Statistical Journal of the IAOS 32 (2016) 167–176
DOI 10.3233/SJI-161007
IOS Press

Adaptive design research for the 2020 Census[1]

Scott Konicki[a,*] and Tamara Adams[b]
[a]*Decennial Statistical Studies Division, U.S. Census Bureau, USA*
[b]*Associate Directorate of Research and Methodology, U.S. Census Bureau, USA*

Abstract. The U.S. Census Bureau is researching and testing new methods to reduce the cost of the 2020 Census while maintaining data quality. One of the most costly components of the 2010 Census was Nonresponse Followup. In this operation, enumerators conducted in-person interviews at housing units that did not return a census questionnaire by mail. For the 2010 Census, enumerators were instructed to visit each of these units up to three times until the case was resolved. Additionally, enumerators were to make up to three contact attempts by telephone. In this paper, we present an overview of current research on determining the number of contact attempts that should be made to nonresponding units with an emphasis on cost containment and improved overall productivity. Rather than the fixed contact strategy employed in the 2010 Census, we consider adaptive approaches that maintain the quality of the data. We present initial results of possible approaches using data from the 2010 Census and discuss the implications of the methods. We also discuss modeling contact probabilities for each hour of the day to support the case management system.

Keywords: Census, adaptive design, Nonresponse

1. Introduction

One of the most costly operations of the 2010 Census was the Nonresponse Followup (NRFU) operation. This operation primarily involved census enumerators interviewing and verifying the status of housing units that did not respond to the 2010 Census by mail. The enumerators were instructed to visit each unit up to three times until the case was resolved and make up to three contact attempts by telephone. The interviews were completed on paper questionnaires. The 2010 NRFU workload included 47 million addresses and cost nearly $1.6 billion [1].

In an effort to reduce costs and improve efficiency while maintaining data quality, the Census Bureau is researching and testing new ways of conducting the NRFU operation. One area of research is adaptive design. In this paper, we discuss current research on adaptive design for two components of the NRFU operation. The first component is the maximum number of contact attempts to make to nonresponding units. Unlike the 2010 Census approach in which all units were treated with the same contact strategy, we investigate an approach that allows the maximum number of contact attempts to vary across areas. The goal of this approach is to contain costs while equalizing a measure of data quality across the areas. The second component deals with what time of day to make contact attempts to nonresponding units. Using paradata from the American Community Survey (ACS), we examine logistic regression models to predict the probability of contacting units at each hour of the day. These predictions fed into a new case management system that was used in the 2015 Census Test as part of the Reorganize Census with Integrated Technology (ROCkIT) initiative.

For both of these components of NRFU, we will explain how we applied the methods during the 2015 Census Test in Maricopa County, Arizona. As will be seen, the application in the 2015 Census Test was done

[1]This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

*Corresponding author: Scott Konicki, Decennial Statistical Studies Division, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, USA. Tel.: +1 301 763 4292; E-mail: Scott.M.Konicki@census.gov.

in a static manner, rather than adaptive. That is, decisions were made before NRFU fieldwork began, and the contact strategies did not change based on results observed during the operation. Because many new procedures were being implemented in this test for the first time, we needed to balance the numerous objectives of the test. In the years ahead, we will continue to investigate enhancements to these approaches and ways to make dynamic decisions by monitoring the NRFU operation.

2. Adaptive Design in the 2013 and 2014 Census Tests

Adaptive, or responsive, survey designs leverage paradata to analyze cost-quality tradeoffs in real time and make changes to the data collection protocol [2]. The use of computer-assisted data collection instruments allows for this real-time decision making. For the 2010 Census, self-response forms were mailback questionnaires, and the NRFU operation also used a paper-and-pencil form. This lack of computerization for collecting the census data prevented the use adaptive designs. Following the 2010 Census, an internet self-response option and computerized NRFU instrument began to be implemented in various tests leading up to the 2020 Census. These new data collection methods allowed for adaptive survey procedures to be investigated.

The 2013 Census Test included two adaptive design panels to examine whether data could be collected more efficiently by using administrative records to reduce NRFU fieldwork and an adaptive approach to case management. In the first adaptive design panel, administrative records were used to identify vacant units and enumerate occupied units before fieldwork. The remaining cases were then sent to the field and allowed three personal visits before seeking a proxy. In the second adaptive design panel, administrative records were not used prior to fieldwork. Cases with administrative records available were allowed one personal visit before seeking a proxy, while cases without administrative records were allowed three personal visits. In both adaptive design panels, a daily propensity model produced probabilities of contacting units in the field. Each day, the seven cases per interviewer with the highest predicted probability were assigned as high priority cases. Enumerators were trained to attempt all seven of the high priority cases each day they worked. For more information on the 2013 Census Test, see [3].

The 2014 Census Test included an adaptive design panel that, among other features, assigned a maximum number of contact attempts allowed for each housing unit. Administrative records were used to remove unoccupied units from the workload prior to fieldwork. Cases for which we thought we could reasonably enumerate the unit using administrative records were allowed one personal visit attempt without a chance for a proxy interview. For the remaining cases without administrative records, half of the units were allowed one contact attempt if the unit was in an area with higher response rates in the 2010 Census. The other half in areas with lower response rates in 2010 were allowed three contact attempts. Instead of using a propensity model to assign high priority cases, a set of seven priority cases per enumerator were assigned each day using business rules. The business rules attempted to assign priority cases that were geographically close and had fewer contact attempts already made.

Moving onwards from the 2014 Census Test, the Census Bureau began work on the ROCkIT initiative. The ROCkIT initiative has the goal of developing a new concept of operations for the decennial census that increases 2020 NRFU productivity by a minimum of 20% [4] by decreasing miles traveled per case and total hours spent, among other things. This initiative was implemented in both of the adaptive design panels for the 2015 Census Test. The ROCkIT initiative includes major changes for the field operation such as an enhanced operational control system and an updated field structure with new staff roles and staffing ratios. Some of the components of the ROCkIT initiative replace adaptive design concepts used in previous census tests. For example, instead of assigning priority cases to enumerators each day, attempts are assigned to enumerators based on enumerator availability, workload location, and best contact times. The enumerators receive an optimal routing of attempts to minimize travel. Unlike previous censuses and tests, enumerators will not have ownership of a set of cases that they alone will work until completion.

3. Determining contact attempts for Nonresponse Followup

3.1. Motivation

In the 2014 Census Test adaptive design panel, the decision on the maximum number of contact attempts allowed was subjective. The set of block groups were split in half using the 2010 Census self-response rates. This approach sought to allow more contact attempts to units in areas that would have lower self-response rates in the 2014 Census Test. This is a desirable approach because areas with a higher rate of NRFU units would be given more attempts to resolve the cases. However, the decision on how to split the block groups did not explicitly take into account any measure of data quality.

Initial research for the 2015 Census Test was focused on real-time stopping rules for NRFU. These rules were based on using the observed information at any point in time to impute for the fieldwork cases that were still unresolved. Approaches such as those of [5,6] were considered. Given the numerous new methods being tested in the 2015 Census Test, it was decided to not use stopping rules for NRFU. Instead, we decided to build off of the 2014 Census Test contact strategy and research an approach to this problem that would allow us to assign the maximum number of contact attempts across a continuum rather than the discrete groups of one and three visits. Furthermore, we wanted to use an objective measure of data quality to guide the assignment of contact attempts.

For our current research, the data quality measure we have been working with is the proxy rate. A proxy is someone other than a household member who can help to complete the NRFU interview for a housing unit, such as a neighbor or apartment manager. While proxy respondents can often provide enough information to close a case, the obtained data is often less complete than the data collected from household respondents. For example, the 2010 Census Coverage Measurement results show that 23.1% of census person records enumerated by proxy respondents needed to have all of their demographic characteristics imputed, as compared to only 1.6% of person records when the respondent was a household member [7].

Clearly, proxy respondents provide less complete data than household members. Often, the proxy may know the household size but not the detailed demographic information. Therefore, it would be beneficial to limit the use of proxy respondents as much as possible. The NRFU contact strategy has been for enumerators to seek out a proxy respondent only after the final contact attempt to the unit. Given this contact strategy, the number of attempts allowed directly influences the proxy rates. When more attempts are allowed, there are more opportunities to contact a household member instead of relying on a proxy respondent.

We define the proxy rate as the number of units completed by a proxy respondent out of the total number of units in the area (including self-responding units). Given budget and time constraints, we cannot conduct enough visits to avoid using proxy respondents altogether. In this paper, we assume that we can allocate an average of a three-visit maximum per NRFU case. One approach would be to allocate visits in a way that would minimize the overall proxy rate. However, this approach could result in disparate proxy rates across smaller areas because it only focuses on the total proxy rate rather than the distribution across areas. Instead of minimizing the overall proxy rate, we seek to allocate NRFU visits to minimize the variance of proxy rates across block groups. All other things equal, such an allocation would give more visits to areas with lower self-response rates. The following section describes our methodology. We use Arizona as an example and discuss the implementation for the 2015 Census Test.

3.2. Methodology

We use the 2010 Census results to define the self-response and NRFU universes in Arizona. The Census Bureau has also been researching and testing the use of administrative records to identify vacant units and enumerate occupied units in place of NRFU fieldwork. We incorporate this feature using a rule-based approach to identify units to remove from the NRFU workload via administrative records. In short, vacant units are identified using undeliverable-as-addressed (UAA) information collected by United States Postal Service in 2010. Occupied units are identified using IRS 1040 tax returns received by the end of April 2010 (see the IRS/UAA approach in [8]).

For ease of presentation, we present the methodology for the hybrid administrative records removal approach. In this approach, units determined to be vacant via administrative records are removed from the NRFU workload before fieldwork begins. Administrative records occupied units are allowed one contact attempt to try to resolve the case with a household respondent; if the contact attempt fails, the unit is enumerated using the administrative records. This is one approach the Census Bureau tested in the 2015 Census Test. The other approach, called the full removal approach, removed both vacant and occupied units from the workload prior to the start of fieldwork. In the following discussion, we briefly note the areas in which the methodology would differ under the full removal approach to administrative records modeling.

In each block group, we assume the following probabilities for completing NRFU cases:

- Probability of completing a case with a household member on each attempt: 0.4;
- Conditional probability of completing a case with a proxy respondent given that all attempts to complete the case with a household member have failed: 0.9.

These are simplifying assumptions for our initial research that were derived from 2010 NRFU contact history data. Future extensions of this research can examine ways in which these probabilities vary across block groups.

With these probabilities, we can calculate the projected proxy rate for each block group using the 2010 Census data. Because our concern with proxy respondents is the incompleteness of the data they often provide, we also take into account the quality of the administrative records used to enumerate occupied units. For the purpose of this research, we consider a unit to have incomplete administrative records if none of the persons in the unit have race or Hispanic origin data available from previous census responses. These units will still be removed from the NRFU workload, but will also count towards the proxy rate. The proxy rate for a block group is given by:

$$\textit{Proxy Rate} = \frac{(1 - P_{hh})^{\textit{NumVisits}} \times P_p \times \textit{Remaining Fieldwork}}{\textit{Total Units}} + \frac{(1 - P_{hh}) \times \textit{Incomplete Adrec}}{\textit{Total Units}}$$

- *NumVisits* is the number of visits allocated to the block group. This input will be determined by the optimization.
- *RemainingFieldwork* is the number of NRFU cases that are not removed from the workload via administrative records. These cases need to be visited by an enumerator.
- *IncompleteAdrec* is the number of units identified as occupied by administrative records, but were incomplete as per the definition above.
- P_{hh} and P_p are the assumed probabilities of completing the case with a household member and proxy respondent, respectively. We assume probabilities of 0.4 and 0.9, respectively. We assume independence of the responses.

Under our assumptions, we would expect 40% of the administrative records occupied cases to be completed with a household member. Therefore, only 60% of the incomplete administrative records cases would contribute to the proxy rate. For the full removal approach, all of the incomplete administrative records occupied cases contribute to the proxy rate because these cases are not attempted in the field.

One factor that our projected proxy rate does not account for is vacant units. By definition, a vacant unit does not have any household members and thus can only be completed by a proxy respondent. For our current research, these are not the types of proxy responses we are concerned about. In a way, our projected proxy rate represents the extreme case in which all of the NRFU fieldwork units are occupied units and subject to potential enumeration by a proxy respondent. Future research will consider how best to incorporate estimates of the number of vacant units in the NRFU fieldwork to remove these units from our projected proxy rate.

Our goal is to assign a maximum number of visits to each block group that will minimize the variance of the above proxy rate across the block groups. The constraint for this optimization is the total number of visits we can assign. For this research, and implementation in the 2015 Census Test, we used an average of three visits per NRFU fieldwork case. Table 1 shows the counts of the units in the 2010 NRFU universe for the state of Arizona. There are 862,214 + 173,942 = 1,036,156 cases in the NRFU workload because the administrative records occupied cases also receive one initial visit. This gives 3 × 1,036,156 = 3,108,468 visits to allocate under our constraint of three visits per workload case. This results in an average of 3.40 visits per remaining fieldwork case because units with occupied administrative records receive only one visit by design. Note that about 15% of the administrative records occupied cases are considered incomplete.

To conduct the optimization, we used the OPTMODEL procedure in SAS [9]. In addition to the total visits constraint, we also included the constraint that each block group must receive at least one visit. The optimal solution is calculated subject to the constraint that the maximum number of visits assigned meets the target average and does not involve the actual number of visits that will be attempted. In a multiple visit block group, some cases will be completed on the first visit, leaving the additional assigned visits unused. Such cases will still be counted in the total number of assigned visits against the constraint.

3.3. Results

Table 2 shows the results of the optimal allocation of NRFU visits for the 4,072 block groups in Arizona that

Table 1
Arizona 2010 NRFU universe by administrative records status

NRFU cases	1,201,229
Administrative Records Vacant	165,073
Administrative Records Occupied	173,942
Remaining Fieldwork (no Administrative Records)	862,214

Table 2
Optimal allocation of NRFU visits for Arizona block groups

Maximum visits	Count	Percent
1	377	9.3
2	1,225	30.1
3	1,320	32.4
4	978	24.0
5	168	4.1
6	4	0.1

were included in the 2010 NRFU operation. The optimization procedure returns a non-integer solution, so we used standard rounding rules to assign the number of visits. As desired, this approach assigns a number of visits along a continuum, rather than the two groups of 1 and 3 visits that were used in the 2014 Census Test. Most block groups are allocated 2 to 4 visits. Some areas with high 2010 self response rates and administrative records use are allocated only one visit. Only a few extreme block groups are allocated 5 or more visits.

3.4. Implementation in the 2015 Census Test

To determine the self-response and NRFU universes for the optimization, we use the 2010 Census data. To use this approach in practice, ideally we would use the self-response data from the test itself and run the optimization right before the NRFU operation begins conducting fieldwork. This would use the most up-to-date information on self-response and administrative records removal to inform the allocation, rather than using assumptions and projections based on 2010 Census data.

For the 2015 Census Test, there were additional constraints that made it necessary to make these decisions well in advance of the start of NRFU. The 2015 Census Test had many components, each with its own priorities and requirements. Field supervisors needed to know in advance how many enumerators to train and hire in each area. The new ROCkIT method of optimizing daily workloads to enumerators takes into account where enumerators live and the location of these fieldwork cases. Any significant changes to the initial allocation of NRFU visits would have created a mismatch between the locations of large workloads and the locations of the hired enumerators. The further the enumerators had to travel between home and their workloads, the less efficient the operation would have been.

Table 3
Initial allocation of NRFU visits for 2015 census test block groups

2015 test area	Maximum visits	Count	Percent (of area)
Central	1	0	0.0
	2	0	0.0
	3	45	54.2
	4	33	39.8
	5	5	6.0
Chandler	1	13	19.1
	2	40	58.8
	3	13	19.1
	4	2	2.9
	5	0	0.0
Mesa	1	9	12.2
	2	29	39.2
	3	22	29.7
	4	12	16.2
	5	2	2.7
Outer ring	1	0	0.0
	2	4	15.4
	3	6	23.1
	4	10	38.5
	5	6	23.1

One of the goals of the 2015 Census Test was to provide accurate cost estimates that can be used to plan for the 2020 Census. For this reason, we allocated contact attempts in the context of a broader census environment, rather than the environment specific to this test. We used the state-level results from Table 2 to observe the results for the 2015 Census Test areas rather than optimizing over only those block groups in the 2015 Census Test. Table 3 shows the distribution of visits across the four geographic areas in the 2015 Census Test. These areas were selected based on characteristics such as the 2010 Census self-response rate and concentration of minority populations. Note that Table 3 shows the results for all block groups in the test for illustration, but only a portion of these block groups were in the test panel that received the hybrid administrative records removal treatment.

Table 3 shows considerable variation in the distribution of visits between the areas. The Central area is a high Hispanic and high mobility area that had low self-response rates in the 2010 Census. These block groups were allocated more visits. On the other hand,

Table 4
Average number of visits from initial allocation for 2015 census test areas

Area	Average number of visits
Central	3.39
Chandler	1.91
Mesa	2.65
Outer Ring	3.88

the Chandler area had high self-response rates in 2010 and received fewer visits as a result.

For the 2015 Census Test, we used the results from Table 3 as an initial allocation to inform the field staff for hiring enumerators. To attempt a more adaptive procedure without compromising the other components of the test, we reallocated visits within each of the four geographic areas of the test site. The initial allocation determined an average number of assigned visits for each area. Although an average of 3 visits at the state level was used for the initial allocation, the average number of visits within a small geographic area may differ.

Table 4 shows the average number of visits from the initial allocation for the 4 areas of the test. To reallocate visits for the 2015 Census Test, we ran the optimization procedure within each area subject to the constraint of meeting the averages in the table. This procedure will (approximately) hold constant the total number of visits in each area as determined by the initial allocation. The benefit of this reallocation is that it can at least partially account for differences in the projected and observed self-response and administrative records removal rates. This was a shortcoming of the approach used in the 2014 Census Test. For that test, some of the block groups that were allocated only one NRFU visit based on having high self-response rates in 2010 actually had lower self-response rates in the 2014 test than block groups that were allocated three visits. Our reallocation approach attempted to partially correct for this situation by making changes within the geographic areas.

3.5. Future research and extensions

The initial research shown here offers improvements on the methods from previous census tests. There are many ways to improve upon this method as we move towards the 2020 Census. First, our initial research used simplifying assumptions about the probabilities of completing cases with household members and proxy respondents. These probabilities were assumed to be constant across all areas. We will examine ways of better estimating these probabilities and allowing them to vary across areas. For example, rural areas may have lower probabilities of completing cases with proxy respondents because there are fewer potential proxies in the immediate area. Second, we will examine other objective quality metrics than the proxy rate. One example is to account for the completeness of data provided by proxy respondents. Our current approach treats all proxy responses as equal when in fact some areas may have higher quality proxy respondents. For example, compare a stable, close-knit community as opposed to an area with high turnover. With the latter, the community members may not be able to provide much knowledge about their neighbors.

An important avenue of future research will be how to make this process more adaptive by determining the number of visits for units in real time. Our current approach fixes the maximum number of visits for each block group before NRFU begins. In a truly adaptive setting, paradata and results of the current operation would be used to make dynamic decisions as to whether to make additional contact attempts at a given unit. Included under this topic of research are stopping rules and dynamic resource allocations. Such methods were not possible in the 2010 Census because the NRFU operation was conducted using paper questionnaires. With the use of mobile devices that transmit response data and a wealth of paradata on a daily basis, there are now many possibilities for adaptive designs in the decennial census. Concepts such as the fraction of missing information [10] may be especially useful near the end of the data collection period. These metrics could identify specific areas where the limited remaining resources should be focused.

4. Contact time modeling

4.1. Motivation

As previously mentioned, the operational control system for the ROCkIT initiative assigns a set of cases for enumerators to attempt each day. Enumerators are given an ordered case route which minimizes travel. There are many inputs into this optimization, including enumerator availability, the location of cases to be worked, and the probability of successfully contacting cases and various hours of the day. In this section, we discuss the research for developing contact time predictions.

In the framework of Groves and Couper [11], the first step in the process of survey participation is contacting the sample household, which is a function of when the contact attempt is made and when the household members are at home. The importance of understanding at-home patterns of households has been acknowledged by the Census Bureau for decades. U.S. Census Bureau [12] presents hourly contact probabilities for various subgroups using interviewer data from the Current Population Survey. These results show distinct patterns by age of the contacted person and also by characteristics of the general area, such as the poverty rate. The authors note that the results of the study could be applied to improve interviewer planning and reduce costs of data collection. The goal of the present research was to provide the ROCkIT team with predicted probabilities of contacting units in the 2015 Census Test at each hour of the day between 10am and 9pm. These predictions would guide the workload assignments of the enumerators in order to increase the chance of successful contact and thus lower the cost of the fieldwork operation.

4.2. Methodology

We used contact history paradata from the ACS to develop logistic regression models for these predictions. The dependent variable for our analysis was derived from the outcome codes of cases in the ACS personal interview followup sample. We used ACS data from 2012 and 2013 and restricted our analysis to the months during which the decennial census NRFU is conducted (May, June, and July). Furthermore, we used only ACS cases that had a final outcome of occupied. Since interviews for vacant units could not be completed with a household respondent, we did not want these outcomes to influence the predictions.

Each record in our data represented a contact attempt from the ACS followup sample. Using the outcome codes of these contact attempts, we assigned each record as either a successful completion or a noncontact. For each contact attempt, we also had information on the day and time of the attempt. This allowed us to form predictions for different hours of the day. For this initial research, we combined Monday through Friday to predict a weekday probability for each hour, and likewise we combined Saturday and Sunday to predict a weekend probability for each hour.

While the dependent variable information came from the ACS paradata, we needed to use independent variables available for the model data and the units in the 2015 Census Test so that we could apply the predictions to the test units. We compiled multiple administrative records sources to create variables that describe the household composition. For example, we used variables indicating the age groups of the administrative records persons associated with the housing unit. The idea is that units with young children or senior citizens may have someone at home more often and be more likely to be contacted successfully during the day. On the other hand, units with only a single adult or working age adults may have lower probabilities of contact until the evening. In addition to age, we included variables indicating the race and Hispanic origin of the members of the household. We also included indicator variables for the administrative records sources from which the person records came.

This approach of using administrative records to describe the household composition is only useful for units that have administrative records persons associated with the unit. Even though our modeling data only used ACS occupied cases, many of these housing units did not match to any of the administrative records sources we used. For these cases, there was very little household-level information available for the unit. We addressed this issue by running separate models for units with administrative records and for units without. The latter included some household-level information, such as housing unit type and whether the unit was owner-occupied, renter-occupied, or vacant in the 2010 Census. The other variables used in these models were ACS 5-year estimates available at the block group level. The ACS estimates included the percent of people in the block group in various age and race and Hispanic origin categories. Additional area-level estimates for characteristics that are not available via administrative records, such as education and poverty, were included in both sets of models. Durrant et al. [13] show that area-level characteristics, such as unemployment rates and percentage of the population in age groups, are predictive of contact when household-level characteristics are not available.

To apply the predicted probabilities to the 2015 Census Test units, we created the same independent variables for each of the test units using the most recent vintage of the administrative records and ACS 5-year estimates. The result was a data set with each unit in the test and its predicted weekday and weekend contact probability for each hour of the day. These probabilities were passed to the ROCkIT team for use in the case assignment optimization.

4.3. Results

Initially, we approached the research by using the set of cases attempted in each hour to predict probabilities of completing the case at that hour. Figure 1 shows the distribution of predicted probabilities for the 2015 Census Test units from this approach for weekdays using box plots. There is very little difference in the distributions across the hours. Each hour has an average predicted probability near 0.4. The distributions for weekend hours showed similar results. The ROCkIT team raised concerns that these results were not very useful for the case assignment optimizer due to the lack of differentiation across the hours.

These initial results were unintuitive. We would expect to have greater chances of successfully contacting most households during the evening rather than the early morning. The issue arises because the ACS contact attempts were not randomly assigned to different hours of the day. ACS enumerators are instructed to visit units when they are most likely to obtain an interview. As a result, the vast majority of the ACS contact attempts in our data set were during late afternoon and evening hours. For example, of the contact attempts made on weekdays, only 5.6% were attempted before 11:00am while 12.7% were attempted between 6:00pm and 7:00pm. It is our assumption that enumerators are successful in using their knowledge of the local areas to decide when to make contact attempts at units. That is, enumerators have some idea as to which units are likely have someone at home during the day and use this information to guide their attempts. This would help explain the flat distribution of successes over the hours of the day despite the great differences in the distribution of attempts.

To account for the nonrandom nature of our data, we proceeded with a two-step process. First, we used our entire data set of contact attempts to predict whether each individual unit was attempted at each hour of the day. Call this P_h (attempt) where h denotes the hour. To do this, we ran a logistic regression model for each hour of the day, where the dependent variable noted whether the contact attempt was made during that hour, regardless of whether the attempt resulted in a completed interview. As with the models described previously, we ran separate models based on whether the unit had administrative records data available.

Second, we ran the logistic models to predict successful contacts as described previously. Because the ACS contact attempts are not randomly attempted, these models are really predicting the conditional probability of contacting a unit given that the case was attempted at the given hour of the day. Call this P_h (contact | attempt). Table 5 summarizes the dependent variables for the models. To obtain a final probability, we multiplied the probabilities from the first and second step:

$$P_h(contact) = P_h(attempt)P_h(contact|attempt)$$

The resulting probabilities were very small due to the generally low probabilities of attempting at each hour. To obtain useful results, we transformed the probabilities back to the scale of Fig. 1 using the mean and standard deviation of those original results. Figure 2 shows the results for weekday hours. Here, we see a more intuitive distribution of probabilities. At early hours, the probabilities are more variable. Some units may have very high probabilities (e.g., those with retired persons) while others are much lower. As the day goes on, the probabilities rise on average. There is less variability later in the evening when many different types of people are likely to be home. These results show a similar pattern to the analyses of U.S. Census Bureau [12] and Weeks et al. [14].

4.4. Future research and extensions

Future research will be dedicated towards improving the models and refining the approach for dealing with the nonrandom nature of the ACS contact attempts. To simplify the process, we used the same set of covariates in each model, but this does not have to be the case. Some variables were strong predictors of both attempt and contact across many hours of the day, such as whether the unit contained any person from the Medicare Enrollment Database. Other variables appear significant at only some hours of the day, so it will be important to understand whether these are actual effects or the natural result of making comparisons across multiple models and multiple times of the day. We will also consider how to improve the models without administrative records data. While many of the area-level covariates are strong predictors, the issue is that when scoring the model results to the 2015 test units, all units in the same area had very similar predicted probabilities. Further complicating the issue is the fact that many of the units with administrative records available were removed from the workload as part of the test; the units that remained to be visited were less likely to have any administrative records available.

For the 2015 Census Test, the predicted probabilities remained static throughout the data collection period. A future extension to consider is how to update the probabilities based on the outcomes of contact attempts made during the test. Durrant et al. [13] show

Table 5
Dependent variables for logistic models

Model	Universe	Dependent variable
P_h(attempt)	All ACS contact history cases	1 – Case attempted at given hour
		0 – Case not attempted at given hour
P_h(contact | attempt)	Cases attempted at given hour	1 – Successful contact
		0 – Unsuccessful contact

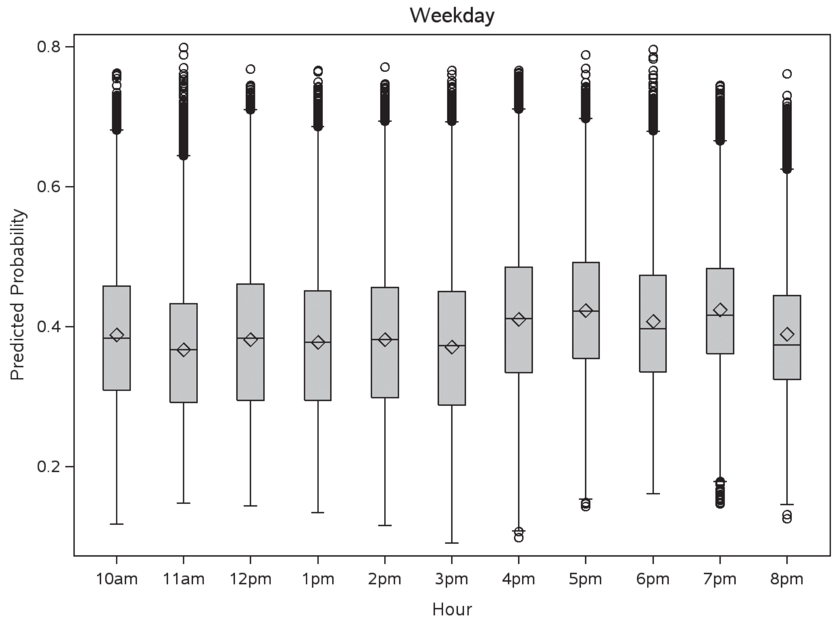

Fig. 1. Predicted contact probabilities for weekday hours: Initial approach.

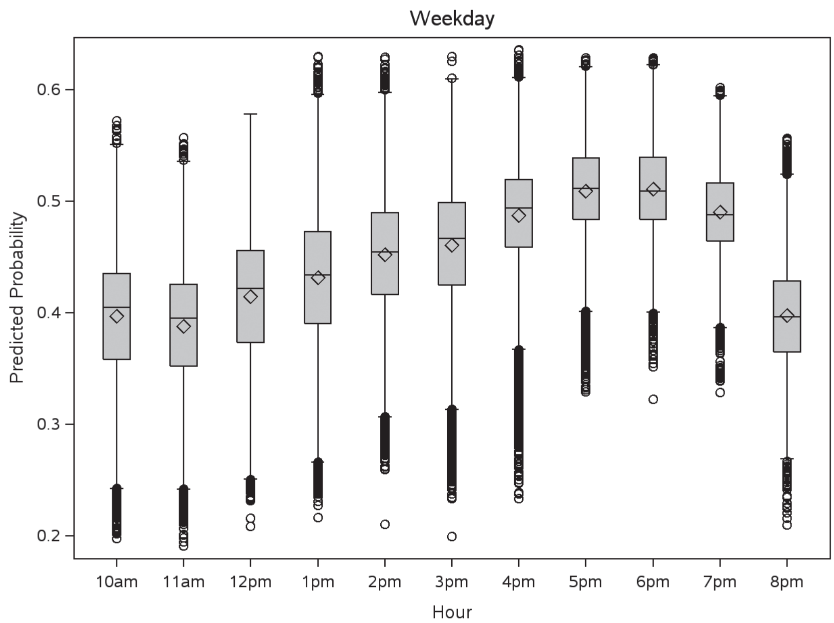

Fig. 2. Predicted contact probabilities for weekday hours: Two-step approach.

that paradata such as interview observations from early contact attempts are useful in predicting best times of contact. One of the issues with this in the current set up, as learned during the 2013 Census Test, is that the data collection period for the census is short. As discussed in Section 3, many units in the 2015 test received only one or two contact attempts. Therefore, there may not be enough time for the additional data obtained during the data collection to be used effectively for cases that are still open. We will need to investigate this further to understand the feasibility of such an approach.

5. Conclusions

This paper presents initial research on adaptive design for the 2020 Census. We showed how an optimization approach can be used to assign the maximum number of contact attempts to make during NRFU in different areas. This approach allows the number of contact attempts to vary across areas in order to equalize an aspect of data quality subject to workload constraints. This paper also discussed how ACS paradata were used to predict contact probabilities for housing units at each hour of the day. These predictions are an input for the ROCkIT case management system and were used to assign contact attempts to enumerators during the 2015 Census Test.

The research summarized here represents some first steps into adaptive design methodologies that could be used in future censuses. Future research is needed to understand how best to integrate these ideas and methods into the data collection operation so that adaptive decisions can be made in real time. Also, we must investigate how to best implement these methods on the larger scale of a census.

References

[1] S. Walker, W. Winder, G. Jackson and S. Heimel, 2010 Census nonresponse followup operations (NRO) assessment. 2010 Census Planning Memoranda Series, no. 190 [Internet]. 2012 Apr. Available from: https://www.census.gov/ 2010census/pdf/2010_Census_NRFU_Operations_Assessment.pdf.

[2] R.M. Groves and S.G. Heeringa, Responsive design for household surveys: Tools for actively controlling survey nonresponse and costs, *J R Stat Soc Ser A* **169**(3) (2006), 439–457.

[3] U.S. Census Bureau. 2020 Research and testing: 2013 Census test assessment. 2020 Evaluation, Analysis, and Assessment [Internet]. 2014 May. Available from: https://www.census.gov /content/dam/Census/programs-surveys/decennial/2020-census/2013_Census_Test_Assessment_Final.pdf.

[4] U.S. Census Bureau. 2020 Field reengineering concept of operations (Con Ops). Presentation to the 2020 Census Program Management Review [Internet]. 2014 June. Available from: http://www2.census.gov/census_2020/pmr_materials/2014-06-06/05_PMR_Field_Reengineering_Conops_6-6-14_v1-0.pdf.

[5] R.S. Rao, M.E. Glickman and R.J. Glynn, Stopping rules for surveys with multiple waves of nonrespondent follow-up, *Stat Med* **27** (2008), 2196–2213.

[6] J. Wagner and T. Raghunathan, A new stopping rule for surveys, *Stat Med* **29** (2010), 1014–1024.

[7] T. Mule, 2010 Census coverage measurement estimation report: Summary of estimates of coverage for persons in the United States. DSSD Census Coverage Measurement Memorandum Series #2010-G-01 [Internet]. 2012 May. Available from: https://www.census.gov/coverage_measurement/pdfs/g01.pdf.

[8] T. Mule, Administrative records modeling update. Presentation to the 2020 Census Program Management Review [Internet]. 2013 June. Available from: http://ftp2.census.gov/census _2020/pmr_materials/2013-06-21/5_PMR_Administrative_ Records_Modeling_v1-0_final.pdf.

[9] T. Huang and E. Hughes, Nonlinear optimization in SAS/OR® software: Migrating from PROC NLP to PROC OPTMODEL. Proceedings of the SAS Global Forum [Internet]. 2010; Paper 242–2010. Available from: https://support.sas.com/resources/papers/proceedings10/242-2010.pdf.

[10] J. Wagner, The fraction of missing information as a tool for monitoring the quality of survey data, *Public Opin Q* **74**(2) (2010), 223–243.

[11] R. Groves and M. Couper, Nonresponse in household interview surveys. New York: Wiley, 1998.

[12] U.S. Census Bureau. Who's home when. Working paper No. 37. U.S. Government Printing Office, Washington, D.C.; 1972.

[13] G.B. Durrant, J. D'Arrigo and F. Steele, Using paradata to predict best times of contact, conditioning on household interviewer influences, *J R Stat Soc Ser A* **174** (2011), 1029–1049.

[14] M.F. Weeks, B.L. Jones, R.E. Folsom and C.H. Benrud, Optimal times to contact sample households, *Public Opin Q* **44**(1) (1980), 101–114.