A Longitudinal Analysis of PPI Response & Nonresponse Patterns November 2017

Joseph Valentine
U.S. Bureau of Labor Statistics
2 Massachusetts Avenue NE, Washington, DC 20212
valentine.joseph@bls.gov

Abstract
For various reasons, survey respondents fail to provide data after periods of cooperation that vary in length. Likewise, after a period of nonresponse, respondents may begin providing data again. The purpose of this paper is to provide statistical estimates of response length, defined as consecutive months with a price reported, and its converse, nonresponse length, for items reported to the Producer Price Index (PPI), a principal federal economic indicator. The estimates are produced via the use of Cox Proportional Hazard models and their derived Survival Curves, which are modified to predict the occurrence of both nonresponse and return to cooperation from nonresponse over different, finite time horizons. Subsequently, estimates of response and non-response length are produced for a set of two-digit NAICS sectors representing up to 564 six-digit NAICS industries, which are then plotted in the x-y plane for comparison. The results suggest modest variation in response and nonresponse across NAICS sectors, but also the existence of seasonal variation in the occurrence of nonresponse. The estimates produced may provide guidance in handling nonresponse.

Key Words: Price index, Survivor function, Censored observations, Nonresponse, Hazard function

Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.

1. Introduction

The Producer Price Index (PPI) is a family of economic indicators that use voluntarily supplied survey data. As a result, nonresponse occurs frequently at the item level and at varying lengths.[1] Some items enter nonresponse status and respondents never provide data again, due to personnel changes, going out of business, or simply reporter fatigue.[2] Conversely, there are also respondents who never fail to provide data. Of all unique items initiated between 2009 and 2013, there are exactly 4,764, or roughly five percent of all items initiated in that time span, that have yet to provide data as of July 2016.

One of the most important questions regarding nonresponse is: how long does it last? That is, how long do respondents fail to provide prices, for a given item, before they begin providing them again? This question matters for several reasons. One reason is because PPI index movements are derived from a set of price changes from individual goods and services. As item nonresponse increases, the number of available and usable prices declines, and index movements could be influenced by large price swings from a (relatively)

[1] An item is defined as an individual good or service, for which prices are being provided in a given month, e.g. corporate management consulting services.
[2] The terms reporter, respondent and establishment will be used synonymously throughout.

small group of items.[3] An issue closely related to nonresponse length is response length, or the number of consecutive months in which a respondent provides data before they stop responding.

The purpose of this paper is to provide statistical estimates of response length and nonresponse length, measured in months. A response history of each item is constructed, from which the length of both response and nonresponse can be calculated. All results are produced via calculation of Kaplan Meier (KM) survival curves and their corresponding (derived) hazard functions. Survival curves are based on a cox proportional hazard model with a robust sandwich variance estimator. The paper is organized as follows: part 2 provides background for the PPI program, while part 3 discusses the data used, and part 4 provides a brief discussion of survival analysis techniques used. Part 5 discusses the results and provides interpretation and analysis, while part 6 offers suggestions for alternative nonresponse strategies.

2. Background

Establishments selected for potential participation in the PPI survey are identified via their Employment Identification Number (EIN) in the Unemployment Insurance (UI) System, in which most firms are legally obligated to participate. Details on the selection process and the statistical methodology behind it can be found in the BLS Handbook of Methods, Chapter 14. Once an establishment is selected, a BLS field representative visits to solicit cooperation. If the establishment's management agrees to cooperate, the field economist proceeds to select goods and services to be priced over time from among all revenue producing transactions. Each good or service for which prices are to be provided is called an item. During the initial visit, prices are provided to the field economist to initiate the establishment and its items into the PPI survey.

Following initiation, establishments are contacted to provide prices on an agreed-upon schedule, a process called "repricing." Typically establishments provide prices monthly, though some provide less often, for example quarterly. Nonresponse occurs when an establishment does not provide prices to PPI when requested. In this case, a PPI analyst contacts the establishment via phone—a nonresponse call. In many cases the respondent provides the analyst with pricing data over the phone and this is then considered a response. If the analyst is unable to speak with the respondent during the nonresponse call, and the respondent still has not provided data, they will be contacted again the following month. This process continues until the respondent provides prices or they are discontinued from the survey. After 12 consecutive unfilled requests a respondent is typically dropped unless there is sufficient justification for them to remain. Follow-ups of the aforementioned respondents is then done at 18 requests and if a given respondent still has not provided data they are then discontinued. The current PPI nonresponse procedure, as designed, is responsible for producing and shaping the data used in the subsequent analysis. In this regard, for the purposes of analysis, it may be considered part of the data generating process.

3. Data

The analysis below is performed using item-level data from establishments that agreed to participate in the PPI. Table 1 below provides a generic example of the data used to calculate the aforementioned survival curves. The index reference month (irm), price, item, and cycle variables represent the month for which a given price would be used in index calculation, the price provided by the respondent or entered by the analyst, a unique identifier for each specific item, and an indicator of whether the respondent was expected to provide data in the given month. The cycle variable takes one of three values, ("O", "M", and "N") each of which has a different meaning in the context of survey response and nonresponse. A cycle value of "M" indicates that the PPI requested data for the item in a given month, while "N" indicates that data was not

[3] Despite this, there is no statistical evidence that PPI indexes suffer from nonresponse bias. For more, see Chopova, Hesley, et al.

requested. A value of “O” indicates that the PPI did not request data if the most recent PPI request was already fulfilled. If the respondent had not provided data during the most recent request period, data would be requested until it is provided. For example, if cycle was “M” in April and no data was provided by May, when cycle was equal to “O”, another request would be made until data for April was provided. If the data was provided for April prior to the May request no additional request would be made. The cycle variable largely determines nonresponse status.

The variables response and nonresponse (nr) are dummy variables created to indicate if the respondent provided the PPI with a price in a given month or if such a price was entered by the analyst after speaking with the respondent or performing research. Using the item, response and nr variables, the cum.response and cum.nr variables were created to count the number of consecutive months in which a respondent did or did not provide data for a given item. Using the cum.response and cum.nr variables, the response history of every item can be constructed. This dataset can then be subset to collect the months prior to an item entering nonresponse status or resuming cooperation after a period in nonresponse status. For example, in the case below, the data ranges from September 2013 through January 2015.

The table shows that for the item in question, the respondent provided prices for six consecutive months before entering nonresponse status, in March 2014. They remained a nonrespondent for three months, after which they provided data for two months. Subsequently, they never provided data again. This is quantified as five months of nonresponse. It should be noted that, though prices are missing for nine months total, the item is only considered to be in nonresponse status for eight months. The difference is due to the August 2014 month, during which the cycle variable was “O” and a price was reported in June 2014, when the most recent request was made. Because the item was off cycle in August 2014, and a price was reported during the month when it was most recently requested in June 2014, it is not in nonresponse status for August 2014.

Table 1: Example Item Response History

item	irm	rpc.length	cycle	price	response	cum.response	nr	cum.nr
A01	201309	1	M	113.1	1	1	0	0
A01	201310	2	O	113.1	1	2	0	0
A01	201311	3	O	113.1	1	3	0	0
A01	201312	4	M	89.5	1	4	0	0
A01	201401	5	O	79.99	1	5	0	0
A01	**201402**	**6**	**O**	**79.99**	**1**	**6**	**0**	**0**
A01	201403	7	M		0	0	1	1
A01	201404	8	O		0	0	1	2
A01	**201405**	**9**	**O**		**0**	**0**	**1**	**3**
A01	201406	10	M	82.33	1	1	0	0
A01	**201407**	**11**	**O**	**82.33**	**1**	**2**	**0**	**0**
A01	201408	12	O		0	0	0	0
A01	201409	13	M		0	0	1	1
A01	201410	14	O		0	0	1	2
A01	201411	15	O		0	0	1	3
A01	201412	16	M		0	0	1	4
A01	**201501**	**17**	**O**		**0**	**0**	**1**	**5**

Note: irm is index reference month. response & nr are indicator variables equal to one if a price is and isn't reported, respectively, in the given index reference month; cum.response & cum.nr are cumulative number of months with prices reported and not reported, respectively, through the current irm. Cycle is the repricing cycle of the item. M indicates that a price was requested. Rpc.length indicates the time, in months, the item has spent in the PPI survey. Unrefined data frame contained approximately 4.4 million observations.

Because survival analyses require a fixed start and end point, the data used for this analysis is comprised solely of items that were initiated between January 2009 and December 2013. However, not all items enter the survey at the same time, so a total history of at most 60 months (or five years) has been superimposed. Five years was chosen because typically respondents are asked to provide data for five to seven years. The data used for analysis spans January 2009 through July 2016. The latter date was chosen because the PPI employs a four month revision window, during which additional information and data may be provided. Using item histories outside of the revision window is done in order to minimize censoring, thus shifting the focus to observed survival times. The data was further refined according to the goal of estimation. For estimating the length of response, or time-to-nonresponse, observations resembling the sixth and eleventh lines of Table 1 were compiled for each item. For estimating the length of nonresponse, observations similar to the ninth and seventeenth row of Table 1 were compiled for each unique value of the item variable. Data was compiled from all sectors with the exception of NAICS[4] sectors 11, 22, 55, 61, and 81. Sector 11 was not included due to being heavily weighted by item level data collected through alternative survey methods. Sector 22 was not included because items in this sector follow a discontinuation procedure that is different from the one mentioned above. Sector 55 was not included because there are no PPI indexes included in this sector. Sector 61 was not included because the PPI covers only one NAICS industry in this sector, and Sector 81 was not included due to a low number of observations.

4. Methodology

The goal of survival analysis is to estimate the length of time, T, until a clearly defined event (e.g. failure to provide prices) occurs. Survival analysis is best suited to data with distinct characteristics such as clearly defined beginning and end times, censoring and skew. Right censoring occurs when the event of interest has not occurred as of the end time, and skew occurs partly due to T being strictly non-negative. Using the cumulative distribution function (cdf) for time T, $F(t) = \Pr\{T \leq t\}$, we arrive at the probability that the event in question has occurred prior to time t. The survival function, $S(t) = \Pr\{T > t\}$ is the complement of the cdf, such that $S(t)$ gives the probability that the event does not happen by time t. Put another way, it is the probability that the event in question will take more than t days to occur, though t could be measured in days, hours or any unit of time measurement. In this case time is measured in consecutive months with a price reported. $S(t)$ is often referred to as the reliability function, which in this context may be a better interpretation. Within the context of data resembling line 5 of Table 1, $S(t)$ would give the probability of an item providing at least t consecutive months of data without interruption and, respectively, the probability of an item being in nonresponse status at least t months.

The Kaplan Meier (KM) estimator works as follows. Assume the existence of n observations or records of the underlying survival time T, with $m \leq n$ recorded times at which nonresponse occurs. Following this, a rank ordering of survival times, $t_{(1)} < t_{(2)} < \cdots < t_m$ is performed. Let the number of items at risk for nonresponse at time t be denoted by n and the observed number of items that become unresponsive denoted by d. The KM estimator is defined by the formula,

$$S(t) = \prod_{t_i \leq t} \frac{(n_i - c_i - d_i)}{n_i}$$

where

[4] North American Industry Classification System (NAICS)

$$S(t) = 1 \; if \; t < t_{(1)}$$

and c_i denotes the number of censored observations at time t. It should be noted from the above formulation that in the event of censoring, the numerator is reduced, thereby inflating the survival probabilities. As a result, censoring should be avoided if possible, by selecting fixed start and end times.

$S(t)$, when viewed graphically, produces a smooth, non-increasing curve in theory but in actuality is more like a downward sloping step function. Each downward step represents an occurrence of the event in question—in this case failure to provide data. Conversely, if we assume there is some interval of time ε, then a modified hazard function $k(t)$ can be defined by the probability that an item enters nonresponse status between t and $(t + \varepsilon)$ months of consecutive responses and can be calculated as:

$$k(t) = \frac{S(t) - S(t + \varepsilon)}{S(t)}$$

Because $S(t)$ is in the denominator, $k(t)$ is interpreted as the conditional probability of failure given that it has not yet occurred at time t. To give another interpretation, it is the probability that an item that has been responsive for t months becomes nonresponsive between t and $(t + \varepsilon)$ consecutive months of repricing. Applied to the question of nonresponse length, $k(t)$ would provide the likelihood that an item which was in nonresponse status for t months began providing data during some later, clearly defined interval of time.

Based on $S(t)$, different quantiles for the distribution of survival times can be established. The median survival time, for example, is defined as the time at which half the risk set (50%) has experienced the event, or when $S(t) = 0.50$. Similarly, the expected survival time or mean can be estimated by calculating the area under the survival curve, or integrating $S(t)$ itself such that,

$$E(t) = \int_0^\infty S(t)$$

The upper limit for the integral can vary according to industry, sector or some method of grouping, or be a point common to an entire family of different curves. Using data similar to Table 1, lower bounds can be established for length of nonresponse and response length through the use of Survival Analysis.

Because the data used for estimation includes multiple records for certain items, the framework must be adjusted to account for recurrent events. To this effect, the KM curves and derived hazard functions will be based on a cox proportional hazards model employing a robust sandwich variance estimator to account for clustering within each unique item. The cox model is specified as:

$$h_i(t) = \; h_0(t)\exp(\beta_i x_{it} + \cdots + \; \beta_i x_{ik}),$$

where $h_i(t)$ is the hazard function for the i^{th} unit or item, $h_0(t)$ is the baseline hazard and β a vector of coefficients for relevant covariates x_i.

5. Results

5.1 Response Length

Figure I: PPI Continued Response Probability, Items initiated 2009 - 2013

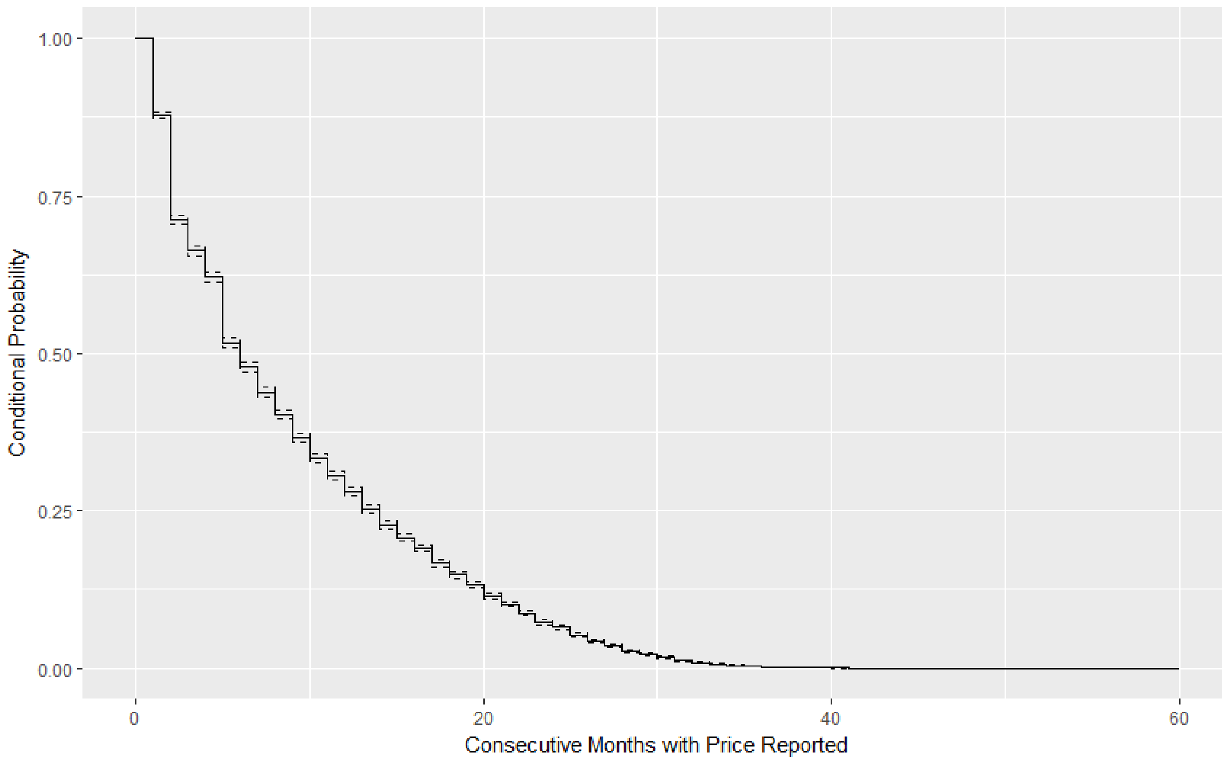


Figure I above plots the probability of continued response for an average item under consideration. Approximately 12 percent of all incidences of nonresponse occur after providing data for a single month. Part of this is due to certain respondents dropping out after initiation, as mentioned earlier. Twenty eight percent of nonresponse cases occur after providing prices for two months. Thirty five percent of nonresponse cases occur after three months, and thirty eight percent after four months. The median time to nonresponse is six months, which means that half of all incidences of nonresponse occur prior to a respondent providing data for six consecutive months, while 75 percent of nonresponse cases occur prior to 13 consecutive months. Therefore, roughly a quarter of all incidences of nonresponse occur after data is provided for 13 consecutive months. The median figure amounts to half of a year's worth of consecutive responses. The mean time to nonresponse is estimated to be nine months, indicating that on average, considering all incidences of nonresponse for items initiated between 2009 and 2013, an item will have prices reported for nine months before they enter nonresponse status.

A useful question involves whether or not there exists seasonal or other variation in nonresponse patterns. Table 3 below provides the results for a cox proportional hazard model fitted with dummy variables representing three of four seasons, in addition to dummies for respondent burden and item weight. Seasons were defined by their index reference month, or the fifth and sixth characters in their "irm" value, according to when the first day of a particular season typically falls. For example, the first day of winter occurs in December, and the first day of spring occurs in March, therefore, December through February are classified as winter months, March through May are considered spring months and so on. Respondent burden refers to the number of items requested from a respondent. Item weight is a measure of importance in index estimation. Due to the wide range of values in both variables, items were broken into categories for both, listed in Table 2 immediately below.

Table 2: Item breakdown by weight and respondent burden, Response Length Sample

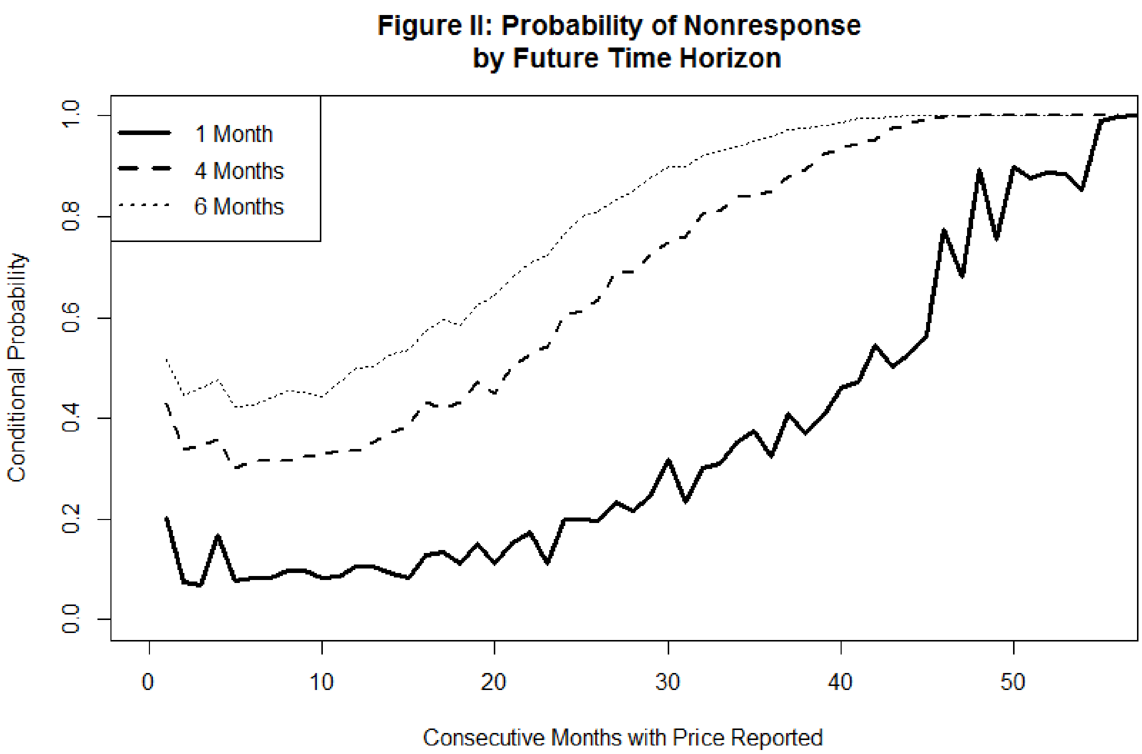
Weight quintile ($)	**Class**	**N**	**Uncensored**	**Censored**
Weight <= 4,085,115	1	5206	5202	4
4,085,115 < Weight <= 11,422,200	2	6411	6401	10
11,422,200 < Weight <= 25,420,000	3	6532	6507	25
25,420,000 < Weight <= 69,395,904	4	6360	6359	1
Weight > 69,395,904	5	5583	5581	2
Burden quintile	**Class**	**N**	**Uncensored**	**Censored**
<= 3 Items	1	3260	3237	23
4 items	2	11082	11063	19
5 -7 Items	3	9902	9902	0
8-10 Items	4	4715	4715	0
> 10 Items	5	1133	1133	0

Table 3: Results for Augmented Cox PH Model for Response Length

	coeff	hazard ratio	robust se	chi square	pr > Chi. Sq
winter	**-0.1551**	**0.856**	**0.0165**	**-9.38**	**<<0.00001**
spring	**-0.0959**	**0.908**	**0.0164**	**-5.83**	**<<0.00001**
summer	**-0.1167**	**0.889**	**0.0156**	**-7.44**	**<<0.00001**
Burden Class 2	-0.0273	0.973	0.0206	-1.32	0.185
Burden Class 3	**0.0898**	**1.094**	**0.0214**	**4.18**	**<<0.00001**
Burden Class 4	**0.1416**	**1.152**	**0.0247**	**5.72**	**<<0.00001**
Burden Class 5	**-0.4275**	**0.652**	**0.0414**	**-10.32**	**<<0.00001**
Weight Class 2	-0.0143	0.986	0.0195	-0.73	0.463
Weight Class 3	0.0096	1.009	0.0199	0.48	0.630
Weight Class 4	0.0092	1.009	0.0204	0.45	0.654
Weight Class 5	0.0193	1.019	0.0216	0.89	0.373
Likelihood ratio test	434.7on 11 df				p<< 0.0001
Wald Test	394.6 on 11 df				p<< 0.0001
Robust Score test	382.1 on 11 df				p<< 0.0001
Dep. Var: cum.response; N = 30,092 with 30,050 occurrences of nonresponse; model stratified by NAICS sector.					

By the hazard rates in Table 3, nonresponse is 11 percent less likely to occur in the summer relative to fall, 14 percent less likely in the winter, and roughly nine percent less likely in the spring. After controlling for respondent burden, aside from what is clear seasonality in the occurrence of nonresponse, it appears that statistically there is no significant effect for item weight. In fact, it appears that respondent burden plays a much larger role in the occurrence of item nonresponse, judging by the hazard ratios for groups three and four. Statistically, these hazard ratios imply that items from respondents that are asked to provide data for five to ten items are between nine and fifteen percent more likely to enter nonresponse status relative to items from respondents asked to provide data for three or fewer items, respectively. The hazard ratio for group five suggests that items from respondents that are asked to provide data for more than 10 items are about 35 percent less likely to enter nonresponse status, compared to those with the smallest burden. One possible reason for the discrepancy is that many respondents with the largest burdens often provide data directly to an analyst via encrypted communication rather than through the PPI's online data collection system, or have a government affairs office which is responsible for interacting with the Federal Government.

Using the estimates of the survival curve in Figure I, Figure II below plots $k(t)$ for differing values of t.



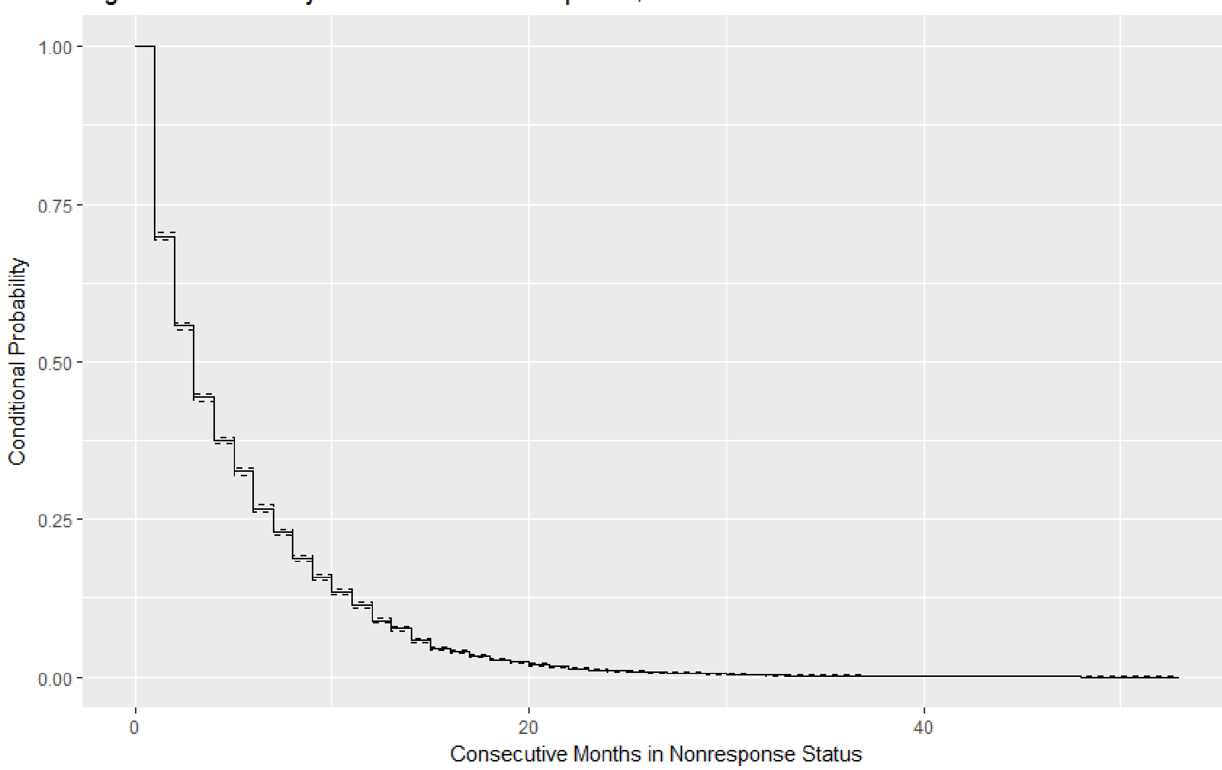
The ranges used to calculate the probabilities in Figure II are best understood via example. At 20 months on the x-axis, the six month curve provides the probability that an item will enter nonresponse status after providing data for 20 to 26 months, conditional on having provided data for 20 months consecutively. What is also clear from Figure II is that, irrespective of months reported, the likelihood that an item enters nonresponse status later on is greater than the likelihood of them entering nonresponse status in the immediate future. One interpretation of this is that over time certain conditions change: a particular point of contact may change positions within their company, review responsibilities may shift between analysts, and so on.

The one month curve in Figure II is largely flat until about 15 months, after which it rises steadily, eventually reaching one at 60 months. This is to be expected: as the corresponding survival curve in Figure I approaches zero, $k(t)$ will approach one. The four and six months curves are different, and somewhat mirror one another. This suggests that when considered over a short—say, month to month—time horizon, PPI respondents are quite reliable, up to a point, since the probability of nonresponse is largely uniform on a month-to-month basis, through about 15 months. However, as the number of consecutive responses rises, so, too, does the probability of nonresponse. This suggests that respondent fatigue may begin to occur around the 15th consecutive response. Overall Figure II suggests that roughly half of all items will spend some time in nonresponse status after providing data between one and seven consecutive months.

5.2 Nonresponse Length

Focusing on the set of items for which there was not a price reported in a given reference month, estimates can be arrived at for length of nonresponse and the likelihood of return from nonresponse. Figure III below plots the conditional probability of continued nonresponse, using 27,991 records from 18,070 unique items.

Figure III: Probability of Continued Nonresponse, Items initiated 2009-2013



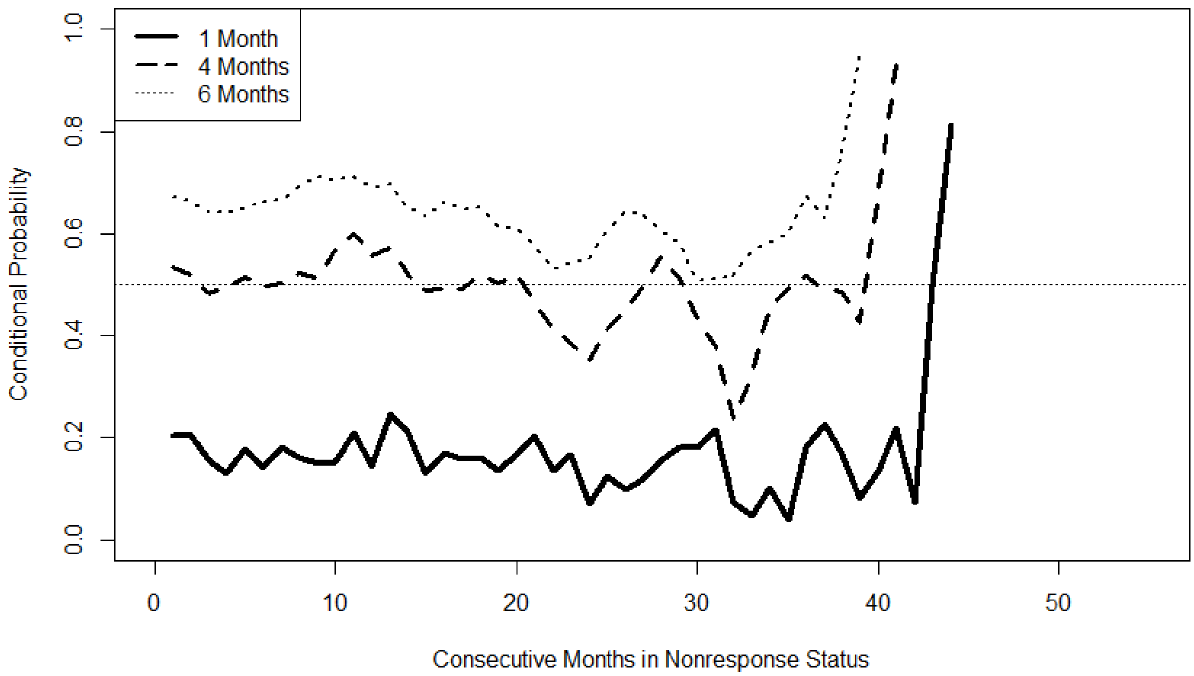
Approximately 30 percent of all incidences of nonresponse are successfully resolved after a month, while 45 percent are resolved after two months, 62 percent after four months, 81 percent after 8 months and 99 percent after 22 months. The mean time in nonresponse is five months, and the median is three months, indicating that, half of all cases of nonresponse lasted less than three months. The data set used contained 27,991 unique items, representing 428 different six-digit NAICS industries, and 14 of 20 NAICS sectors. Table 4 below outlines the distribution of items by weight class and respondent burden. The results of a Cox proportional hazard model applied to time in nonresponse status are displayed in Table 5. The hazard rates suggest that, effectively speaking, weight has a minor effect on time in nonresponse. More specifically, it appears that items from the middle of the weight distribution are more likely to resume providing data earlier than others. However, it is also abundantly clear that respondent burden is a better indicator of whether a respondent will resume providing data after a period of nonresponse. What Table 4 and Table 5 show is that, for a given period of nonresponse, respondents that are asked to provide data for fewer items are more likely to resume providing data for given item. Additionally, the seasonal dummy variables indicate that season has no significant effect on nonresponse length.

Table 4: Item breakdown by weight and respondent burden, Nonresponse Length sample				
Weight quintile ($)	**Class**	**N**	**Uncensored**	**Censored**
Weight <= 4,085,115	1	4764	4738	26
4,085,115 < Weight <= 11,422,200	2	6048	6019	29
11,422,200 < Weight <= 25,420,000	3	6045	6030	15
25,420,000 < Weight <= 69,395,904	4	5889	5881	8
Weight > 69,395,904	5	5228	5245	17
Burden quintile	**Class**	**N**	**Uncensored**	**Censored**
<= 3 Items	1	2988	2942	46
4 items	2	10044	10020	24
5 -7 Items	3	9363	9346	17
8-10 Items	4	4580	4572	8
> 10 Items	5	1016	1016	0

Table 5: Results for Augmented Cox PH Model for Nonresponse Length					
	coeff	hazard ratio	robust se	chi square	pr > Chi. Sq
winter	0.0218	1.0220	0.0169	1.29	0.197
spring	0.0113	1.0114	0.0186	0.61	0.542
summer	0.0190	1.0192	0.0172	1.10	0.270
Burden Class 2	**-0.0785**	**0.9245**	**0.0228**	**-3.45**	**0.0005**
Burden Class 3	**-0.2050**	**0.8145**	**0.0231**	**-8.86**	**<<0.00001**
Burden Class 4	**-0.2589**	**0.7718**	**0.0261**	**-9.91**	**<<0.00001**
Burden Class 5	**-0.0867**	**0.9169**	**0.0442**	**-1.96**	**0.0499**
Weight Class 2	0.0385	1.0393	0.0205	1.88	0.060
Weight Class 3	**0.0532**	**1.0546**	**0.0209**	**2.54**	**0.011**
Weight Class 4	0.0020	1.0020	0.0210	0.09	0.924
Weight Class 5	0.0356	1.0362	0.0233	1.52	0.126
Likelihood ratio test	198.5 on 11 df				p<< 0.0001
Wald Test	179.3 on 11 df				p<< 0.0001
Robust Score test	179.1 on 11 df				p<< 0.0001
Dep. Var: cum.dlq; N = 27,991 with 27,896 occurrences of return from nonresponse; model stratified by NAICS sector.					

Figure IV below plots the conditional probability of return from nonresponse for a series of finite time horizons differing in length. It is analogous to Figure II above in that it provides the probability that a respondent will begin providing data again during a fixed time interval, given a predetermined number of months in nonresponse status. For example, the four month curve, at 20 months on the x-axis, provides the probability that an item that was in nonresponse status for 20 months will return within the next four months of nonresponse.

Figure IV: Probability of Return from Nonresponse, by Future Time Horizon



What is clear from Figure IV is that an item is more likely to return later than sooner, judging by the distance between the curves. One reason for this is that over time multiple calls are made, often by more than one analyst. Overall, this suggests that time and persistence are most effective in attempting to resolve incidences of nonresponse. The horizontal line in Figure IV is drawn at the point where the probability of return is equal to the probability of continued nonresponse—the point of even odds for both states. Clearly, the point is different, on the x-axis, depending upon which time horizon is being considered on the x-axis. Over a four month horizon, the probability of return within four months of continued nonresponse is 0.53 after one month in nonresponse status and 0.51 after two months. The four month curve is roughly at or above the even odds mark for the first 20 months in nonresponse status. What this means is that if a respondent is to begin providing data again, for a given item, within a four month correction window, it should happpen within the first twenty months of nonresponse, after which the probability falls below 0.5 and quickly approaches the tails of the distribution. The four and six month curves begin to turn up around and approach one at about 30 months on the x-axis. This is because the curves listed in Figure IV are derived from those in Figure III: absent censored data in the tails of the distribution of survival times, as the survival curve in Figure III approaches zero, those in Figure IV will begin to approach one, as is shown with the four and six month curves. Therefore, the interpretation is that more than 50 percent of all cases of nonresponse that last 40 or more months end within four months, while nearly 80 percent of those cases end within six months. These estimates appear to be relatively high because items are typically dropped after 18 consecutive months in nonresponse status, so very few items are given the opportunity to reach 40 months in nonresponse status. As mentioned above, less than one percent of all incidences of nonresponse last more than 22 months. Nevertheless, Figure IV does provide evidence that long-term cases of survey nonresponse can be resolved. Of the items in question, most of them were from the largest weight group.

5.3 Comparison

Table 6 below lists the adjusted mean response length and nonresponse length for 14 of the 20 unique NAICS sector groups represented in the PPI family of indexes. The table also lists the ratio of response length to nonresponse length for the mean.

Table 6: Mean & Median Response & Nonresponse Length, by NAICS Sector

		response length, months		nonresponse length, months		
Sector	Title	Mean	Median	Mean	Median	Mean ratio
44-45	Retail Trade	13.35	8	3.49	2	3.83
42	Wholesale Trade	12.08	6	3.2	2	3.78
48-49	Transportation and Warehousing	13.42	7	3.75	2	3.58
53	Real Estate and Rental and Leasing	12.19	6	4.13	2	2.95
71	Arts, Entertainment and Recreation	11.94	9	4.33	3	2.76
52	Finance & Insurance	8.5	3	3.3	2	2.58
62	Health Care and Social Assistance	14.75	7	6.06	3	2.43
56	Administrative and Support and Waste Management and Remediation Services	11.09	5	4.56	3	2.43
54	Professional, Scientific and Technical Services	15.61	9	6.86	5	2.28
31-33	Manufacturing	11.8	5	5.38	3	2.19
51	Information	10.16	2	5.46	2	1.86
21	Mining, Quarrying, and Oil and Gas Extraction	10.7	2	6.95	5	1.54
23	Construction	8.13	5	5.32	3	1.53
72	Accommodation and Food Services	11.63	7	8.4	5	1.38

The table shows that each two digit NAICS sector has its own distribution of time-to-nonresponse as well as time-in-nonresponse, though many mirror one another in certain respects. How each distribution is evaluated depends on what point on the curve one considers. A reasonable measure of central tendency is the median, which provides the number of months at which a given survival curve equals 0.50. For sector 52, Finance and Insurance, half of all incidences of nonresponse occur after 3 consecutive responses where half of said incidences last for less than two months. As discussed earlier, the mean represents the area under each survival curve and it takes into account all of the data available for a given sector. By this, a larger mean value would indicate a wider range of values for response and/or nonresponse length, but also the extent of censoring. However, since each survival curve starts at one month of response or nonresponse, and the maximum value is capped at 60 months, what the mean essentially provides is an indicator of which sectors have especially long tails in their distribution and the degree of censoring. As mentioned above, there is censoring present in the data used for estimation, but it is certainly not excessive, as shown in Table 2 and Table 4. For example, in the case of response length, the mean values across all sectors are strong, indicating that at least six months of consecutive prices can be expected before an item enters nonresponse status.

Overall, the variability in the mean estimates is greater than that of the median estimates. Similarly, the variability in the response length statistics is greater than those of nonresponse length, indicating more cases of long response lengths relative to nonresponse lengths. This is to be expected: there are no procedures in place to handle long response lengths, as they would serve no purpose, but the PPI has, in recent years, implemented several to address the problem of long nonresponse lengths. As mentioned above, respondents that have failed to provide data following 12 requests are typically dropped from the survey, unless there is sufficient justification for them to remain; these respondents, and therefore items, are discontinued at 18 requests if they remain in nonresponse status. In short, the procedures implemented are responsible for shaping the data used in the analysis, making them part of the data generating process. That the distribution of nonresponse lengths is relatively tight is a reflection of such procedures.

Figure V: Mean Response & Nonresponse Length, by NAICS Sector

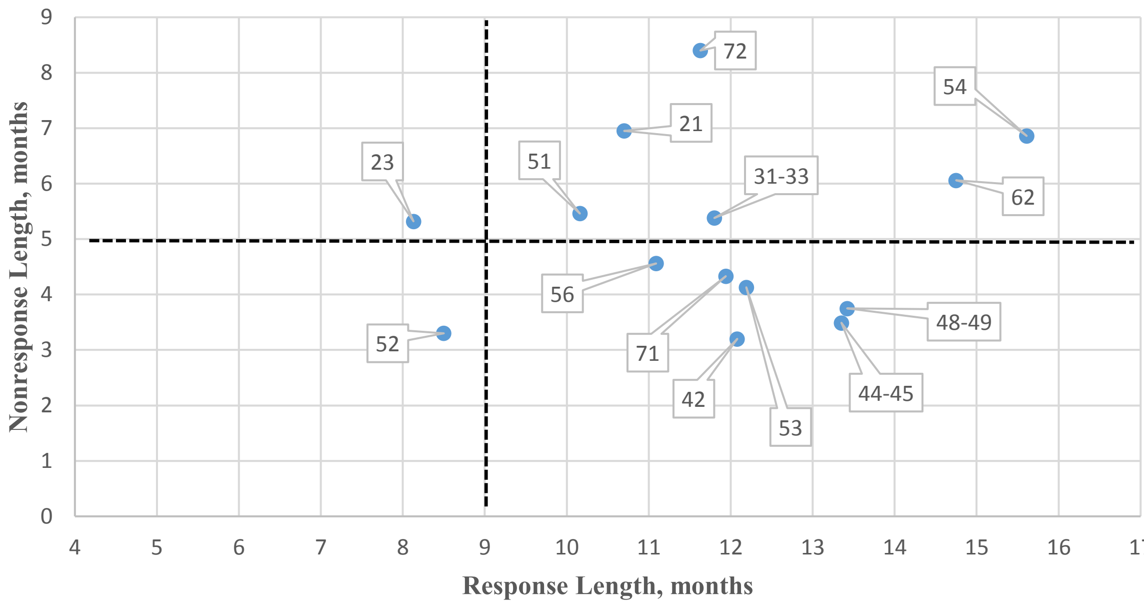


Figure V provides a straightforward, graphical way of assessing the information in Table 6 based on the ordered pair that each dot, or sector, represents. The two black lines, at nine months on the x-axis and 5 months on the y-axis, represent the mean response and nonresponse lengths. Viewed on a quadrant-by-quadrant basis, Figure V presents which sectors contain the best respondents, judging by their place in the x-y plane, as it represents item level nonresponse behavior. Those in the upper left quadrant, e.g. sector 23, are those with relatively low time-to-nonresponse and high time-in-nonresponse. They represent items with relatively short times between incidences of nonresponse, but also those that stay in nonresponse status for relatively long periods of time. Note that these items do not perform particularly well in both nonresponse and response length. Conversely, those in the bottom right quadrant represent sectors with items that provide relatively large amounts of data before entering nonresponse status, which doesn't last long. Sectors in the upper right quadrant are high-benefit, high-cost, and those in the bottom left are low-benefit, low-cost. The final column in Table 6 provides the ratio of mean response length to mean nonresponse length. It can be interpreted as a basic cost-benefit measure. For example, for every month in nonresponse status, respondents from sector 42 provide data for roughly four months in a row. Overall, Figure V suggests that, on an item-by-item basis, at the sector level, respondents from most sectors represented by the PPI family of indexes spend more time providing data than they do not.

6. Conclusion

The goal of this paper was to provide statistical estimates of the amount of time, in consecutive months, until the occurrence of nonresponse and, conversely, the time in nonresponse. This was done via survival analysis. It is estimated that the mean time to nonresponse is 9 months, with a median of 6 months, and the time in nonresponse is estimated at a mean value of 5 months and a median of 3 months. It was also shown that each two–digit NAICS sector is characterized by its own distribution of response and nonresponse length, such that survival curves were estimated for 14 of the 20 sectors represented by the PPI family of indexes.

The results presented above provide valuable context regarding both the length of nonresponse and what may be driving it. Table 2 through Table 5 provide compelling evidence regarding the distribution of items by weight and overall respondent burden. In short, it appears that most respondents are asked to provide data for between four and ten items. Furthermore, the data suggests that nonresponse is most likely to occur if an item is from a respondent who is asked to provide data for between five and ten items. This is especially compelling evidence because item weight was not found to be a statistically significant predictor of an item entering nonresponse status, after controlling for respondent burden.

In addition to the respondent burden and item weight, there are still many other factors, only some of which may be quantifiable, that play a role in the occurrence of nonresponse and return from nonresponse status. For example, one factor that may be quantifiable is if there has been a change in the point of contact for PPI. If the individual reporting data leaves their company, periods of nonresponse are almost certain to follow. One factor that may be just as important as respondent burden is attitudes toward government. If the individual tasked with providing data to the PPI has, for example, a negative attitude toward government surveys then a potentially extended period of nonresponse may also be likely to follow.

Regarding nonresponse length, Table 5 provides evidence that an item's "importance" does not make its period in nonresponse status more likely to end, but rather respondent burden does. The evidence in Table 5 suggests that items from the middle of the weight distribution are more likely to return, while also suggesting a fairly straightforward relationship between respondent burden and the length of time in nonresponse status. For the vast majority of items, smaller respondent burden is associated with a greater likelihood of ending a period in nonresponse status. Coupled with the results from Table 3, it appears that the best way to prevent items from entering nonresponse status, and to ensure that the amount of time spent in nonresponse status is short, is to aggressively seek ways to minimize respondent burden.

Given these results, a question remains regarding how best to address the problem of PPI nonresponse. One common method would be to employ the use of algorithmic tools via Machine Learning. Such tools could produce any number of outputs that could be used to structure a more targeted nonresponse procedure. For example, one such question involves the optimal point of deletion for a respondent in nonresponse status: for a given number of months in nonresponse status, is there a point after which, statistically speaking, a respondent should no longer be contacted? To account for factors such as respondent burden and item weight, a decision tree algorithm could provide a simple "yes/no" answer based on an item's place in the distribution of any variables of interest.

Overall, the above results provide a useful starting point for analytical approaches to the more general problem of survey nonresponse. Given the estimates above, a natural question is whether or not the same patterns hold upon aggregation of item-level data. Research into this matter is currently underway via the use of Multi-State Markov, semi-Markov and survival models, which will be employed to estimate transition probabilities for each state of aggregate response and nonresponse. For example, if a respondent is asked to provide data for four items, and two of them are in nonresponse status, what is the likelihood that they will all be in nonresponse status the following month? These and similar questions are currently being investigated and will be the subject of a later paper.

Acknowledgements

Andy Sadler, Kathleen Frawley, Melanie Santiago, Bonnie Murphy, Jeff Hill and Steve Paben all provided valuable comments and suggestions. Any errors or omissions are my own.

References

Chopova, B., Hesley, T., Kelly, G., Paik, S., Sadler, A. and Slack, D. 2012. "Analyzing Nonresponse Bias in PPI Data." Retrieved from: https://www.bls.gov/osmr/pdf/st120080.pdf

Hosmer, David W. and Lemeshow, Stanley. 1999. *Applied Survival Analysis: Regression Modeling of Time to Event Data.* New York: John Wiley.

U.S. Census Bureau. 2012. North American Industry Classification System (NAICS). Retrieved from: https://www.census.gov/eos/www/naics/

U.S. Dept of Labor, Bureau of Labor Statistics. 2016. BLS Handbook of Methods. Chapter 14, "Producer Prices." Washington, DC. Retrieved from: https://www.bls.gov/opub/hom/pdf/homch14.pdf