Optimizing Unit Nonresponse Adjustment Procedures After Subsampling Nonrespondents In The Economic Census

Laura Bechtel, Katherine J Thompson

U.S. Census Bureau[1]
Washington, DC 20233

1. Introduction

With budgets throughout the federal government tightening, many agencies have begun to consider adaptive strategies for data collection. One such strategy is to select a probability sample of nonresponding units for follow-up, instead of attempting a complete follow-up. Besides saving cost, this can improve the tabulation quality by replacing an essentially self-selected sample with a representative sample. In this framework, one can design adjustment procedures that mitigate or essentially eliminate unit nonresponse bias.

The Economic Directorate of the U.S. Census Bureau is investigating nonrespondent subsampling strategies for usage in the 2017 Economic Census. Although no unified methodology has been fully established, the design will include a systematic sample of nonrespondents, sorted by a unit measure of size. The primary goal of our research is to determine how best to improve the quality of the adjusted tabulations, given a systematic subsample of nonrespondents. A secondary goal is to determine the maximum allowable subsampling rate to achieve quality tabulations, given performance requirements on unit response rates. Whitehead et al. (2013) further study allocation strategies for selecting a subsample.

Technically, the Economic Census attempts a 100% follow-up of unit nonrespondents using a variety of procedures. These include mail-out reminders (letters and packages), automated phone reminders for selected cases (robot calls), and personalized telephone follow-up for other selected cases. Contact strategies vary by type of unit (single or multi-unit) as well as by size of unit and are designed to improve the quality of industry level tabulations by focusing on large units that contribute substantially to the totals. Thus, the intensive and costly follow-up operations are confined primarily to the largest cases, and the unconverted nonrespondents tend to be the smaller cases that receive little (if any) personal contact.

This report presents the results of a simulation study that uses empirical 2007 data from selected industries from seven of the trade areas included in the Economic Census[2]. In particular, we consider the quality effects on industry-level tabulations of key items using three separate adjustment strategies: (1) imputing a complete record for all nonrespondents (sampled or otherwise), sometimes referred to as "mass imputation"; (2) using adjustment cell weighting on respondents in the subsample; and (3) a combination of the two methods, applied to the subsample. Our design is very simple – a systematic subsample of all nonrespondents by industry – and does not incorporate any targeted allocation strategies.

By simplifying the sample selection procedure and essentially ignoring the effects of alternative contact procedures, we can examine the interactions between the sampling rate, the adjustment strategies, the nonrespondent conversion rates (the rate at which the sampled nonrespondents respond), and the assumed response mechanisms on the quality of the estimates over repeated samples. For our analyses, we rely extensively on historic reporting patterns for a program that has had a mail-out/mail-back collection with elective internet reporting available in selected industries. That said the 2017 Economic Census is planned as a 100% internet collection.

The 2012 Economic Census is well underway, and it is too late in the data collection cycle to implement a new adaptive collection strategy. Fortunately, there is plenty of time to prepare for the 2017 Economic Census. We conduct this research under the assumption that there will likely be a probability subsample of nonrespondents in the

[1] Any views expressed in this paper are those of the authors and not necessarily those of the U.S. Census Bureau.
[2] We exclude the Construction sector, which is a probability sample.

2017 Economic Census. The time is ripe to plan a more adaptive collection strategy. Collection and processing costs should be cheaper with the electronic collection, so budget can be diverted. We also expect large cases will respond by internet (parallel to current surveys), so the environment may be better for subsampling nonresponding smaller units and actually obtaining response data from them.

2. Economic Census Background

The U.S. Census Bureau conducts an Economic Census in years ending in 2 and 7, mailing out over four million census forms to business establishments that provide commercial services to the public and other businesses. Data are collected at the establishment level and are classified according to the North American Industry Classification System (NAICS). The Economic Census coverage extends to establishments in eighteen non-farm economic sectors, including wholesale trade, retail trade, finance, insurance, real estate, services, transportation, communication, utilities, manufacturing, mining, and construction. For processing purposes, the finance, insurance, and real estate industries are grouped into a single trade area (FIRE), as are the transportation, communication, and utilities industries (Utilities). Census data are tabulated and released to the public on a fixed schedule, with industry level tabulations released first, followed by state and other selected geographic breakdown industry tabulations and summary/series reports.

Although there is one Economic Census, the collected items differ by trade area, and may differ by industry within trade area. Each establishment provides values for four *general statistics items*: annual payroll (payroll), 1st quarter payroll, 1st quarter employment or average employment[3] (employment), and total receipts/value of shipments (receipts). Often, the Census Bureau's business register contains administrative data for these items. Additional general statistics items (requested from each establishment) differ by sector: for example, the wholesale trade requests beginning and ending inventory value from all establishments, and the services industries trade area requests operating expenses from all establishments in tax-exempt industries. In general, usable administrative data are not available for these other items. In addition, the Economic Census collects industry-specific items. Industry-specific total items – for example, number of hotel rooms – are referred to in-house as trailer data; detailed category breakdowns of totals that differ by industry are referred to as product line data. In these cases, administrative and auxiliary data are not available, and the item response rates for these items tend to be low.

The Economic Census data undergo extensive review at the micro- and macro-levels. The micro-level review procedures are designed to obtain accurate national industry level tabulations. Consequently, the national-level ratio edit and imputation parameters can be quite different from the industry-average parameters and edit parameters in subdomains such as industry-state. Each sector develops ratio edit and imputation parameters from the prior economic census data ("cold deck" parameters). Imputation cells differ by trade area, ranging from the very broad six-digit industry level to finer breakdowns such as tax-exempt status or legal form of operation. Likewise, data inclusion rules for parameter development vary by trade area: for example, some trade areas use only reported or administrative data, whereas others use the complete imputed data set. The effect of inflation or price change between censuses is somewhat mitigated in dollar comparisons (e.g. annual payroll/1st quarter payroll) but can be quite pronounced in wage per employee ratios. Consequently, many trade areas update their ratio editing and imputation parameters mid-processing cycle using current census data ("warm deck").

With the exception of the construction sector, the Economic Census is a cut-off sample that mail forms to establishments that exceed a specified trade-area specific unit size cut-off threshold (Lineback et al, 2012). "Complete" records are created via imputation for the cases that are not mailed a form, as well as for unit nonrespondents. When possible, administrative data are substituted for the missing general statistics items, and ratio or other model imputations are used to complete the remainder. Treatment of missing trailer data and product line data differs by trade area and is not further discussed. For single unit (SU) establishments, administrative data are often an excellent substitution option. This is rarely the case with the multi-unit (MU) establishments, as tax data are provided at the company or employer identification number (EIN) level, comprising several establishments. In these

[3] 1st quarter employment is collected from the wholesale trade, retail trade, finance, insurance, real estate, services, transportation, communication, and utilities sectors, whereas 1st – 4th quarter employment is requested separately and averaged across quarters for the manufacturing, mining, and construction sectors.

cases, the aggregated data are allocated to the individual establishment level, and the expected administrative data value is not necessarily equal to the census value.

The primary purpose of the Economic Census is to provide reliable benchmark industry-level tabulations. Unit nonresponse follow-up is conducted in several stages by calendar date, and follow-up procedures vary by type of unit (SU or MU) and by size of unit. Account managers maintain personal contacts with the largest companies. Otherwise, there are four phases of follow-up that begin on fixed calendar dates. Follow-up treatments at the first phase range from reminder letters, to packages (reminder letter plus form), to robot phone call reminders, and lastly to personal calls. As follow-up phases progress, the reminder letters become more stringent, the package delivery may be certified mail, and the likelihood of a personal phone call from a centralized telephone center for large SU and MU unit establishments increases. Because phone calls are quite expensive, they are limited to the largest cases in low responding industries at the later follow-up phases, determined subjectively by trade area experts.

The Economic Census is a very large program, and nonresponse follow-up is costly. The currently-used follow procedures are designed to maximize two sets of response rates: the unit response rate (unweighted proportion of eligible units) and the total quantity response rates (approximately 1 – imputation rate) for the general statistics items (Thompson and Oliver, 2012). To maximize the total quality response rates, the follow-up procedures focus on obtaining data from the larger units. The response set is therefore self-selecting, except for the large cases, and not necessarily representative of the parent population. This in turn has “ripple effects” on the imputation procedures, particularly ratio imputation.

In addition to reducing cost, there are quality advantages to replacing the self-selecting sample with a probability sample. First, having a smaller pool of units for follow-up makes it easier to obtain respondent data from a variety of units, instead of just the large ones. With a small well-designed sample, program managers can easily track the respondent pool by unit characteristic (e.g., unit size, industry). This opens the door, so to speak, for implementing adaptive or responsive data-collection procedures. From a survey design perspective, the probability sampling allows the design of adjustment procedures that mitigate nonresponse bias effects on totals.

3. Simulation Study Design

In practice, adaptive designs require consideration of many layers. Examples include developing allocation strategies, creating rules for eligible units for subsampling, determining the timing for subsampling during data collection (for example, it might be best to maintain the current procedure of sending reminder letters to all delinquent units early in the collection process), and developing data collection strategies. Our research ignores many of these factors. Instead, we consider a very simple scenario, where any single or multi-unit establishment that was mailed a form and did not respond is eligible for the systematic sample as presented in Figure 1.

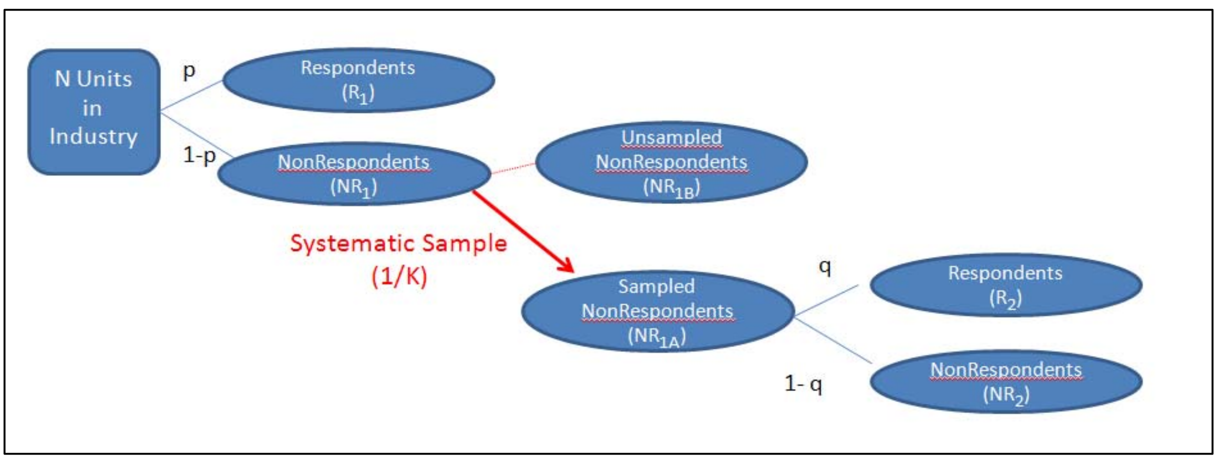

Figure 1: Big Picture of Simulation Study in a Single Industry

Here, the subsampling begins at the point in the collection process where approximately *p*-percent of the forms have returned [Note: we suppress an industry subscript for simplicity, but allow different response probabilities by industry in the study]. A portion of the sampled nonrespondents will eventually be respondents. In contrast to Whitehead et al. (2013), we do not attempt to track converted rates over time. Throughout the remainder of the paper, we use the following notation:

y = value of the data item
R_{1i} = an indicator variable for unit i's response before subsampling
R_{2i} = an indicator variable for unit i's response after subsampling
p = Probability that a unit in the industry has responded before implementing subsampling = $P(R_{1i}=1)$
q = Probability that a subsampled unit will eventually provide a response = $P(R_{2i}=1 \mid R_{1i}=0)$
A = Probability that a subsampled unit will respond = $(1-p)*q$
B = Probability that a unit responds to the Economic Census= p+A = $P(R_i=1)$
R_1 = Number of units that responded before subsampling
NR_1 = Number of units that did not respond before subsampling (i.e., the size of the subsampling frame)
K = subsampling interval (systematic sample)
NR_{1A} = Number of nonrespondents that were selected for follow-up
NR_{1B} = Number of nonrespondents that were not selected for follow-up
R_2 = Number of subsampled nonrespondents that were converted to respondents (the number of followed up cases that responded)
NR_2 = Number of unconverted subsampled nonrespondents

For the simulation, we use data from the 2007 Economic Census in the selected set of industries provided by the trade area experts. We produce tabulations using the trade areas' respective editing cells instead of the published six-digit North American Industry Classification Series (NAICS) code, since many editing cells are subdomains of their respective industries. The industry datasets contained full-year reporter establishments whose response data were included in the 2007 Economic Census tabulations. Table 1 provides the number of studied industries and editing cells within each trade area.

Table 1: Number of Studied Industries and Editing Cells by Trade Area

Trade Area	Number of Industries	Number of Editing Cells	Trade Area	Number of Industries	Number of Editing Cells
FIRE	8	10	Services	8	10
Manufacturing	6	6	Utilities	8	11
Mining	3	3	Wholesale	7	34
Retail Trade	8	8			

Each industry micro-dataset contains the final tabulated values of payroll, 1st quarter payroll, employment, and receipts as well as the available corresponding administrative data item values (there are some missing values in the administrative data).

We randomly induce unit nonresponse in the population data using response mechanisms discussed in Section 3.1 below, repeating the process independently 1,000 times per response mechanism variation. In each replicate, we select three systematic subsamples of nonrespondents per response mechanism of K=1, 2, and 3 (see Section 3.2). Within subsample and replicate, we use three *adjustment procedures* to account for unit nonresponse:

Impute Impute a complete record for *all* nonrespondents, so that there are N units with complete data in the imputed dataset (mass imputation). See Section 3.3 for details on imputation models.

Impute/Weight Impute a complete record for all unconverted *sampled* nonrespondents, then multiply the weight of all units in the subsample by K, so that there are $R_1 + NR_{1A}$ with complete data in the imputed dataset. See Section 3.3. for details on weighting models.

Weight Increase the weights of the R_2 subsampled respondents (converted nonrespondents) to account for the NR_2 unconverted sampled nonrespondents, then multiply the weight of the R_2 respondents by K, so that there are $R_1 + R_2$ units with complete data in the final dataset.

We study the combined effects of adjustment procedure, response mechanism, subsampling rate, and adjustment method (imputation or weighting) on three general statistics items: payroll, employment, and receipts.

3.1. Response Mechanism

We consider three separate response mechanisms:

(1) Uniform response before and after subsampling (all units equally likely to respond with probabilities p and q, respectively);
(2) Nonignorable response before subsampling (i.e., p is directly dependent on the value of a collected data item) and uniform response after subsampling; and
(3) Covariate dependent response before subsampling (i.e., p is directly dependent on a variable that is collected but is not under study) and uniform response after subsampling.

The first response mechanism is the most tractable and the least realistic for a business program. To simulate response mechanisms (2) and (3), we fit logistic regression models on our empirical data with unit response to the 2007 census as the dependent variable and payroll as independent (predictor) variables. Thus, the probability of response is nonignorable for the payroll estimates and covariate-dependent for the employment and receipts estimates (See Table 2).

Table 2: Response Mechanism Applied to Studied Data Items

Response Mechanism	Payroll	Employment	Receipts
Uniform	X	X	X
Nonignorable	X		
Covariate-dependent		X	X

3.2. Allocation

Under a uniform response mechanism for both p and q, it can be shown that the overall probability of responding given a systematic sample is given by $B_K = p + (1/K)*(1-p)*q$. With the Economic Census, follow-up starts when approximately 40% of sample has responded [Note: this is approximate because reminder letters are sent on a fixed calendar date]. Subject matter experts suggested that the unit nonresponse conversion rate (A) from later stages of follow-up is approximately 20%. Table 3 provides B_K for K = 1, 2, 3, 4, and 5 under a uniform response mechanism at both stages, considering three values of p and nine values of q. The highlighted cells indicate where A≈ 20%, before subsampling.

Table 3: Response Probabilities Given a Uniform Response Mechanism and a Systematic Subsample

					B								B		
p	q	A	K=1	K=2	K=3	K=4	K=5	p	q	A	K=1	K=2	K=3	K=4	K=5
0.4	0.1	0.06	0.46	0.43	0.42	0.42	0.41	0.6	0.1	0.04	0.64	0.62	0.61	0.61	0.61
0.4	0.2	0.12	0.52	0.46	0.44	0.43	0.42	0.6	0.2	0.08	0.68	0.64	0.63	0.62	0.62
0.4	**0.3**	**0.18**	**0.58**	**0.49**	**0.46**	**0.45**	**0.44**	0.6	0.3	0.12	0.72	0.66	0.64	0.63	0.62
0.4	**0.4**	**0.24**	**0.64**	**0.52**	**0.48**	**0.46**	**0.45**	0.6	0.4	0.16	0.76	0.68	0.65	0.64	0.63
0.4	0.5	0.30	0.70	0.55	0.50	0.48	0.46	**0.6**	**0.5**	**0.20**	**0.80**	**0.70**	**0.67**	**0.65**	**0.64**
0.4	0.6	0.36	0.76	0.58	0.52	0.49	0.47	0.6	0.6	0.24	0.84	0.72	0.68	0.66	0.65
0.4	0.7	0.42	0.82	0.61	0.54	0.51	0.48	0.6	0.7	0.28	0.88	0.74	0.69	0.67	0.66
0.4	0.8	0.48	0.88	0.64	0.56	0.52	0.50	0.6	0.8	0.32	0.92	0.76	0.71	0.68	0.66
0.4	0.9	0.54	0.94	0.67	0.58	0.54	0.51	0.6	0.9	0.36	0.96	0.78	0.72	0.69	0.67
0.5	0.1	0.05	0.55	0.53	0.52	0.51	0.51								
0.5	0.2	0.10	0.60	0.55	0.53	0.53	0.52								
0.5	0.3	0.15	0.65	0.58	0.55	0.54	0.53								
0.5	**0.4**	**0.20**	**0.70**	**0.60**	**0.57**	**0.55**	**0.54**								
0.5	0.5	0.25	0.75	0.63	0.58	0.56	0.55								
0.5	0.6	0.30	0.80	0.65	0.60	0.58	0.56								
0.5	0.7	0.35	0.85	0.68	0.62	0.59	0.57								
0.5	0.8	0.40	0.90	0.70	0.63	0.60	0.58								
0.5	0.9	0.45	0.95	0.73	0.65	0.61	0.59								

The 2006 Federal Register Notice issued by the Office of Management and Budget guidelines states that samples should be designed to achieve the maximum response rate possible. Using the B_K from Table 3 as a proxy for the expected unit response rate (URR), we conclude that (1) subsampling should not begin until at least 60% of the forms in an industry have been received and (2) subsampling rates need to be 3 or smaller. Moreover, the larger sampling rate (K=3) is only feasible – given the historic unit conversion rate -- if the average URR in most of the Economic Census industries is 60% or larger before follow-up begins. [Note: this parallels the independent findings presented in Whitehead et al. (2013).] In our simulation study, we consider p = 0.4, 0.5, 0.6, and 0.7 with q varying from 0.2 to 0.7 (by 0.10) within value of p.

Under the nonignorable and covariate-dependent response mechanisms described in Section 3.1, each replicate has a single value of p, but the values of q can differ. We optimistically varied q from 0.4 to 0.7 (by 0.1), although we focus in Section 4 on q = 0.5, which yields an expected unit conversion rate of 0.2 when p = 0.6. Table 4 presents summary statistics by trade area on these estimated values of p with q = 0.5.

Table 4: Expected p ($P(R_{1i}=1)$), A ($=(1-p)*q$), and B ($P(R_i=1)$ Under Nonignorable Response Mechanism

Trade Area	Number of Editing Cells		p	A	B	Trade Area	Number of Editing Cells		p	A	B
FIRE	10	Median	0.51	0.25	0.75	Services	10	Median	0.57	0.22	0.78
		Min	0.32	0.34	0.66			Min	0.25	0.38	0.62
		Max	0.93	0.03	0.97			Max	0.78	0.11	0.89
Manufacturing	6	Median	0.60	0.20	0.80	Utilities	11	Median	0.54	0.23	0.77
		Min	0.56	0.22	0.78			Min	0.49	0.25	0.75
		Max	0.65	0.18	0.82			Max	0.81	0.10	0.90
Mining	3	Median	0.44	0.28	0.72	Wholesale	34	Median	0.56	0.22	0.78
		Min	0.33	0.33	0.67			Min	0.35	0.32	0.68
		Max	0.86	0.07	0.93			Max	0.73	0.13	0.87
Retail Trade	8	Median	0.70	0.15	0.85						
		Min	0.53	0.23	0.77						
		Max	0.83	0.09	0.91						

3.3. Imputation/Weighting Procedures and Imputation/Weighting Models

The Economic Census uses *composite imputation* to account for missing and invalid data items as well as for unit nonresponse, and imputation methods and models vary by item (Wagner, 2000). Imputation for each item is attempted in a pre-specified sequence, with the deterministic methods attempted first (e.g., rescaled reported data, logical edits, administrative data), and model imputation attempted only after all direct substitution methods have been exhausted. Many programs use historic data to develop the initial set of model imputation parameters, although several do create “warm deck” imputation parameters later in the processing cycle. With unit nonresponse, administrative data substitution is the most frequently employed imputation method when such data are available.

Several imputation models are available for the studied data items in the production setting. For simplicity, we use the following three imputation models in our composite imputation simulations:

Administrative Data $y_i' = \tilde{y}_i$, where $\tilde{y}_i$ is the administrative data value for unit i

Clerical Change: $y_i' = f(y_i)$, where $f(y_i)$ is a simple algebraic operation on the reported value

Ratio Impute: $y'_i = f^m x_i$, where x_i is an auxiliary variable available for all units (may be an imputed value) and m refers to the data used for parameter development (cold deck = historic or warm deck = current)[4]. The formula that we used to develop ratio imputation parameters differs by trade area, as shown in Table 5.

Table 5: Formulae For Ratio Imputation Parameters (Developed Within Edit Cell)

Trade Area	Parameter Computation (f^m)	Method
Manufacturing and Mining	$f^C = median_{\forall i \in N_{2002}}(x_i / y_i)$	Cold Deck
	$f^W = median_{\forall i \in R_1 \cup R_2}(x_i / y_i)$	Warm Deck
All Others	$f^C = \sum_{i \in N_{2002}} x_i / \sum_{i \in N_{2002}} y_i$	Cold Deck
	$f^W = (\sum_{i \in R_1} x_i + \sum_{i \in R_2} Kx_i) / (\sum_{i \in R_1} y_i + \sum_{i \in R_2} Ky_i)$	Warm Deck

The imputed estimate of item y is given by $\hat{y}^v = \sum_{i \in R_1} y_i + \sum_{i = NR_1} w_i^v \tilde{y}_i^m I_i^v$, where $I_i^V \equiv 1_{\forall i}$ and $w_i^v = 1$ when v =Impute and $I_i^V = 1$ for $i \in R_2$ (and =0 otherwise) and $w_i^v = K$ when v = Impute/Weight.

Table 6 provides summary frequencies of usage from each imputation model by trade area. To obtain these frequencies, we categorized all model imputations into the "ratio imputation" category[5]; we grouped administrative data substitution, logical edit substitutions, and direct substitution (current data) into the administrative data category; and classified the remaining changes as clerical.

Table 6: Median Frequencies of Imputation Model by Trade Area (In Percentages)

	Payroll			Employment			Receipts		
Trade Area	Admin.	Ratio	Clerical	Admin.	Ratio	Clerical	Admin.	Ratio	Clerical
FIRE	91.66	1.65	4.55	89.19	8.13	3.72	34.26	44.09	19.72
Manufacturing	76.25	6.76	11.97	3.25	82.65	10.54	0.88	92.65	1.23
Mining	84.34	9.09	0.00	66.67	33.33	3.63	2.41	93.42	1.20
Retail Trade	94.97	0.68	2.97	91.55	4.18	3.82	40.78	48.15	7.72
Services	89.14	2.70	3.49	87.58	3.16	6.61	39.71	40.98	17.91
Utilities	96.41	3.59	0.00	90.74	6.29	1.92	37.94	47.62	19.17
Wholesale Trade	93.09	2.80	3.68	88.09	9.79	2.39	33.37	45.69	22.69

In the simulation, we randomly assigned imputation models to nonresponding units by the editing cell (not trade area) frequencies, independently determining imputation model by item within unit. The ratio imputation model for payroll (imputed first) uses 1st quarter payroll as auxiliary variable when available, then administrative receipts, and lastly administrative employment, whereas the other two imputation models use the current value of payroll (imputed or reported) as auxiliary variable. Because of difference in the population distributions, the ratio parameters for the manufacturing and mining industries are computed differently than the other sectors (see Table 5). The manufacturing and mining trade areas compute ratio imputation parameters as the average value of the studied ratios; the other trade areas compute ratio imputation parameters as the sum of the numerator value divided by the sum of the denominator value (Thompson and Sigman 1996).

---

[4] For payroll, we use administrative data as an auxiliary variable, usually 1st quarter payroll. For the other variables, we use the imputed/reported value of payroll as auxiliary variable.

[5] Actual methods include edit-cell average parameters, auxiliary trend imputation, regression model imputation, and machine-rescaling corrections (e.g., divide by 1000).

To prevent atypical values from overly influencing the editing cell averages, we excluded outlying values from cold and warm deck parameter imputation basing using resistant fences rules (Hoaglin et al, 1983). The general statistics items collected by the Economic Census are subjected to ratio edits as part of the micro-data review process, and the "Plain Vanilla" system that implements these edits ensures that the tabulated microdata satisfy *all* of the ratio edits (Wagner, 2000). Unfortunately, we did not have a comparable method of evaluating the administrative data used in our study. In practice, the administrative data for the general statistics items would be pre-edited. Because we could not do this, the effects of administrative data imputation on the imputed tabulations' bias (see Section 4) is inflated in a few cases [Note: in general, the ratio of the administrative data to tabulated data item was between (0.95, 1.05)].

Table 7 presents summary statistics on correlation coefficients by trade area for the latter two ratio models (the correlation of payroll to annual payroll is always ≈ 0.99). Notice that the strength of association for these models is not consistent, especially for wholesale trade and utilities, which comprise several very different types of industries (e.g., wholesale trade industries include agents/broker classifications and merchant wholesaler classifications).

Table 7: Summary Statistics on Correlation for Ratio Imputation Models by Trade Area

Trade Area	Payroll and Employment			Payroll and Receipts		
	Median	Minimum	Maximum	Median	Minimum	Maximum
FIRE	0.80	0.58	0.97	0.72	0.41	0.99
Manufacturing	0.94	0.90	0.96	0.85	0.82	0.92
Mining	0.95	0.85	0.99	0.86	0.52	0.97
Retail Trade	0.84	0.80	0.98	0.87	0.55	0.95
Services	0.90	0.16	0.99	0.94	0.77	0.99
Utilities	0.92	0.82	0.99	0.75	0.11	0.99
Wholesale Trade	0.89	0.22	0.98	0.73	0.26	0.98

With payroll and employment, ratio and clerical imputation are rarely used. Consequently, the composite imputation simulations primarily assess the statistical properties of administrative data imputation, especially for payroll. With receipts, administrative data and ratio imputation are often more evenly balanced, but there is a high usage of administrative data except for the manufacturing and mining establishments. Recall that administrative data are not available for all Economic Census data items. Consequently, it would be unwise to make any broad statements about imputation or adjustment procedures based solely on composite imputation simulation. To study the unconfounded effects of adjustment procedure alone (given an imputation or weighting method), we also create fully imputed data sets using only ratio imputation.

Adjustment cell weighting is used instead of imputation to account for unit nonresponse in many applications. If the nonrespondents are a random subsample, then weighting the respondent sample units by the inverse response rate will yield essentially unbiased estimates (Kalton and Flores-Cervantes, 2003). Similarly, if the administrative data are viable substitute values for missing reported data, then more precise results would be obtained by imputing as many cases as possible using administrative data, then using adjustment cell weighting to account for remaining nonresponse in the subsample. We refer to the latter method as the "hybrid weighting" approach. Note that nonresponse weighting adjustment options are only considered for the systematic subsamples (K=2 and K=3) and that the hybrid method is only considered in comparison with composite imputation.

The weighted estimate of item y is given by $\hat{y}^v = \sum_{i \in R_1} y_i + \sum_{i = NR_1} w_i^v f_i^v y_i I_{si} I_{ri}$, where $I_{si} = 1$ when unit i is included in the nonrespondents subsample and 0 otherwise, $I_{ri} = 1$ when subsampled unit i responds and 0 otherwise, $f_i^v = NR_{1A} / R_2$ when v = Weighting (Traditional) and $f_i^v = NR_{1A} / (R_2 + NR_{2A})$ where NR_{2A} is the number of nonresponding sampled units that are imputed with administrative data when v = Weighting (Hybrid), and $w_i^v = K$.

3.4. Simulation Design, Part II

Table 8 summarizes the adjustment method combinations produced for each response mechanism in our simulation study.

Table 8: Adjustment Method Combinations in Simulation

Adjustment Procedure	Imputation or Weighting Model	Composite Imputation			Ratio Imputation		
		K=1	K=2	K=3	K=1	K=2	K=3
Mass Imputation	Cold Deck	X	X	X	X	X	X
	Warm Deck	X	X	X	X	X	X
Impute/Weight	Cold Deck		X	X		X	X
	Warm Deck		X	X		X	X
Weight	Traditional		X	X		X	X
	Hybrid		X	X			

There are important differences between each adjustment procedure:

- **Impute**: A complete record is created for all units, regardless of sampling. Subsampling effects are only present via the warm deck imputation parameters. For composite imputation, subsampling effects are essentially absent otherwise. Mass imputation is a less variable adjustment procedure than the other two because of the larger sample size. With composite imputation, the variability is further reduced because the same administrative data values are used in each replicate;
- **Impute/Weight**: The subsampling effects are present in the tabulations since only sampled nonrespondents are included in the tabulations. The reduced sample size and the weight adjustment for subsampling will increase the variability of this procedure over mass imputation. Subsampling effects will be increased when warm deck parameters are used (over cold deck); and
- **Weight**: This is the most variable of the three adjustment procedures due to the reduced sample size. If the response rate from the subsample is very low, then this procedure will have very high variability from the adjustments alone. Moreover, if the converted sampled respondents are not a representative subsample, then this procedure will yield biased estimates as well.

Section 4 summarizes the results of the simulation. There are several caveats. First, we randomly assign response propensity in both stages. In reality, with 100-percent follow-up (K=1), larger units are simply more likely to provide usable response data than small units. The nonignorable/covariate dependent response mechanism scenarios are more realistic, but may allow for more nonresponse in large units than in reality. Moreover, we assume that the response mechanism for respondent conversion is uniform. Of course, the purpose of adaptive data collection strategies is to ensure that this is not the case. The simulation procedure assumes that the tabulated micro-data are correct, which is patently unrealistic, but is much easier to model and less reliant on unprovable assumptions than a more realistic model. With the imputation procedures, we randomly assign imputation models to items within unit, whereas in practice, the determination of imputation model is more purposive – and many of the implemented imputation models use historic data for the same establishment. Finally, the simulation design ensures that the *impute* adjustment procedure will be the least variable. Likewise, the *weight* procedure will always be the least biased of the three procedures because of the random response propensity assignment at the second stage.

4. Simulation Study Results

4.1. Evaluation Criteria

Let T denote the response mechanism (uniform-uniform, nonignorable-uniform, covariate dependent-uniform) and p and q be the realized values of the first and second stage response probabilities. For each response mechanism $(pq)^T$ and subsampling rate K, we compute the following statistics by item within trade area editing cell:

Relative Bias $$\left((\sum_r \hat{Y}_{Kmr}^{v(pq)^T} / 1000) / Y \right) - 1$$

Relative Variance $\left(\left(\sum_r (\hat{Y}_{Kmr}^{v(pq)^T} - \hat{\bar{Y}}_{Kmr}^{v(pq)^T})^2 / 1000 \right) / \sigma_y^2 \right) - 1$

where v = adjustment method (Impute, Impute/Weight, Weight), m = parameter type (cold or warm deck) when v = Impute or Impute/Weight and m = weighting method (traditional or hybrid) otherwise, r indexes the replicate, Y is the true population value, and σ_y^2 is the population variance.

The objective of the adjustment procedures is to minimize the bias induced by unit nonresponse i.e. to maximize the probability that the relative bias approaches zero. Comparing the variance of each adjustment procedure provides insight into the instability induced by subsampling and the precision of each strategy.

4.2. Comparison of Adjustment Methods

In the sections below, we present results for the nonignorable-uniform and covariate dependent-uniform response mechanisms with respondent conversion probability q = 0.5. This set of results uses a 1-in-2 subsample (K=2), since the patterns discussed below are generally the same for K=3.

4.2.1. Relative Bias

Figures 2 and 3 present the median relative biases by trade area for payroll using composite imputation and ratio imputation. In these figures – and all others presented in this section – I_C and I_W represent Impute with cold deck or warm deck, IW_C an and IW_W represent Impute/Weight with cold or warm deck, and W_H and W_T represent weighting hybrid and weighting traditional. The hybrid method is not available for the ratio imputation data sets.

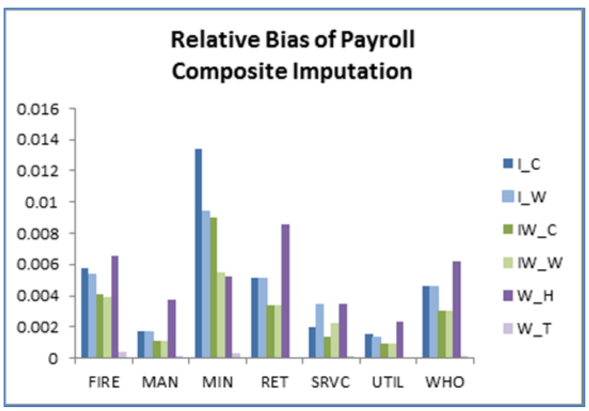

Figure 2: Relative Bias of Payroll with Composite Imputation and Nonignorable Nonresponse (K=2)

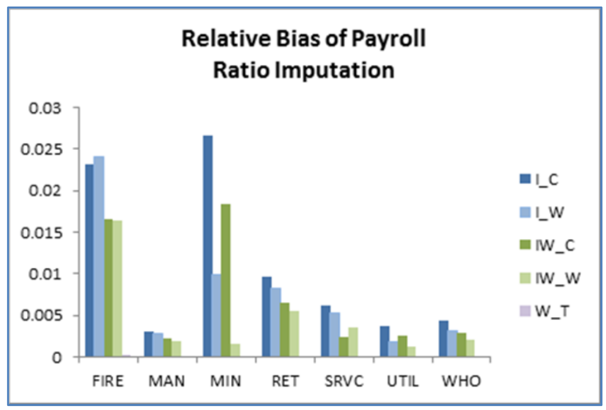

Figure 3: Relative Bias of Payroll with Ratio Imputation and Nonignorable Nonresponse (K=2)

With composite imputation, the relative biases are all very low. Recall that payroll is primarily imputed using administrative data in most trade areas when composite imputation is used. Even so, the traditional weighting (W_T) consistently yields the least biased estimates. This is largely a function of the simulation design, which always employs a uniform response mechanism at the second phase. With weighting, the effect of the differences between census and administrative data is demonstrated by the higher hybrid weighting method biases and in the Impute tabulations. As mentioned in Section 3.3, the administrative data used in the simulation are not edited, hence a portion of the bias in the simulation may be induced by a few outlying values used in substitution. That said, the Impute/Weight results are not a bad compromise, especially when warm deck imputation is used. Note that within adjustment procedure, warm deck tends to be *slightly* less biased than cold deck again another consequence of their heavy use of administrative data and light use of ratio imputation. The patterns are very similar when only ratio imputation is used. However, the relative biases greatly increase and the difference between warm deck and cold deck biases within adjustment method are clearer and we can see the benefits of developing parameters from a representative subsample.

Figures 4 and 5 present the relative bias results for employment. With the composite imputation, it is difficult to find a clear pattern with any of the trade areas except for Manufacturing and Mining, which rely primarily on ratio imputation. In these trade areas, using warm deck parameters within method is preferable. However, when the confounding effects of administrative data are removed, a clearer pattern emerges, with warm deck parameters tending to yield less biased estimates than their cold deck counterparts within adjustment procedure, and Impute/Weight adjustment procedure likewise yielding less biased estimates than the Impute procedure.

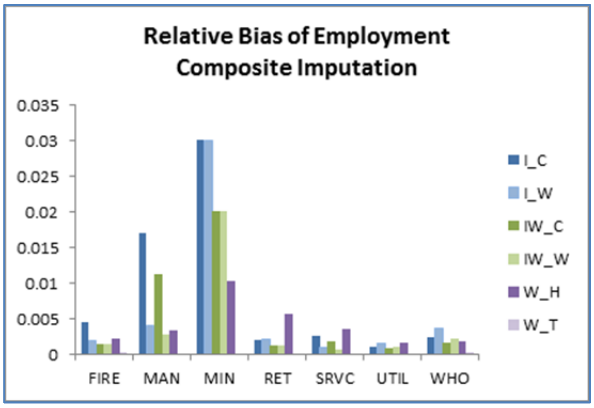

Figure 4: Relative Bias of Employment with Composite Imputation and Covariate Dependent Nonresponse (K=2)

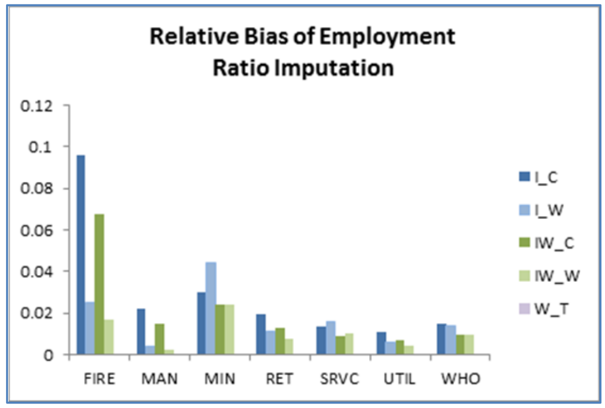

Figure 5: Relative Bias of Employment with Ratio Imputation and Covariate Dependent Nonresponse (K=2)

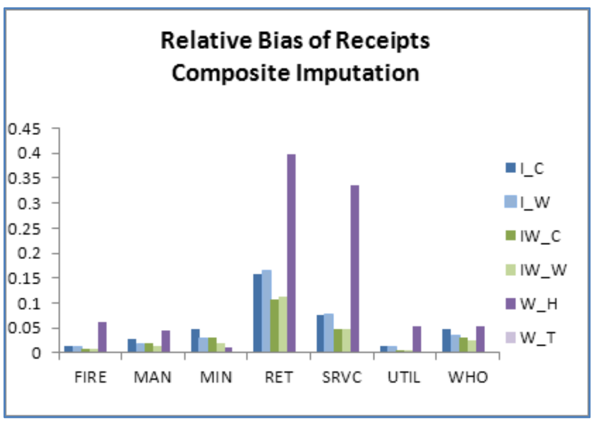

Figure 6: Relative Bias of Receipts with Composite Imputation and Covariate Dependent Nonresponse (K=2)

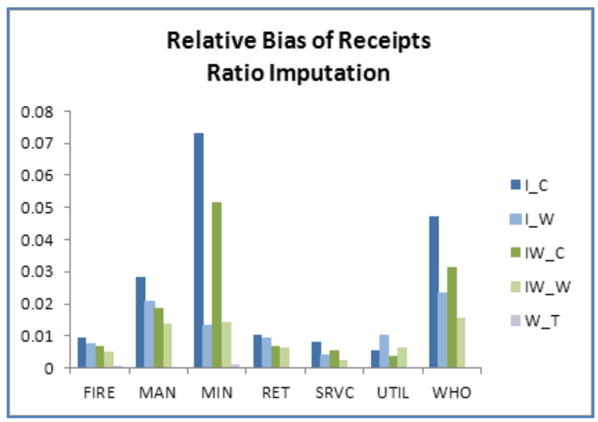

Figure 7: Relative Bias of Receipts with Ratio Imputation and Covariate Dependent Nonresponse (K=2)

Finally, Figures 6 and 7 present the relative bias results for receipts. In general, the relative bias of the composite imputation is much larger. However, when only ratio imputation is used, we see the same patterns as with employment.

To summarize, traditional weighting is always the least biased of the considered adjustment procedures, but adding administrative data imputation to the weighting procedure has very unpredictable results. In all these comparisons, the composite imputation muddles the picture a bit. However, when only one imputation method is used, a very clear pattern with imputation methods and adjustment methods emerges. Warm deck parameters – which take advantage of a representative subsample of recent data – yield less biased estimates than their cold deck counterparts. The Impute/Weight adjustment procedure is less biased than the Impute procedure and the differences in relative bias between the two procedures is not trivial.

4.2.2. Relative Variances

Table 9 provides the median relative variance estimates (K=2) for composite by trade area, using the color scheme and abbreviations from Figures 2 through 7.

Table 9: Median Relative Variance (K=2) by Trade Area with Composite Imputation

Item	Trade Area	I_C	I_W	IW_C	IW_W	W_H	W_T
Payroll	FIRE	11.69	11.09	17.71	16.83	24.64	44.57
	Manufacturing	6.12	5.98	13.59	13.41	59.69	286.95
	Mining	1.87	1.82	19.21	20.47	30.42	35.95
	Retail Trade	231.95	231.96	331.17	330.71	383.27	1147.75
	Services	0.64	0.69	7.64	7.68	53.65	96.93
	Utilities	-0.72	-0.72	-0.44	-0.44	3.15	13.00
	Wholesale Trade	0.77	0.75	2.31	2.28	8.51	35.28
Employment	FIRE	52.14	48.57	152.96	48.17	141.33	325.17
	Manufacturing	38.27	29.29	62.57	49.63	114.78	459.76
	Mining	8.72	9.10	17.56	17.57	26.21	53.33
	Retail Trade	227.90	222.60	674.31	670.51	979.75	2340.94
	Services	12.21	11.60	83.22	70.68	125.41	309.30
	Utilities	0.69	0.69	11.56	10.57	17.16	44.39
	Wholesale Trade	6.77	6.28	22.41	22.26	30.10	68.17
Receipts	FIRE	937.11	937.42	1354.86	1355.27	8157.74	89.49
	Manufacturing	290.53	286.79	116.87	113.65	59274.51	356.38
	Mining	2.97	2.66	19.91	20.69	82.41	119.89
	Retail Trade	391775.04	391832.65	640271.81	640225.43	2246010.54	1814.45
	Services	19768.35	19764.71	31241.63	31238.16	553828.25	119.70
	Utilities	221.71	222.03	329.90	330.17	2061.33	41.56
	Wholesale Trade	33.47	34.62	76.56	76.21	131.17	101.82

The variance of each adjustment procedure is compared to the true population variance for the item in the editing cell. We expected these empirical variances to be larger than their population counterparts, at least due to sampling effects. Nevertheless, we were alarmed by the extremely large Retail trade statistics, especially for receipts. Table 10 below presents the corresponding statistics using ratio imputation.

Removing the administrative data component changes the picture, but not consistently. If the imputation model is strong and administrative data are not heavily used, then the relative variances do not change much. In the Retail Trade receipts case, the variance is greatly reduced by using the ratio estimator because (1) the composite estimator heavily utilizes administrative data substitution and (2) our administrative data were not validated and extremely large values were inadvertently substituted. Furthermore, the comparability at the unit level of the administrative data and the census data definitely has an effect on the variance, but the magnitude of the effect is not always predictable.

Even so, for all of the items studied, the results presented in Table 9 and 10 can be summarized as follows:

- Adjustment cell weighting causes unacceptable increases in variance, showing the cumulative effect of probability subsampling, unit nonresponse in the subsamples, and the nonresponse adjustment procedure.

- Creating a fully imputed data set using composite imputation is the least variable procedure. However, it is also the most biased.
- Imputing values for nonresponding sample units, then weighting the completed subsample is not an unreasonable compromise between the two extremes (Weight or Impute), especially when using warm deck imputation. However, we can only recommend this procedure when the administrative data are reliable.

Table 10: Median Relative Variance (K=2) by Trade Area with Ratio Imputation

Item	Sector	I_C	I_W	IW_C	IW_W	W_T
Payroll	FIRE	48.19	41.03	83.60	64.39	42.90
	Manufacturing	14.52	14.57	28.45	27.16	274.52
	Mining	3.04	4.14	23.61	30.51	40.87
	Retail Trade	414.08	340.41	650.61	526.95	1132.80
	Services	5.14	2.35	17.44	13.44	86.96
	Utilities	-0.21	-0.15	0.44	0.41	12.70
	Wholesale Trade	2.24	2.40	6.61	5.46	34.21
Employment	FIRE	567.06	347.76	792.37	485.30	171.58
	Manufacturing	56.30	40.32	86.66	63.34	463.79
	Mining	8.81	8.81	26.06	37.00	54.93
	Retail Trade	1129.80	924.82	1822.06	1461.54	2415.99
	Services	96.07	43.00	186.45	85.07	287.00
	Utilities	10.40	9.16	18.60	15.57	41.07
	Wholesale Trade	16.13	16.26	28.80	24.89	69.34
Receipts	FIRE	20.65	24.82	37.74	39.31	88.87
	Manufacturing	44.50	47.75	76.48	73.63	339.69
	Mining	2.73	3.32	22.39	26.58	129.00
	Retail Trade	633.92	723.15	1060.40	1104.03	1755.45
	Services	5.76	3.62	15.12	12.85	109.62
	Utilities	5.23	6.49	9.60	10.45	43.33
	Wholesale Trade	17.06	18.26	26.26	26.83	106.49

We had expected large reduction in relative variance using the warm deck parameters over the cold deck parameters. Warm deck imputation is ratio estimation, and the ratio models used for the general statistics items are often strong (see Table 7). However, the differences between corresponding empirical variances for warm and cold deck within adjustment procedure are very small.

4.3. Allocation Results

On average, using the higher subsampling rate increased the empirical variance of each procedure by a factor less than the expected ($3^2/2^2$). The systematic subsampling is providing representative subsamples. There are substantial cost savings with a 1-in-3 subsample of nonrespondents compared to a 1-in-2 subsample, and we can anticipate the corresponding loss in precision.

Although the bias effects of subsampling can be largely controlled through the adjustment cell procedure, the subsampling effects on the empirical variance or MSE is unpredictable (going from 100% follow-up) and can be quite large. We illustrate this unpredictability using "typical" plots of MSE from a FIRE editing cell (the industry in this case) assuming a uniform response mechanisms with p = 0.6.

Figures 8 through 10 graph the MSE for payroll for values of q = 0.2 through 0.8 for the two Impute/Weight methods (warm and cold deck) and for weighting with subsampling rates of K=1, K=2, and K=3, respectively.

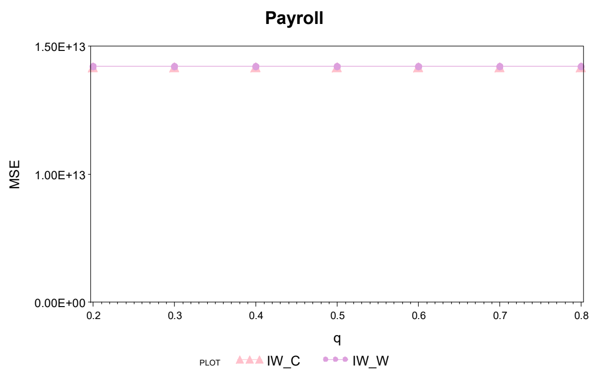

Figure 8: MSE Plot of Payroll Obtained Using Composite Imputation with *p*=0.6 (Uniform Response) for K=1

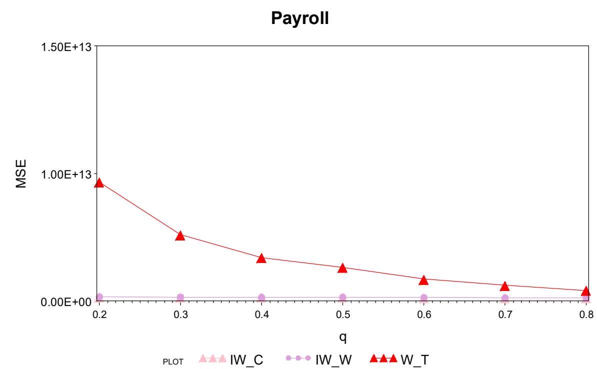

Figure 9: MSE Plot of Payroll Obtained Using Composite Imputation with *p*=0.6 (Uniform Response) for K=2

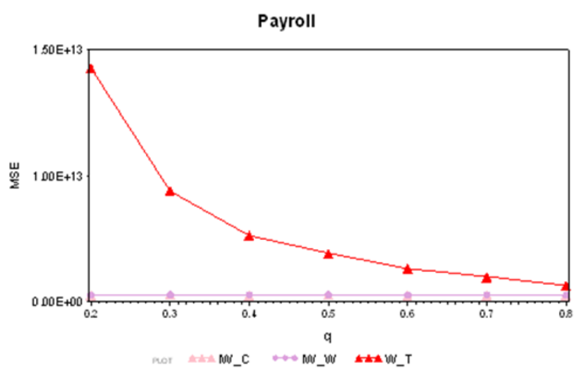

Figure 10: MSE Plot of Payroll Obtained Using Composite Imputation with *p*=0.6 (Uniform Response) for K=3

With payroll, the maximum MSE occurs with full imputation (Impute/Weight in this context creates a complete record for all nonrespondents, then weights each by 1). The subsamples that are treated with the Impute/Weight procedures have nearly equivalent MSE's. The weighted tabulations' MSEs decrease as q increases (the probability of unit nonresponse conversion increases). However, the unit conversion rate (q) needs to be unrealistically high for the three sets of MSE's in the treated subsamples to be comparable.

Figures 11 through 13 present the corresponding MSE graphs for receipts. With receipts, the pattern completely reverses. With no subsampling (K=1) or with weighting (K=2 or K=3), the MSEs are minimized and approximately equal until $q \geq 0.7$ (K=2), and the Impute/Weight procedure with subsamples has much greater error.

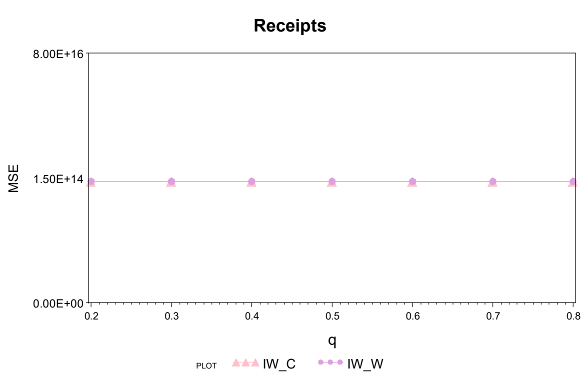

Figure 11: MSE Plot of Receipts Obtained Using Composite Imputation with *p*=0.6 (Uniform Response) for K=1

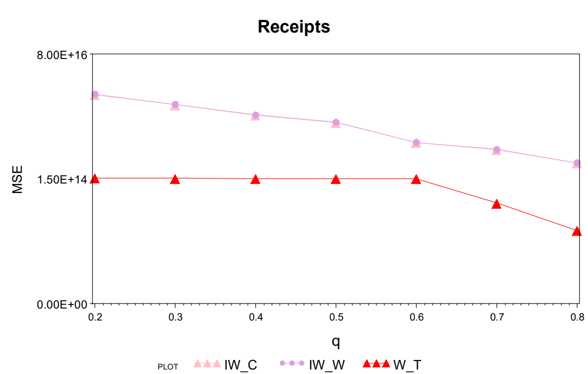

Figure 12: MSE Plot of Receipts Obtained Using Composite Imputation with *p*=0.6 (Uniform Response) for K=2

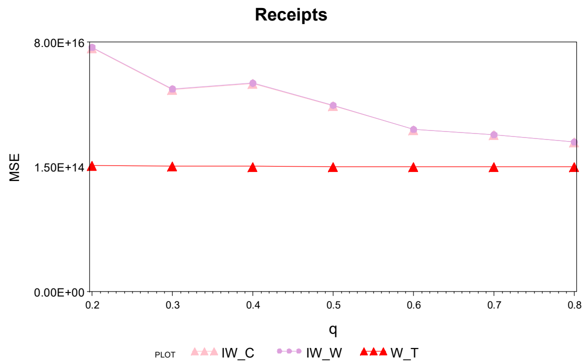

Figure 13: MSE Plot of Receipts Obtained Using Composite Imputation with *p*=0.6 (Uniform Response) for K=3

This demonstrates the importance of examining the effects of subsampling on the tabulations *before* comparing alternative subsampling rates. The jump from *not subsampling* to *subsampling* may behave inconsistently and can go in different directions for different items. Program priorities may come into play. For example, we often found that subsampling did not overly affect payroll tabulations, but did have large effects on receipts. It is also important to note that these patterns can and do vary by trade area.

5. Conclusion

The Economic Census is a large, comprehensive program that covers literally thousands of industries and collects data from millions of establishments. The size of the program and the diversity of the data collected have always made standard processing a challenge. On one hand, from a processing perspective, it would best to use the same methods for editing, imputation, tabulation, and adjustment. On the other hand, each industry is its own population. From a modeling perspective, the optimal adjustment method for a given item in a given industry could be quite different from the optimal method for the same item in a different industry even when both industries come from the same sector. When the sectors are different, it is even less likely that the optimal adjustment methods are the same. For example, consider the different weighted regression models used to develop ratio imputation parameters discussed in Section 3.3. Prior to each census, the regression models are reviewed and re-validated. The manufacturing and mining sectors always have an error model that is consistently different from the other trade areas and the resultant best linear unbiased estimator used for ratio parameter development in these trade areas is different. Here, the best practice is to develop ratio imputation parameters using the model appropriate for the data, and that is what is done. As a compromise, the census processing uses generalized software. However, each trade area develops its own processing script, and editing and imputation parameters are developed at an industry subdomain level.

At the beginning of this project, we hoped to find a single "best" adjustment method or even a single "best" sampling method. Instead, the results differed by trade area. To understand the problem, we needed to consider each trade separately then examine the patterns from the collective set of results. This journey to understand these patterns was slow, and we took several steps forwards and backwards. Exploratory graphical analyses at the editing cell level proved misleading, as extreme differences in scale obscured common (or distinct patterns). Subtle differences in variance caused by differing sampling rates within the same method appeared to be larger than in reality. Fortunately, using the side-by-side comparisons within adjustment method, sampling rate, item, response mechanism, and editing cell forced us to look at the data at a fine level and to see larger patterns by trade area. As soon as we found ourselves in a metaphorical forest – no longer lost in the trees – the across-the-board differences were more obvious, and we could begin to explore possible causes.

Each of us had a preconceived idea of the "best adjustment" method. One of us preferred adjustment cell weighting, because of its ease of implementation and fairly robust bias correction properties. The variance increases due to the combined impact of subsampling and nonresponse were unacceptable for any subsampling rate that would achieve a tangible cost savings (i.e. at least a one-in-two subsample). This agrees with the research conducted by Whitehead et al. (2013). One of us preferred mass imputation because it allows for easy computation of domain estimates. Unfortunately, the bias increases due to incomparability of auxiliary data or weaker imputation models were likewise unacceptable. Verret (2013) had similar findings for mass imputation with Statistics Canada's National Household Survey.

The primary purpose of nonresponse adjustment is to reduce the effect(s) of nonresponse bias on the estimates. Many imputation studies use one replicate (the original tabulated data set) to compare alternative adjustment methods and assess the competing methods by measuring deviations between imputed and original tabulated data (Nordholt, 1998; Thompson and Williams, 2003). There are several advantages to using available historical data such analyses. It allows the evaluator to estimate the amount of time actually required by the edit process. It allows the program managers to "see" the effects of the alternative methods on the published estimates. Most important, it uses a "gold standard" accepted by the customers (subject-matter experts generally have great confidence in their publication data). However, there are just as many disadvantages. First, this approach assumes that the tabulated (edited and imputed) data are entirely correct. Second, it is difficult to examine relationships between the adjustment method and specific conditions (for example, composite versus ratio imputation or uniform verses non-ignorable response mechanism). Conclusions are highly dependent on a single sample and – more important – a single set of nonrespondents. Cross-validation is a possible way of addressing this concern, but it greatly reduces the size of both the model-building and validation sets (if current data are used) and introduces time-period comparability issues with census data because of the five-year collection cycle. Using simulated data (modeled on real data) eliminates the last concern by using a large number of replicates and by varying the response mechanism to assess sensitivity.

Ultimately, one could argue that there is no single optimal adjustment procedure for a program with as diverse data as the Economic Census. That said, selecting a representative probability subsample of nonrespondents, focusing on maximizing their response, completing the records via warm-deck imputation for the unconverted nonrespondents, and weighting the completed subsample (otherwise known as "Impute/Weight") is a good middle ground. There are several advantages to this approach. The imputation models profit from the representative subsample, which should contain units of varying sizes in direct contrast to the historic cold deck parameters. In theory, with the smaller set of nonrespondents, the subject matter experts could devote more time to obtaining valid auxiliary data for substitution, further increasing precision. The subsampling does increase the variance even under 100% response, but the advantages of more precise imputation might offset these increases if the models are strong. Moreover, the warm deck parameters are used throughout the census processing to impute values for missing and invalid items, and so the quality benefits of developing warm deck parameters from a representative sample are not limited to unit nonresponse adjustment. That said, using this compromise adjustment procedure is not necessarily uniformly the best for all item in all industries: indeed, cold deck mass imputation or adjustment cell weighting has better statistical properties in some cases.

If subsampling is implemented, then Census Bureau Standards require that measures of reliability be provided along with the tabulation. If only sampling variance is accounted for in the reliability measures, many of the domain estimates – or the trailer data or product line items – may not be sufficiently reliable. Adding in the imputation variance will exacerbate the situation. The bias issues of the imputed estimates are not new but have not been highlighted. A logical next step is to compare alterative imputation methods in terms of their contribution to the variance.

Acknowledgements

The authors thank Xijian Liu, Eddie Salyers, and Yukiko Tomabechi Ellis for their valuable contributions to this research project.

References

Federal Register Notice (2006). OMB Standards and Guidelines for Statistical Surveys.

Hoaglin, D.C., Mosteller, F., and Tukey, J.W. (eds.) (1983). Understanding Robust and Exploratory Data Analysis. New York: Wiley.

Kalton, G. and Flores-Cervantes, I. (2003). Weighting Methods. *Journal of Official Statistics*, **19**, pp. 81–97.

Lineback, J.F., Oliver, B., Willimack, D. (2012). Developing Response Metrics for the Economic Census. *Proceedings of the 2012 Federal Committee on Statistical Methodology Research Conference.*

Nordholt, E.S. (1998). *Imputation: Methods, Simulation Experiments and Practical Examples. International Statistical Review,* ***66****, pp. 157-180.*

Thompson, K.J., and Oliver, B. (2012). Response Rates in Business Surveys: Going Beyond the Usual Performance Measure. *Journal of Official Statistics*, **28**, 0pp. 221-237.

Thompson, K.J. and Sigman, R.S. (1996). Evaluation of Statistical Methods for Developing Ratio Edit Module Parameters. *Internal Census Bureau Report (document ESM-9601) available upon request.*

Thompson, K.J. and Williams, Q. (2003). Developing Imputation Models for the Services Sectors Portion of the Economic Census. *Proceedings of the 2003 Federal Committee on Statistical Methodology Research Conference.*

Verret, F. (2013) The Estimation Methodology of the 2011 National Household Survey. *Proceedings of the Section on Survey Research Methods,* American Statistical Association.

Wagner, D. (2000). Economic Census General Editing – Plain Vanilla. *Proceedings of the 2nd International Conference on Establishment Surveys*.

Whitehead, D., Oliver, B., Thompson, K.J. (2013). Determining Allocation Requirements for Subsampling Nonrespondents from the Annual Survey of Manufactures (ASM). *Proceedings of the 2013 Federal Committee on Statistical Methodology Research Conference*