Evaluating the Impact of the Loss of Telephone Follow-Up (TFU) Operations on the Service Monthly and Annual Sample Maintenance (SMASM) Quarterly Birth Sampling

Micah Tyler, Michael Owen, Franklin Sell
Katrina Washington[1]
[1]U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746

Abstract
The Business and Professional Classification Report (SQ-Class) is a survey whose data is used in quarterly sample maintenance operations for a subset of current business surveys at the Census Bureau. These quarterly sample maintenance operations supplement the current business surveys with business births and help ensure proper industry classification in the Economic Census. Companies selected into sample for the SQ-Class receive a form that requests two months of sales or receipts data, as well as industry classification and company structure information. This data is used to conduct a second round of sampling, with selects being added to their appropriate current business survey sample(s). Over the years, the amount of money required to conduct the SQ-Class has increased steadily, with TFU operations cited as one of the costliest aspects. To control cost, the idea of removing TFU while compensating for the loss of additional response data by increasing the total sample size was proposed. This paper outlines the research conducted to assess both the feasibility of increasing the sample size, and the impact that a loss of TFU operations would have on SMASM's quarterly birth sampling.

Key Words: two-phase sampling, telephone follow-up, industry classification, measure of size, business births, sample maintenance

Disclaimer: Any views expressed are those of the author(s) and not necessarily those of the U.S. Census Bureau. All counts and percentages presented are rounded values, per disclosure avoidance guidelines. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the application of disclosure avoidance. (Approval ID: CBDRB-FY20-ESMD003-010)

1. Introduction

The U.S. Census Bureau conducts several business surveys that measure the nation's economic activity for the retail, wholesale, and service sectors of the economy on an annual, quarterly, or monthly basis. Samples for these surveys are selected approximately once every five years as part of a process known as the Business Sample Revision (BSR). In an effort to ensure that these samples are representative of their respective target populations throughout their approximate five-year lifespan, the samples are updated quarterly to reflect the changes taking place in the overall population. These updates are done so that the samples can reflect new businesses that come into existence (referred to as births), existing companies that become inactive or go out of business (referred to as deaths), companies or parts of companies that change ownership or association, and companies that go out of business and then become active again (referred to as

reactivations). The Service Monthly and Annual Sample Maintenance (SMASM) system is a two-phase sampling process designed to identify, track and apply these changes, based mostly on administrative updates to the Census Bureau's Business Register (BR) and the results of the Business and Professional Classification Report (SQ-Class).

The Business and Professional Classification Report is a quarterly survey whose results are used to supplement retail, wholesale, and service annual and sub-annual surveys with business births and to help ensure proper industry classification in the Census Bureau's Business Register. The Business Register is a master list of businesses and a data resource for business statistics. Every quarter, a list of new or reactivated companies is sent to the Business Register from the IRS. Each company, whether it be an employer, sole proprietor, corporation, partnership, estate, trust, individual, or other entity, has applied for an employer identification number (EIN) for tax filing and reporting purposes. The data received from the IRS, which we refer to as administrative data, is linked to each company's EIN. EINs that are in-scope to our surveys are extracted from the Business Register, and are subjected to first phase sampling. There are approximately 14,000 EINs selected in the first phase sample using stratified systematic sampling with strata based on their North American Industry Classification System (NAICS) code and quarterly administrative payroll. EINs are selected into the Business and Professional Classification Report and receive a form that requests two months of sales or receipts data and establishment, as well as classification and company structure, information to ensure proper NAICS classification at the 5- or 6-digit level. This data is then used to create a measure of size for second phase sampling. EINs that are eligible for inclusion into the retail, service, or wholesale surveys are subjected to second phase sampling the subsequent quarter and selected EINs are added to their appropriate current business survey sample(s).

2. Problem Description

First phase sampled EINs, also referred to as firms, are mailed a letter with a username and password to use to report/complete the Business and Professional Classification Survey. Reports not received by the scheduled due date, which is typically 30 days from mail-out, are considered delinquent. Reminder letters are sent approximately one week before the due date. Telephone follow-up (TFU) begins approximately six weeks after the initial mailing to respondents and about three weeks after the reminder letter has been sent. TFU runs for six weeks. During the 6 weeks of TFU, analysts at the National Processing Center (NPC) conduct research on firms to obtain a phone number. Once a phone number is obtained, analysts call delinquent firms and attempt to collect data over the phone.

Over the years, the cost required to conduct the Business and Professional Classification Report has increased steadily, with telephone follow-up (TFU) operations cited as one of the costliest aspects. In an effort to get cost under control, the idea of eliminating TFU while compensating for the loss of additional response data by increasing the total sample size has been proposed. In order to determine whether it is feasible to consider this, an evaluation is needed to assess the impact that a loss of TFU operations would have on the second phase sample.

3. Analysis

Due to time and budget constraints, an actual experimental design and implementation was considered out-of-scope to this research. However, as a result of the United States federal

government shutdown of 2018-2019, a de facto implementation did occur for quarterly births selected into the Business and Professional Classification Report sample in 2018 Quarter 4 (2018Q4) and scheduled for second phase sampling in 2019Q1 since TFU, scheduled for the month of January 2019, was cancelled due to the shutdown.

Without an experimental design, the most appropriate analysis was to conduct simulations of second phase sampling without the inclusion of TFU data. These simulations were run on second phase input files from 2017Q1 up to 2018Q4. Although inventory and ecommerce inquiries are included in SQ-Class, we excluded these items from this analysis due to time constraints. Instead, we focused on the potential impact that eliminating TFU would have on the measures of size, industry classification, and number of births subjected and selected through birth sampling. A couple of assumptions were made in the process.

3.1 Second phase doesn't happen in a bubble
The results from one quarter can have an impact on subsequent quarters. It was possible to build some of this impact into the simulations. For instance, the second phase is conducted using stratified systematic sampling, a procedure that involves a fixed sampling interval but a random starting point. With sampling taking place each quarter, these "random starts" are carried forward from quarter to quarter in order to maintain the sampling interval across quarters. This feature of second phase was likewise built into the simulation; however, it was not feasible to build all such interdependencies into the simulation.

For instance, once an EIN has been selected into the Business and Professional Classification Survey, it will continue to remain eligible for second phase until it either has enough information (measure of size, NAICS classification, etc.) to subject it to sampling or remove it from sampling (e.g. associated with a company already in sample, out-of-scope NAICS, etc.). Until then, it remains eligible to be recanvassed every other quarter until it responds or the Business Register is updated with better classification information. Therefore, in practice, the EIN would continue to be recanvassed until such definitive information is received, whether it's from the Business Register or the survey. However, we did not simulate the recanvassing of any EINs. Thus, no EIN would re-appear in subsequent quarters, and the additional ineligible EINs that would undoubtedly accumulate over the course of the BSR were **not** built into the simulation. While it is possible to model the number of additional recanvassed EINs that would occur as a result, this was not done due to time constraints. Similarly, if an EIN provided data during TFU that would have definitively removed it from sampling, the data was not used in our research, and the EIN was likely subjected to sampling in the simulations.

3.2 We chose to err on the side of caution
A number of liberties needed to be taken with the decision of how to exclude TFU data. Ultimately, we concluded that the most efficient way of handling this was to "blank" all response data received after the scheduled start date of TFU for each quarter, regardless of whether the data was actually collected through TFU or just coincidentally received through our online data reporting tool after the scheduled TFU start date. Furthermore, we blanked all response data, including NAICS and company affiliation data supplied by the staff at NPC. This allowed us to maximize the impact of the loss of this data on the subsequent current survey births.

4. Impact on Number of Births

Using the scheduled TFU dates, we were able to construct a timeline for each quarter of the approximate duration of TFU, seen in Table 1 below. Response data was "blanked" for any EINs having a check-in date on or after the TFU Start Date. In other words, only response data received prior to the scheduled TFU Start Date was used in the second phase simulation runs. Over the first 8 quarters of the most recent BSR (BSR-17), the pre-TFU response rates hovered around 50%. There was a slight drop-off in the pre-TFU response rates between 2017 and 2018. Part of this can been attributed to the migration to a new online reporting platform that took place just prior to the November 2017 initial mail, which added about 15 minutes to the response burden estimate per respondent.

As mentioned at the beginning of this section, a de facto "no-TFU" implementation was realized in 2018Q4 due to the United States federal government shutdown of 2018-2019. In order to provide a comparison with the simulations, approximate response rates were calculated based on an anticipated TFU Start Date of 1/3/19 and the actual NPC closeout date of 2/14/19. During this window, no TFU operations were conducted but response data was allowed to continue to trickle in until second phase sampling was conducted on 2/19/19. The difference between pre- and post- TFU response rates for this quarter illustrates the realized impact of the loss of TFU on response volume.

With the pre- and post-TFU response and the mailed sample count for first phase, we produced an estimate for the percent increase in sample needed to achieve the same response volume without TFU, *% Increase in Sample Needed* in Table 1. The new sample size was determined by the original response volume [1] times the achieved pre-TFU response rate. These numbers varied wildly, but remain above 50% after the introduction of the new online reporting platform. In other words, if we want to achieve the same volume of response data for input to second phase, assuming equivalent pre-TFU response rates, we would need an increase of more than 50% in the initial sample each quarter.

---

[1] Original Response Volume= Original Sample Size*Achieved post-TFU Response Rate

First Phase Quarter	Original Sample Size	Achieved pre-TFU Response Rate	TFU Start Date (mm/dd/yy)	# TFU Working Days	Achieved post-TFU Response Rate	New Sample Size Needed to Produce Original Response Volume	% Increase in Sample Needed
2016Q4	12000	51.5%	12/14/16	33	73.3%	16000	33.3%
2017Q1	11500	52.9%	3/16/17	34	73.9%	15500	34.8%
2017Q2	15000	55.9%	6/15/17	34	76.7%	20000	33.3%
2017Q3	15500	54.4%	9/14/17	35	80.6%	23500	51.6%
2017Q4	11000	45.5%	12/14/17	32	71.8%	17500	59.1%
2018Q1	10000	44.9%	3/15/18	34	75.0%	16000	60.0%
2018Q2	15000	46.6%	6/20/18	30	70.0%	24000	60.0%
2018Q3	13000	46.7%	9/20/18	29	70.8%	21000	61.5%
2018Q4	9500	49.4%	1/3/19	0	57.9%[2]	NA	NA

Table 1: TFU Start Dates, along with response rates (pre- and post-TFU for First Phase).

5. Impact on Second Phase Sampling

The simulations of second phase sampling were run for each quarter to assess the impact of the reduced response rates. This was achieved by removing the response data obtained during the TFU period from input data sets from first phase. Once the input data sets were modified, second phase was rerun once for each quarter. Then, the second phase simulation runs were compared to the production runs. In Figure 1, the blue line represents the actual EINs subjected to second phase sampling from the production run, while the red line represents the hypothetical EINs subjected in the simulation. Counts from both production and the simulation exclude EINs dropped due to an out-of-scope NAICS along with EINs identified as auxiliary or associated with a foreign or government entity, both of which are out-of-scope to the BSR surveys. They also exclude EINs ineligible for second phase due to lack of a NAICS match, lack of data to use to calculate a measure of size (i.e., payroll or revenue), or with a FIPS code outside of the United States.

---

[2] Although no TFU occurred, the response rates were calculated at the end of the scheduled six weeks that were allotted for TFU.
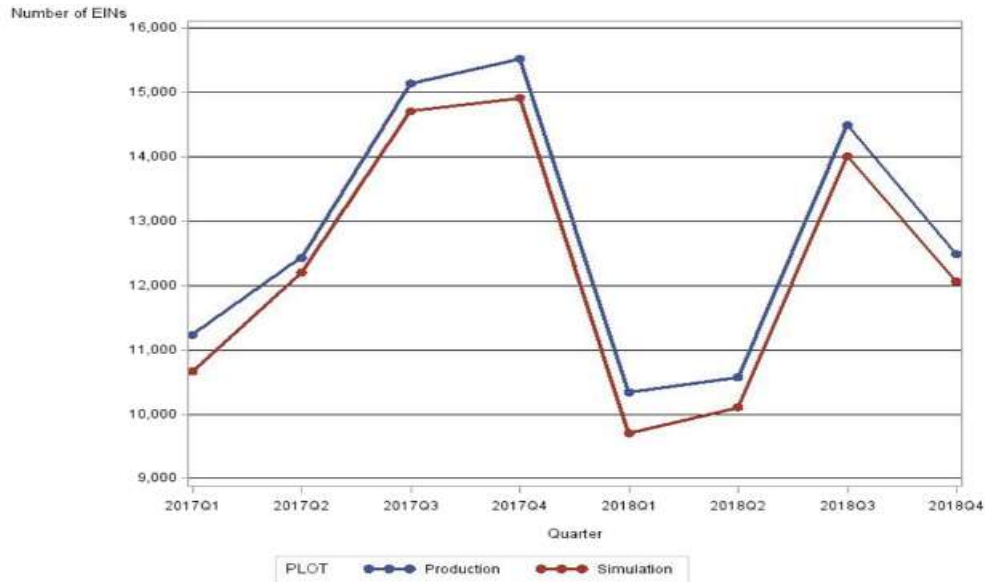
Figure 1: Subjected Counts from Second Phase, 2017Q1 to 2018Q4

The number of EINs subjected between production and simulation are similar, as seen in Figure 1. The simulation counts also follow the same trend as we've come to expect in production, with larger 3rd and 4th quarter counts. However, there is still a gap between production and simulation. The smallest difference happened in 2017Q2 with 250 fewer EINs subjected in the simulation than in production, while 2018Q1 saw the largest difference (650). The average difference is around 450, which, although not shown in these graphs, would have a cumulative effect until the Business Register obtains more detailed NAICS information for these EINs, or until they supply response data through recanvassing.

Likewise, Figure 2 contains the comparison of the number of births added to the surveys (blue line) with the simulated number of births that would have been added without TFU data (red line). The larger gap between production and simulation in 2017Q1-2017Q3 is due to an oversampling of EINs in the Retail sector during that timeframe; otherwise, the counts are fairly comparable. Therefore, we can conclude that, in general, the differences are a direct result of the lower "subjected" counts due to the missing TFU data.
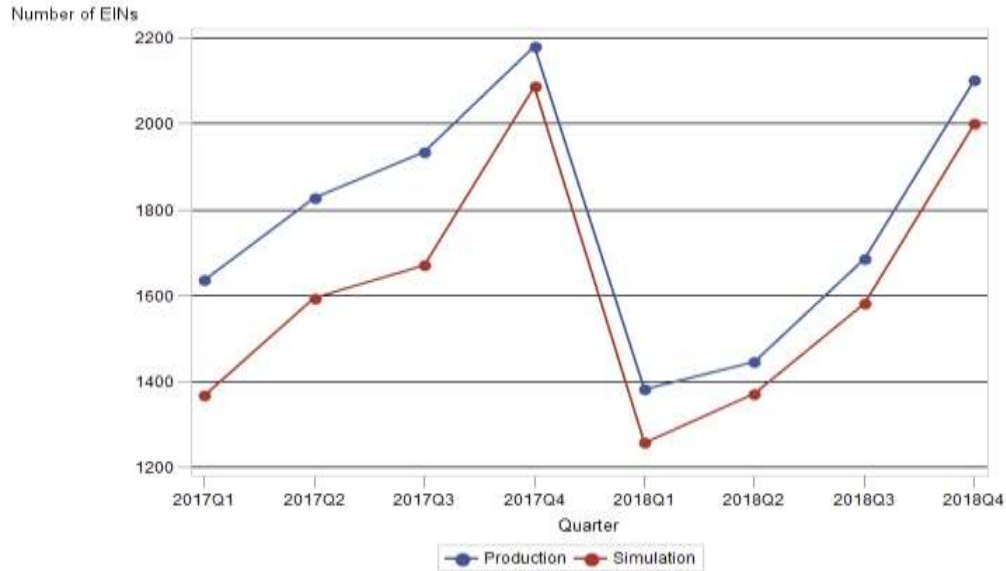
Figure 2: Second Phase select counts from 2017Q1 to 2018Q4.

Likewise, Table 2 further expands on these comparisons by approximating how many births would have been added to each of the BSR surveys each quarter given the lower response rate from the previous quarter's first phase. Production birth counts for 2019Q1 are included for comparison since TFU was not conducted on this sample.

Table 2: Comparison of births added to the surveys, production vs simulation, 2017Q1 to 2018Q4											
Number of Births, by Survey											
Second Phase Quarter	*ARTS*		*MRTS*		*AWTS*		*MWTS*		*SAS*		*QSS*	
	Prod	*Sim*	*Prod*	*Sim*	*Prod*	*Sim*	*Prod*	*Sim*	*Prod*	*Sim*	*Prod*	*Sim*
2017Q1	350	200	200	100	50	30	20	20	1,200	1,100	250	200
2017Q2	400	200	250	100	90	80	50	40	1,300	1,300	300	250
2017Q3	400	200	200	100	50	40	20	20	1,500	1,400	350	300
2017Q4	250	250	150	150	70	60	30	30	1,800	1,800	450	350
2018Q1	200	200	100	90	30	40	20	20	1,200	1,000	250	200
2018Q2	150	200	90	100	40	40	20	20	1,200	1,100	250	200
2018Q3	200	200	100	100	50	30	30	20	1,400	1,400	350	300
2018Q4	250	250	150	150	70	60	30	30	1,800	1,700	450	350
2019Q1	200	*NA*	90	*NA*	30	*NA*	<15	*NA*	1,100	*NA*	250	*NA*

6. Impact of Measure of Size

Also of interest is the amount of estimated dollar volume being added to the surveys through the quarterly birth process. In second phase processing, there are four potential methods that can be used to assign a measure of size, which is an estimation of annual revenue/sales, to an EIN depending on the availability of response data. A quick way to determine which method was used is by looking at the basis code assigned during sampling. The following four basis codes are assigned during second phase:

4 - Measure of size derived from monthly sales reported on the SQ-Class form

5 - Measure of size derived from administrative payroll from the IRS. No monthly sales available

6 - Measure of size derived from administrative payroll from the IRS. Monthly sales are available from the SQ-Class form

7 - Measure of size derived from administrative employment from the IRS. No monthly sales available

Methods 4 and 6 both imply the presence of response data. In method 4, response data is directly used in the calculation of the measure of size. In method 6, the response data is annualized and compared to an annualized payroll-based value with the payroll-based value ultimately being assigned because it is larger. If the response data is unavailable, then the methodology picks between the larger of the payroll-based value (method 5) and the employment-based value (method 7).

Figure 3: Second Phase MOS method comparison between Production and Simulation for 2017Q1 through 2018Q4

As expected, the use of methods 4 and 6 dropped across all quarters due to the lack of response data. As a result, the loss of TFU data would mean that response data is used less frequently in the MOS calculation. Ironically, Figure 4 shows that having less available response data for methods 4 and 6 does not impact the frequency with which response data is used, when available, to calculate MOS.
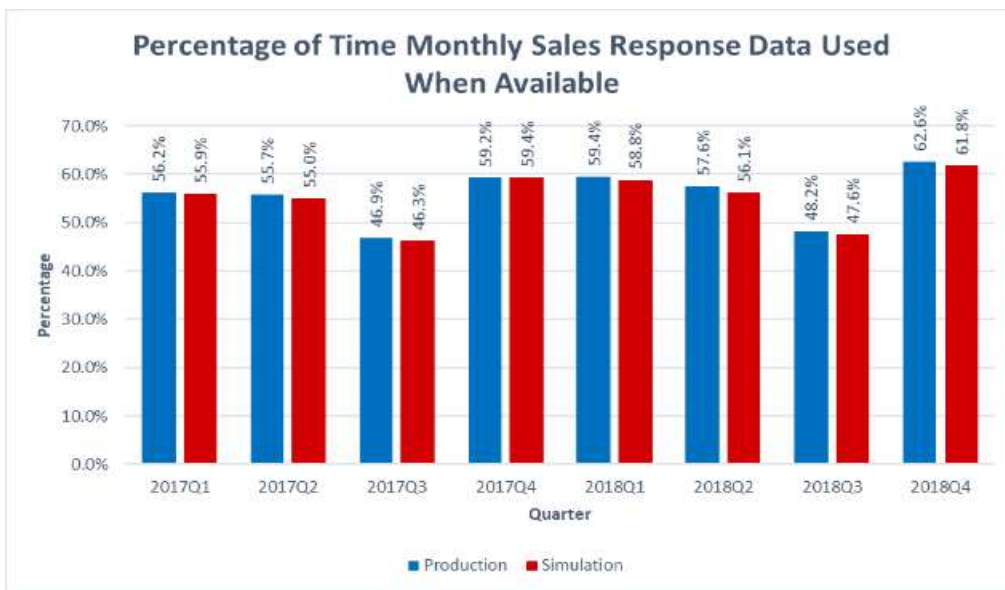

Figure 4: Second Phase MOS method with response data comparison between Production and Simulation for 2017Q1 through 2018Q4

To that end, the next logical step is to compare the MOS assigned in production to the MOS assigned in the simulation. The most efficient means for conducting this comparison is through the stratum that is assigned. After all, the assigned stratum determines the weight that is given. If any differences in the dollar amount of the assigned MOS are negligible enough that the EIN remains in the same strata prior to sampling, then the end result is unchanged with regards to its contribution to the estimate. For instance, stratum 1 for NAICS AAAAAA has a lower bound of $0 and an upper bound of $2 million. If an EIN in NAICS AAAAAA is subjected to sampling in production with a MOS of $500,000 coming from TFU-captured response data, but would have had a MOS of $1.5 million using payroll data in our simulation, the end results are equivalent. The EIN would fall into stratum 1 in both scenarios, have a weight of 88 and, if selected, have the same weight-adjusted revenue in both scenarios.

Table 3, below, summarizes the results of the stratum comparison for each of the eight examined quarters. In order to reduce any outside influences, simulated and production measures of size were only compared for an EIN if its assigned NAICS placed it in the same primary sampling stratum (i.e., NAICS or NAICS grouping, referred to as a recode) in both production and simulation. Table 4 contains a similar analysis regardless of assigned NAICS. From these tables, we can conclude that the MOS assigned in the simulation placed EINs within one stratum of their production stratum between 93 and 96 percent of the time. Again, there is a notable jump between 2017Q3 and 2017Q4. This is due to the retail oversampling issue in which overinflated measures of size were being

assigned prior to 2017Q4. The simulation was done using correct parameters for those three quarters, resulting in more realistic measures of size.

Table 3: Percent of EINs with or without a stratum assignment change between production and simulation for subjected EINs, same NAICS recode.			
Second Phase Quarter	*No Change in Stratum*	*Stratum Change of 1*	*Stratum Change >1*
2017Q1	84.8%	6.1%	9.1%
2017Q2	84.3%	6.1%	9.6%
2017Q3	85.7%	5.0%	9.3%
2017Q4	92.6%	3.3%	4.1%
2018Q1	93.3%	3.3%	3.4%
2018Q2	93.5%	3.3%	3.2%
2018Q3	92.3%	2.7%	5.0%
2018Q4	91.3%	3.5%	5.2%

Table 4: Percent of EINs with or without a stratum assignment change between production and simulation for subjected EINs, regardless of NAICS recode match.			
Second Phase Quarter	*No Change in Stratum*	*Stratum Change of 1*	*Stratum Change <=1*
2017Q1	81.0%	8.1%	10.9%
2017Q2	81.7%	7.9%	10.4%
2017Q3	86.2%	6.9%	6.9%
2017Q4	86.7%	5.7%	7.6%
2018Q1	88.7%	5.2%	6.1%
2018Q2	88.0%	5.5%	6.5%
2018Q3	89.3%	5.0%	5.7%
2018Q4	87.5%	5.8%	6.7%

Figure 5 shows the total subjected dollar volume of measure of size for all industries from 2017Q1 to 2018Q4, and Figure 6 shows the weight-adjusted measure of size for the births added to the surveys over that same period. The simulation totals (red) are, for the most part, not statistically significantly different from the production totals (blue) except in 2017Q1-2017Q3. The differences for the first three quarters is explained by the retail oversampling issue. A couple of outlier EINs in specific industries caused the disparity in 2017Q4. For one outlier EIN, the originally captured NAICS code wasn't detailed enough to allow for subjection in second phase. As a result, its measure of size was not tabbed in the simulation. For the second outlier EIN, the loss of response data meant that it was subjected to sampling in a different industry with a much smaller payroll-based measure of size. Thus, aside from the quarters affected by the retail over sampling issue, the measure of size assignment appears to be comparable despite the loss of response data.
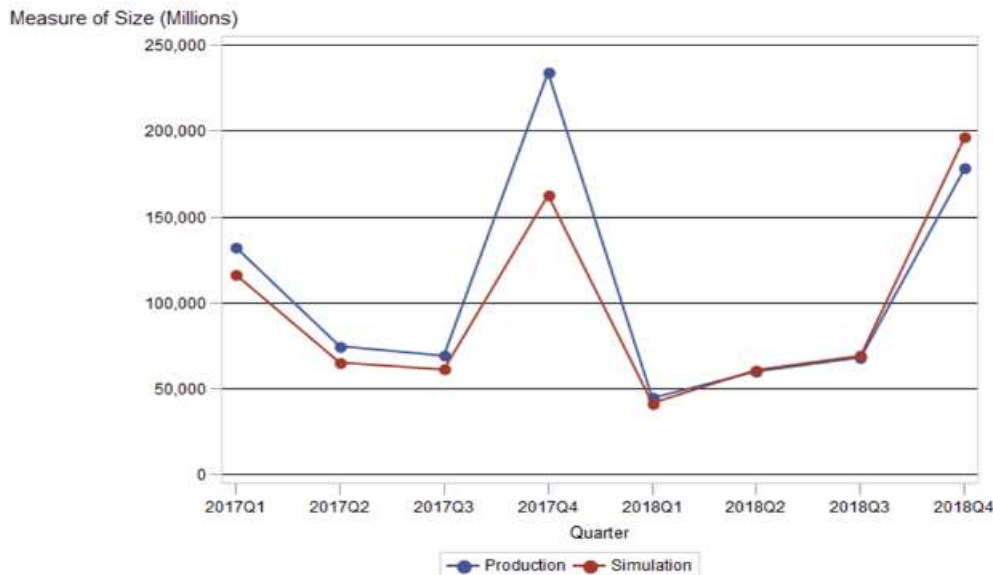
Figure 5: Comparison of subjected Dollar Volume (Measure of Size) from 2017Q1 to 2018Q4 between production and simulation Second Phase runs
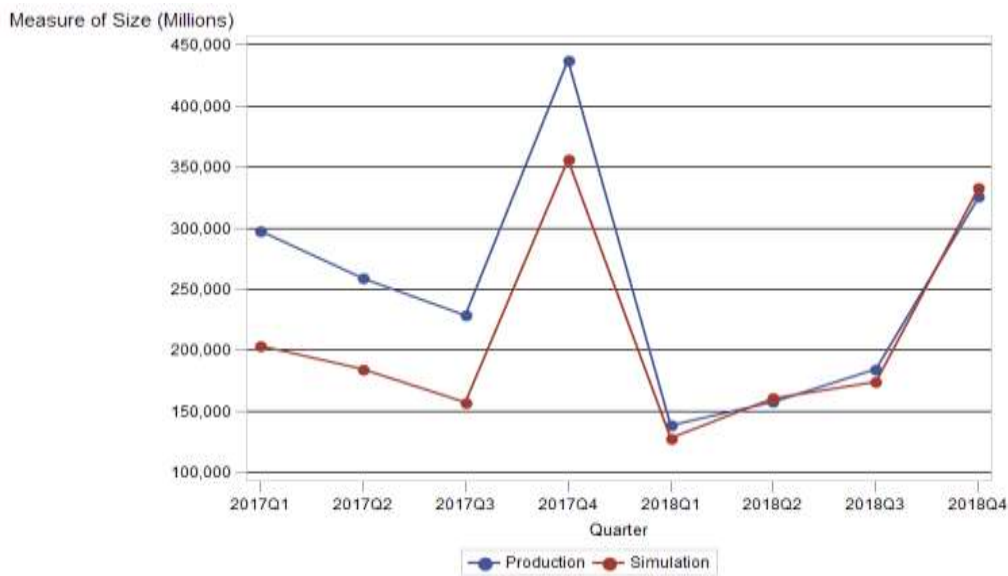

Figure 6: Total Weighted Measure of Size for all Second Phase Selects, by Quarter

7. Impact on Company Affiliation

Another insight was a decrease in the number of EIN 'Drops'. An EIN Drop is an EIN that is being dropped from further processing due to an identified association with an already-sampled company. This affiliation information is generally captured as part of an EIN's response to the Business and Professional Classification Report. Summary counts of these "drops", provided in Table 5 below, indicate an increased potential for subjecting EINs to sampling that are already associated with sampled companies, resulting in an increase in analyst-submitted requests for affiliation changes throughout the life of the sample. The simulation shows that for each quarter, anywhere from 50 to 200 additional EINs could have been subjected to sampling when they should have been drops. This is, in large part,

due to the loss of the company affiliation information provided by NPC analysts as well as the loss of data fields that were set during the TFU window. While the assumption of the loss of this data, in its entirety, may not be completely valid, it was handled this way in order to present a worst-case scenario.

Table 5: The difference in the number of EIN "drops" between production and simulation.			
Second Phase Quarter	*Production Drops*	*Simulation Drops*	*Decrease in Number of Drops*
2017Q1	300	250	50
2017Q2	350	200	150
2017Q3	350	250	100
2017Q4	700	500	200
2018Q1	350	250	100
2018Q2	450	350	100
2018Q3	400	300	100
2018Q4	700	550	150

6. Impact on Company Affiliation

We concluded our research by assessing the impact on the NAICS assignment. We have already seen some of this impact indirectly through the subjected/selected counts in Figures 1 and 2. In the simulation, fewer EINs were subjected to sampling because their NAICS detail from administrative data alone was not detailed enough to be assigned a sampling recode in second phase. However, the other important component to this analysis are the EINs that are subjected/selected. In particular, for EINs whose response data were "blanked" in the simulation, did the Business Register provide a sufficiently close match to the response NAICS we used in production?

Each quarter, as part of the monitoring and analysis of the SMASM process, a table summarizing the level of NAICS agreement between the response NAICS from SQ-Class and the corresponding first phase NAICS with which sampling was originally conducted on the respondents, is produced. We regenerated these tables using the simulation data to give us an idea of the impact on NAICS assignment. A comparison of simulation to production can be found in the graphs in Figure 7. This figure shows that there will naturally be a drop in the volume of NAICS matches, particularly at the 5- and 6-digit level, while the distribution is similar for both scenarios.
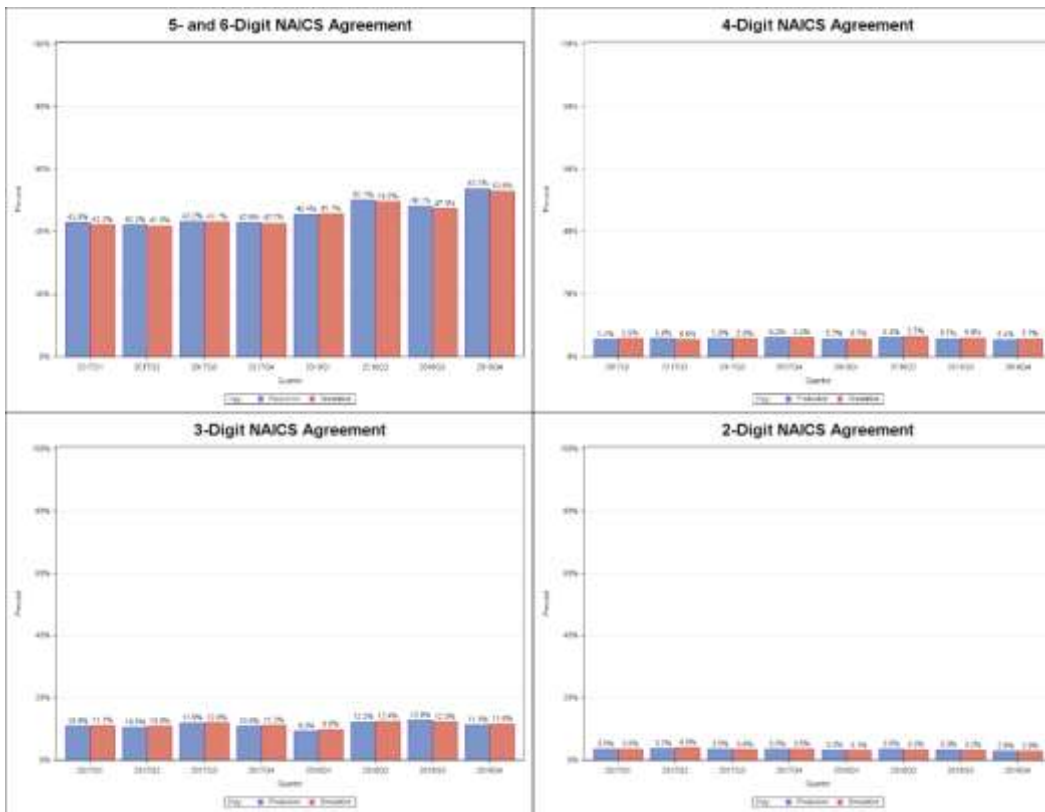
Figure 7: Comparison of First Phase NAICS to Response NAICS, Production vs Simulation

For each simulated quarter, we compared the production NAICS to the simulated NAICS for the EINs with check-in dates after the start of TFU, i.e., the EINs with "blanked" response data. For the simulation NAICS, both response data and NPC-provided NAICS were excluded from consideration for assignment. Only first phase NAICS or the latest NAICS on the Business Register (BR) at the time of second phase sampling could be assigned in the simulations.

Table 6: NAICS agreement between Production and Simulation, "blanked " EINs only					
Second Phase Quarter	*Number of EINs with "blanked" response data[3]*	*Percent 6-digit Match*	*Percent assigned to same sector, but different 6-digit*	*Percent assigned to different sector*	*Percent ineligible*
2017Q1	2700	47.4%	14.5%	14.5%	23.6%
2017Q2	2600	51.9%	16.7%	16.7%	14.8%
2017Q3	3100	47.6%	15.9%	17.5%	19.0%
2017Q4	4200	47.1%	15.3%	16.5%	21.2%
2018Q1	2800	47.3%	10.9%	14.5%	27.3%
2018Q2	3000	47.5%	16.9%	15.3%	20.3%
2018Q3	3600	49.4%	16.4%	16.4%	17.8%
2018Q4	3300	52.4%	13.8%	13.8%	20.0%

---

[3] Number of EINs with "blanked" response data, i.e. with check-in dates after the start of TFU.

For the EINs with "blanked" response data, Table 6 shows the percentage of NAICS agreement between simulation NAICS and production NAICS assigned in second phase. We ended up assigning the same 6-digit NAICS during second phase on average 47 to 52 percent of the time. Conversely, for approximately 15 to 25 percent of the blanked EINs, the BR or first phase NAICS isn't defined enough to assign a NAICS or primary sampling stratum/recode, in which case the EIN would remain ineligible and recanvassed in a subsequent quarter. The remaining percentage of EINs are assigned a different NAICS code than what was assigned during production, and are generally evenly split between being assigned a NAICS within (or outside of) the same sector as the production-assigned NAICS code.

Table 7: Count of EINs selected into a different recode in the simulation than they were subjected with in production			
Second Phase Quarter	*Simulation second phase selects (all trades)*	*Different Recode*	*Percent Different Recode*
2017Q1	1400	100	7.1%
2017Q2	1600	100	6.3%
2017Q3	1700	100	5.9%
2017Q4	2100	200	9.5%
2018Q1	1300	100	7.7%
2018Q2	1400	100	7.1%
2018Q3	1600	150	9.4%
2018Q4	2000	200	10.0%

When looking at EINs selected in second phase for the simulation, 10% or less are being sampled in a different recode from the one they were subjected with in production. See Table 7 for details. Although this number might seem alarming, 23 – 39 percent of second phase selects under our current structure enter second phase sampling operations without a response NAICS. Assuming the same 6-digit NAICS match rates in Table 6, we would expect around half of those 23 – 39 percent of second phase selects without a response NAICS to actually have had a different NAICS, if they had responded. This is not too far from the 10 percent mentioned above. With the loss of the additional response NAICS gained during TFU, from 23 – 39 percent in production to 45 – 59 percent in the simulation (see Table 8), we could expect the proportion of second phase selects without a response NAICS to actually have a different NAICS than what is assigned to them to increase to approximately 25 percent of second phase selects. These estimates are included in Table 8.

Table 8: Comparison of potential misclassified selects between Production and Simulation, 2017Q1-2018Q4

Second Phase Quarter	Production			Simulation		
	Second phase selects (all trades)	*Percent Without Response NAICS*	*Percent of misclassified selects[4]*	*Second phase selects (all trades)*	*Percent Without Response NAICS*	*Percent increase in misclassified selects[5]*
2017Q1	1600	28.1%	15.6%	1400	48.1%	29.6%
2017Q2	1800	38.9%	22.2%	1600	56.3%	31.3%
2017Q3	1900	31.6%	18.4%	1700	48.5%	27.3%
2017Q4	2200	22.7%	13.6%	2100	45.2%	26.2%
2018Q1	1400	28.6%	14.3%	1300	52.0%	28.0%
2018Q2	1400	35.7%	17.9%	1400	59.3%	29.6%
2018Q3	1700	32.4%	17.6%	1600	53.1%	28.1%
2018Q4	2100	28.6%	14.3%	2000	50.0%	22.5%

8. Conclusion

The loss of TFU has the potential to affect the current surveys in a number of different ways:

- A decrease in the number of births added to current surveys
- Incorrect sampling weight assignment due to inaccurate measure of size
- Increase in NAICS misclassifications
- Increase in missing company affiliation

From the simulated second phase runs, we saw differences in counts of the births being added to the current surveys. The decrease in the overall subjected/selected counts are largely attributable to the lack of a 5- or 6-digit NAICS for EINs who would have responded during TFU.

For the EINs who were subjected to second phase in the simulation, we saw in Table 3 and Table 4 that the assigned measure of size would typically keep an EIN within one stratum assignment of it's corresponding production stratum between 89 and 94 percent of the time. However, because weights can be vastly different from stratum to stratum, and from recode to recode, the overall impact is more difficult to measure. For this reason, the dollar volume of subjected EINs and the weighted dollar volume of selected EINs were also examined and compared to their production equivalents. Production and simulation numbers were found to be comparable, with exceptions only for misclassified NAICS.

In the NAICS evaluation, we found that the most recent available NAICS at the time of second phase sampling frame creation only matched the response NAICS on average 47 to 52 percent of the time (see Table 6), which would result in around 10 percent of births being sampled in a different industry than if we had the additional response data (see Table 7).

---

[4] Using 5- and 6-digit NAICS match percentages from Figure 6, applied to *Production selects without Response NAICS*.

[5] Using 5- and 6-digit NAICS match percentages from Figure 6, applied to *Simulation selects without Response NAICS*.

Finally, we saw that the loss of the additional company affiliation data acquired during TFU could result in subjecting to sampling an additional 50 to 200 EINs per quarter that are affiliated with an already-subjected company.

Therefore, we conclude from our research that the impact of the loss of TFU data is on the NAICS and company affiliation (see Table 6 and Table 7). The reduced count of subjected/selected EINs in second phase is largely dependent on the lack of a 5- or 6-digit NAICS for assigning a sampling recode. Because EINs without a NAICS match are recanvassed every other quarter, an increase in the number of recanvassed EINs can be expected. Likewise, an increase in the potential for selecting misclassified births into the current surveys is to be expected when dealing with the loss of response data for over 2,600 EINs per quarter. This could result in inaccurate industry estimates and in an increase in analyst-submitted NAICS change requests, shifting the burden from classification to survey analysts and math stats.

Likewise, the reduced number of drops each quarter increases the likelihood of subjecting and selecting EINs into our current samples when they are actually out-of-scope or already associated with a sampled company. This, in turn, creates more work for the survey analysts in submitting requests to modify the affiliation and representation of a sampled EIN and the math stats who have to review and evaluate these requests.

The effects on the measure of size and stratum assignment appear to be reasonable as payroll and employment-based measures of size achieve the same stratum assignment over 90 percent of the time. Furthermore, reported revenue from the Business and Professional Classification Report only serves as the measure of size 60% of the time when available. The potential still exists for outliers to occur, as the example from Figure 3 revealed; however, based on our research, we can't definitively say that the loss of sales data from the additional responses received during TFU will result in significantly more requests to reassign the sample weight of an EIN due to an incorrect measure of size assignment at the time of sampling.

If the decision is made to discontinue TFU, earlier analysis showed that an increase in the size of the SQ-Class sample would be needed to achieve the same response volume. However, doing this does not guarantee that second phase sampling will include more accurate information. Additionally, modifying the sampling rates for first phase would necessitate a modification to second phase sampling rates as well. First phase is designed to select around 10 percent of the births presented during extraction. Second phase is designed to then select 20 percent of the first phase selects to achieve the approximate 2 percent selection rate of the universe, thereby, matching the overall initial BSR sample selection rate. If the first phase sample were to increase by 50-55 percent in accordance with the numbers in Table 1, this would reduce the selection rate in second phase from 20 percent down to about 13 percent.

9. Future Work

While this research stops short of making recommendations for or against retaining TFU, consideration should be given to recommendations with regards to the Business and Professional Classification Report's relationship to a more robust annual business survey system and the Business Register. The impact of the loss of additional response data provided by TFU was only evaluated for the current BSR surveys. However, NAICS feedback from SQ-Class to the Business Register takes place even for out-of-scope EINs.

This impact, which could not be evaluated, would become more amplified under a survey system that includes even more industries, in which SQ-Class would have a more prominent role in both obtaining and maintaining classification in concert with the Business Register.

It would be worthwhile to look at the impact of losing TFU data would have on the estimates of the surveys receiving these births. If cases have been subjected/selected in the sample that are misclassified. These cases would have been given a wrong NAICS and/or measure of size. These cases will need weight changes or if they are in the wrong NAICS, they are receiving weights and measures of size based on the NAICS they presented as at the time of birth sampling. We could take the simulated and production results to see which cases are misclassified. Then, see how those cases contribute to the survey they were added to.

In lieu of TFU, another proposed idea is to incorporate an additional due date reminder mailing to take the place of TFU. Figure 8 shows the level of responses by day for the 2018Q3 first phase mail (plots for all quarters are available in the Appendix). Note, actual response counts have been removed from the graphs to prevent any risk of disclosure. The responses between the TFU Start (solid red) and Closeout (solid blue) lines are the data that we are at risk of losing without TFU. There are two large peaks on this graph. The left peak corresponds to the response after the initial mailing while the right peak, just prior to the TFU Start date, corresponds to the response received after the due date reminder (DDR) mailing. Assuming that any due date reminder mailing will result in a spike in responses, an additional mailing sometime during the TFU window could help account for some of the data lost by cutting TFU operations.
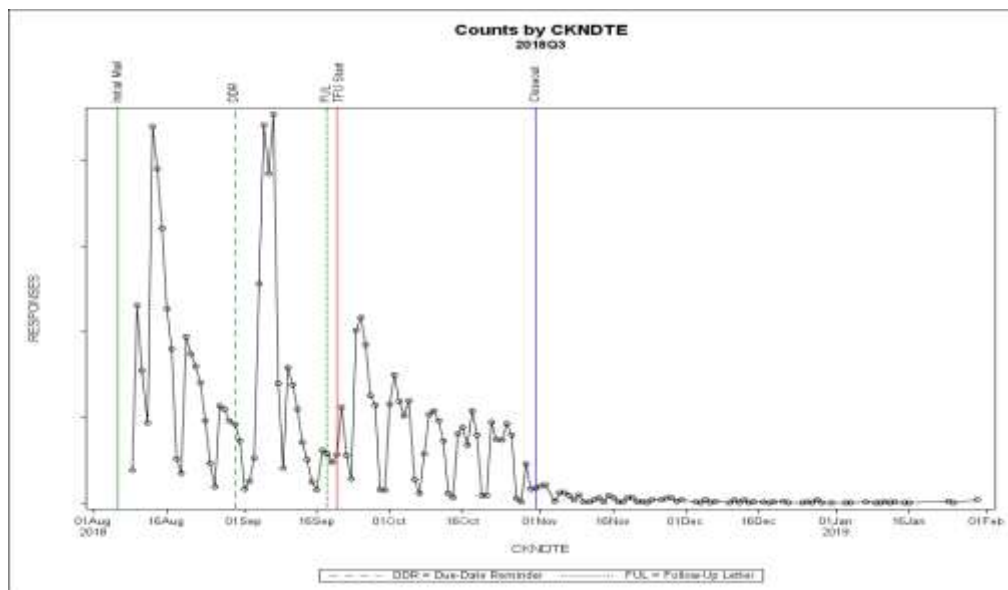

Figure 8: Responses by date, 2018Q3 mailing

Acknowledgements

References

U.S. Census Bureau. (2020). North American Industry Classification System. Retrieved July 21, 2020 from https://www.census.gov/eos/www/naics/.

U.S. Census Bureau. (2020). National Processing Center. Retrieved June 16, 2020 from https://www.census.gov/about/npc.html.

U.S. Census Bureau. (2020). Business and Professional Classification Survey: About the Survey. Retrieved June 16, 2020 from https://www.census.gov/programs-surveys/sq-class/about.html.

Diamond, L. K. (2019). Birth Extraction Specification. Internal U.S. Census Bureau Memorandum. June 26, 2019.

Diamond, L. K. and Owen, M. (2017). First Phase Processing Specification. Internal U.S. Census Bureau Memorandum. December 14, 2017.

Sell, F. (2018). SMASM Overview for BSR-17. Internal U.S. Census Bureau Memorandum. September 20, 2018.

Sell, F. (2014). Business Sample Revisions 2017: Second Phase Processing Specification. Internal U.S. Census Bureau Memorandum. June 25, 2020.

Sell, F. (2019). Second Phase Processing Specification. Internal U.S. Census Bureau Memorandum. June 26, 2019.

Sell, F. (2018). BSR-17 Second Phase Sample Register. Internal U.S. Census Bureau Dataset. February 19, 2019.

Appendix

The following graphs illustrate the pattern of responses to the Business and Professional Classification Report by day. These plots, by quarter, help us to understand the volume of response data received throughout the quarter. Reference lines are added providing key dates in the SMASM processing schedule that coincide with increases (and decreases) in received response data. Counts were removed to comply with the Census Bureau Disclosure Avoidance guidelines.
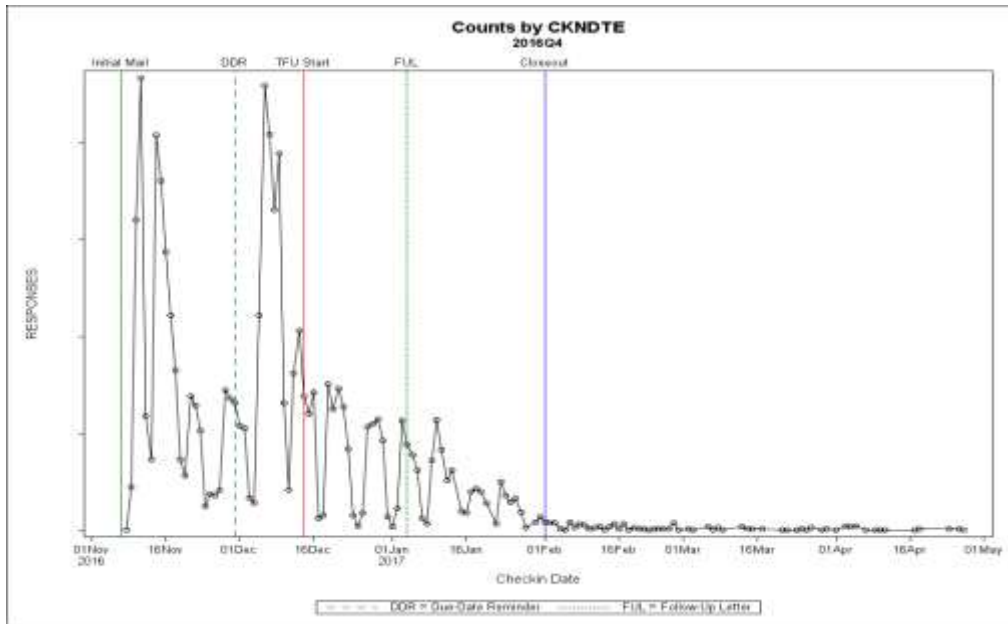

Figure 9: Responses by date, 2016Q4 mailing
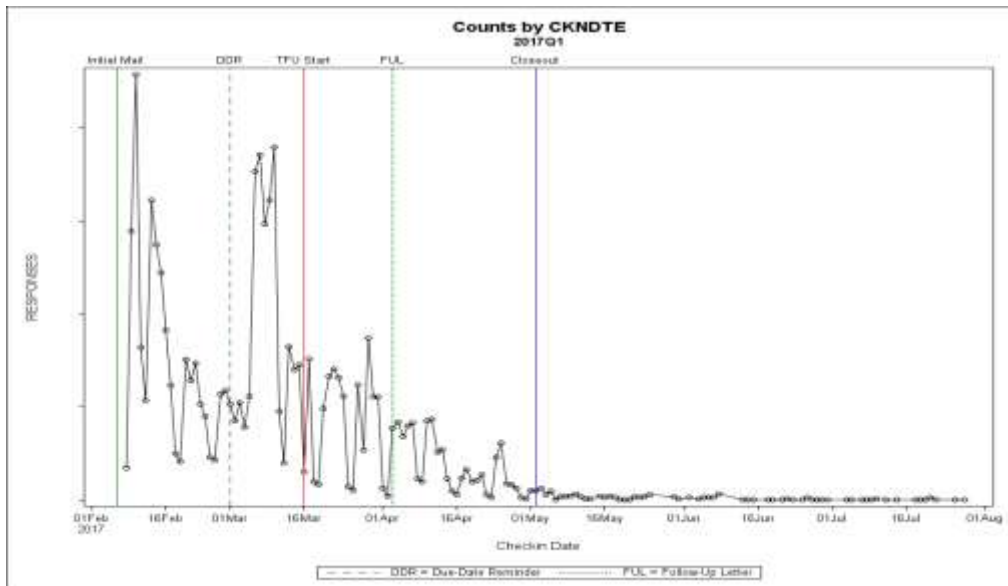
Figure 10: Responses by date, 2017Q1 mailing
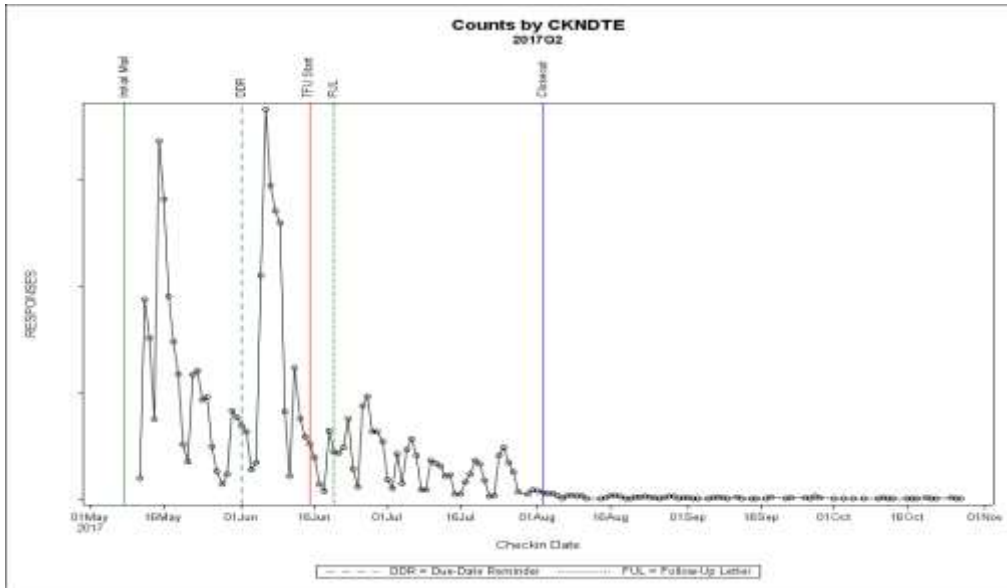
Figure 11: Responses by date, 2017Q2 mailing
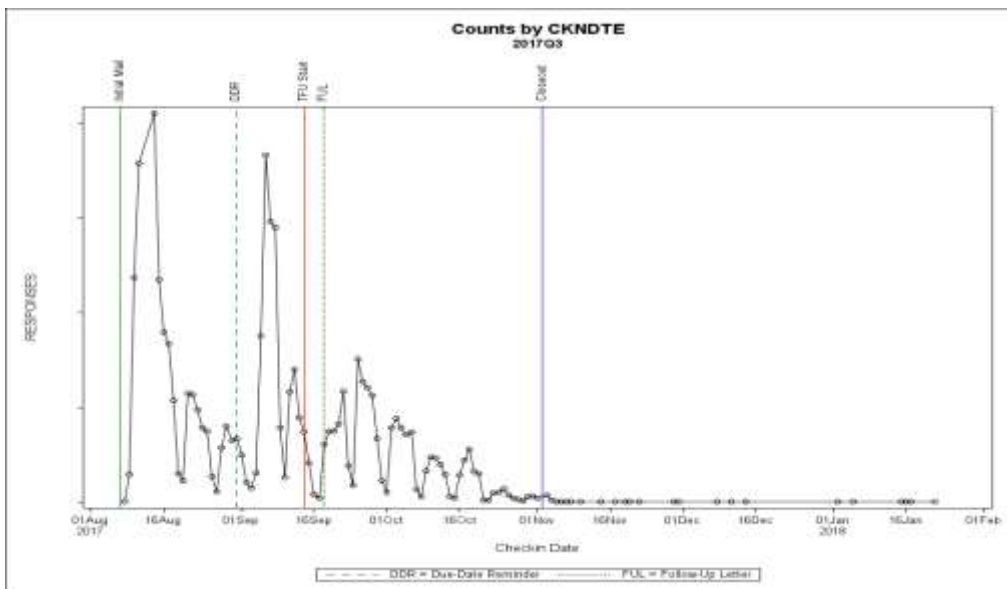
Figure 12: Responses by date, 2017Q3 mailing
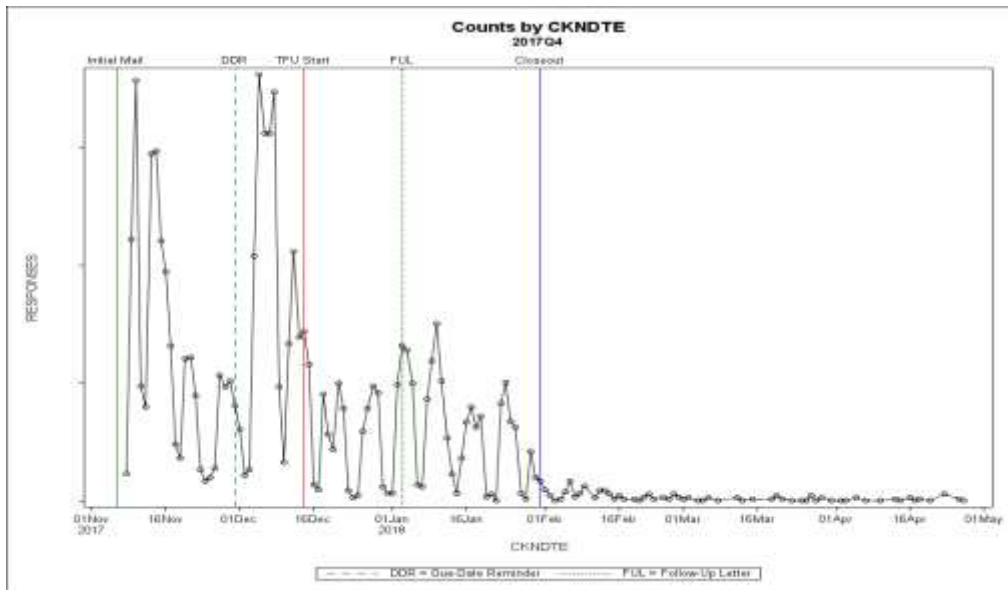
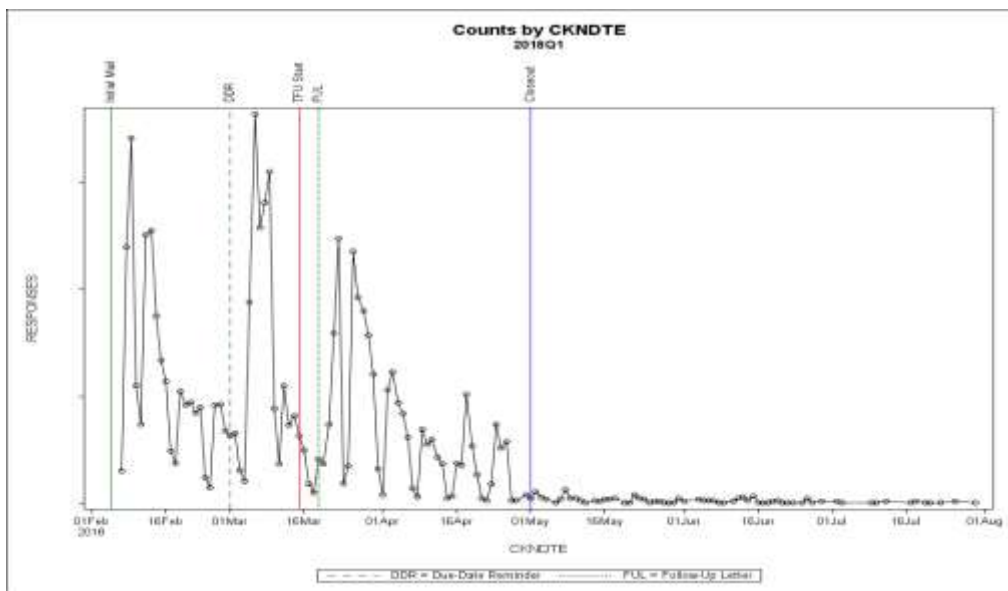**Figure 13:** Responses by date, 2017Q4 mailing

Figure 14: Responses by date, 2018Q1 mailing

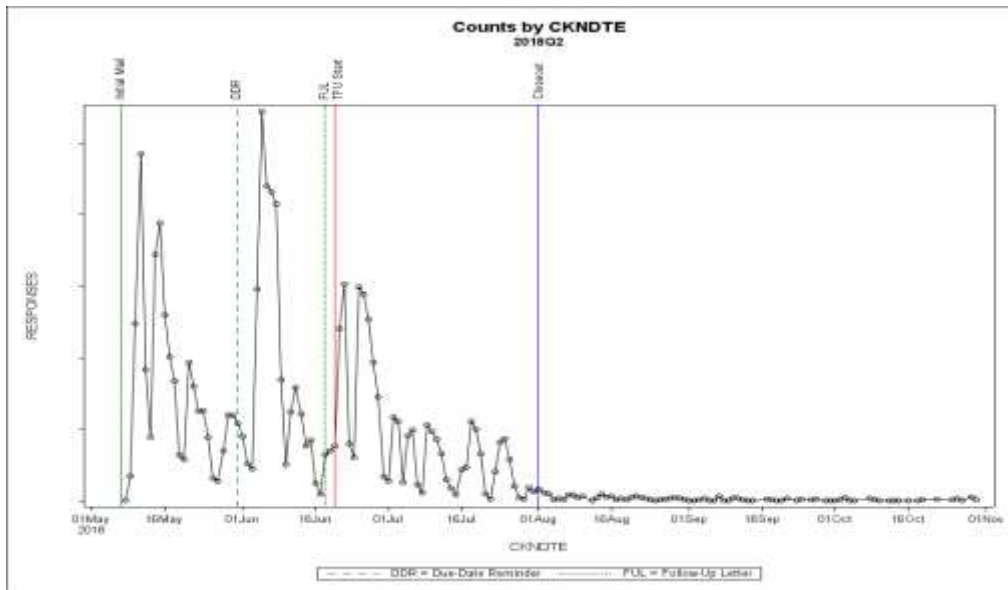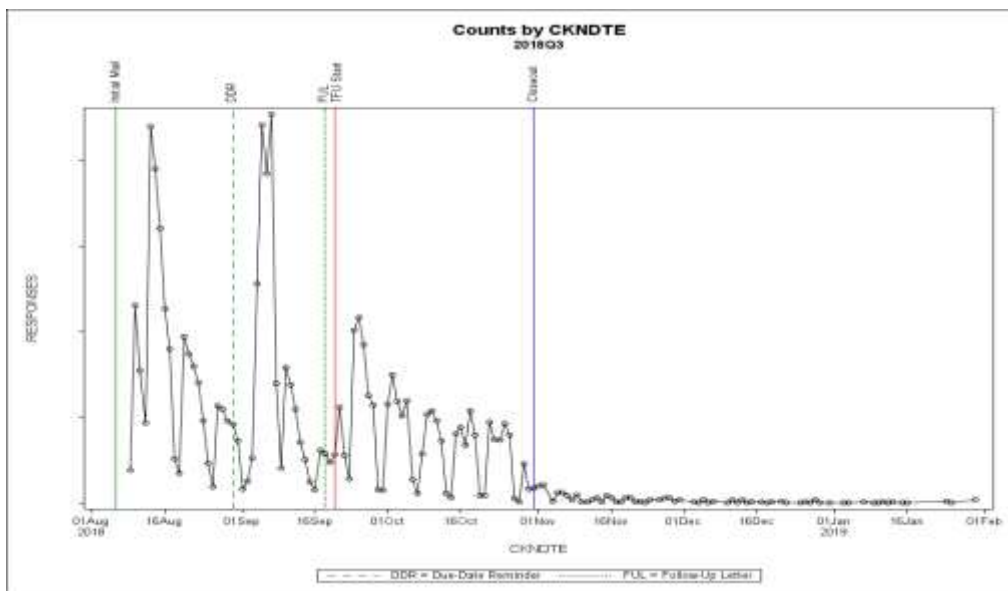**Figure 15:** Responses by date, 2018Q2 mailing

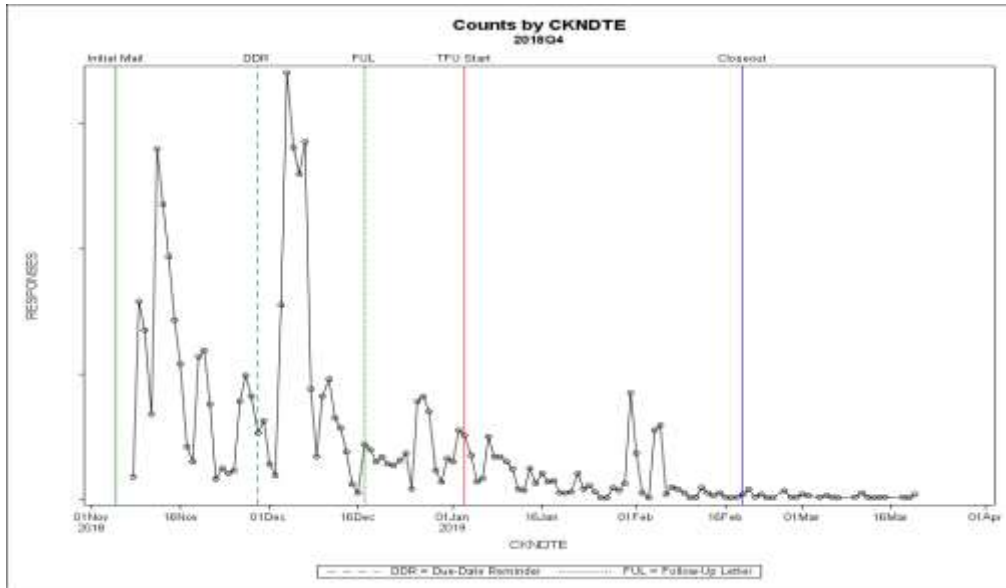**Figure 16:** Responses by date, 2018Q3 mailing

Figure 17: Responses by date, 2018Q4 mailing

The final graph (2018Q4), for comparison, is for the de facto No-TFU quarter. The scheduled TFU start date of January 3 was included as a reference line to indicate the window in which TFU would have taken place without the government shutdown. It should also be noted that the Closeout line (blue) was extended by about 2 weeks as a result of the government shutdown and would have originally been January 31. The small peaks within the "TFU window" appear to correspond to an email reminder that was sent on January 29 after the government shutdown had ended and SMASM processing had resumed.