Author manuscript available in PMC 2014 Jan 9 at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3887122/.

Murphy, S. A. [cited 11 May 2015], "Adaptive Interventions," *The Methodology Center* [online]. Available at https://methodology.psu.edu/ra/adap-inter.

Peugh, J., Langdale, K., and Goyle, A. (2014), "University of Maryland Joint Program in Survey Methodology 2014 Home Energy Use Survey Methodology Report." Unpublished report.

Rosenbaum, P. R. (2002), *Observational Studies*, (2nd ed.), New York, NY: Springer.

——(2010), *Design of Observational Studies*, New York, NY: Springer.

Salinas-Torres, V., Pereira, C. A. B., and Tiwari, R. (2002), "Bayesian Nonparametric Estimation in a Series System or a Competing Risks Model," *Journal of Nonparametric Statistics*, 14: 449-458, DOI: 10.1080/10485250213114. Available at https://www.ime.usp.br/~cpereira/publications/series3.pdf from https://www.ime.usp.br/~cpereira/.

SAS Institute Inc. (2013), *SAS/STAT® 13.1 User's Guide*, Cary, NC: SAS Institute Inc.

Wolchover, N. (2013), "As Machines Get Smarter, Evidence They Learn Like Us," *Quanta Magazine* [online]. Available at https://www.quantamagazine.org/20130723-as-machines-get-smarter-evidence-they-learn-like-us/.

Zugna, D. (2008), "A New Bayesian Approach to Competing Risks Models in Longitudinal Studies," [online]. Presented at Atti della XLIV Riunione Scientifica (Università della Calabria, 25 -27 giugno 2008). Available at http://old.sis-statistica.org/files/pdf/atti/Atti%20pubblicati%20da%20Cleup_109-116.pdf from http://old.sis-statistica.org/index.php?area=main&module=contents&contentid=361.