Developing Variance Estimates for Products in the Economic Census

Jeremy Knutson, U.S. Census Bureau[1]
Matthew Thompson, U.S. Census Bureau
Katherine J. Thompson, U.S. Census Bureau

Abstract
The Economic Census collects data on the revenue obtained from products (product data) from all sampled units. In the 2017 Economic Census, missing product data will be imputed using hot deck imputation, and variance estimates for product data will be published for the first time. Product data pose unique challenges. Often sampled establishments elect not to provide any values (complete nonresponse) and many products are rarely reported. Consequently, the variance estimator must account for sampling variance, post-stratification, and imputation variance. Our recommended multiple imputation variance estimator combines the finite population Bayesian bootstrap (FPBB) with the approximate Bayesian bootstrap (ABB). Using a simulation study, we evaluate the performance of this variance estimator with two hot deck imputation methods (nearest-neighbor and random), seeking the appropriate number of implicates at the FPBB and ABB stages, with the simultaneous objectives of producing variance estimates with good statistical properties while maintaining programmatic efficiency. A secondary consideration is whether the imputation method has an impact on the quality of the variance estimates.

Key Words: Economic Census, finite population Bayesian bootstrap, approximate Bayesian bootstrap, hot deck imputation, products

1. Introduction

The Economic Census is the U.S. Government's official five-year measure of American business and the economy. The term “Economic Census” is a bit of a misnomer; the majority of sectors sample the small single-unit (SU) establishments and survey all of multi-unit (MU) establishments[2]. The Economic Census collects a core set of data items from each establishment called general statistics items: examples include total receipts or value of shipments (“receipts”), annual payroll, and number of employees in the first quarter. In addition, the Economic Census collects data on the revenue obtained from

[1] Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.
[2] A single-unit (SU) establishment owns or operates a business at a single physical location, whereas multi-unit (MU) companies comprise two or more establishments that are owned or operated by the same company.

product sales (hereafter referred to as "products"). With the exception of the construction sector, all sectors construct a *complete universe of general statistics values* by using administrative data in place of respondent data for unsampled units. However, product data are collected from only the sampled establishments. In most sectors, weighted sample estimates are further calibrated to the industry totals for receipts.

Product data collection is challenging. The Economic Census collects information on over 8,000 different products defined by the North American Product Classification System (NAPCS); see https://www.census.gov/eos/www/napcs/more.html. However, many products are rarely reported. Establishments can report values from a long list of potential products in a given industry (some lists span more than 50 potential products), and consequently, many establishments choose not to report any product data (complete product nonresponse). These lists vary by industry and can in fact differ within the broader sector. Furthermore, product descriptions are quite detailed and some products are mutually exclusive. In addition, all reported product values within a given establishment are expected to sum to the total receipts value reported earlier in the questionnaire. Finally, legitimate zero values are expected for the majority of eligible products in an industry, at both the individual establishment and total industry level.

In the 2017 Economic Census, missing product data will be imputed using hot deck imputation (Thompson and Liu 2015; Knutson and Martin 2015), and variance estimates for product totals will be published for the first time. Depending on the industry, random hot deck or nearest neighbor hot deck imputation will be implemented (Tolliver and Bechtel 2015; Bechtel, Morris, and Thompson 2015). The variance estimator must account for sampling variance, calibration weighting, and imputation variance.

From a variance estimation perspective, most of the challenge lies with the poor predictors and high expected zero rates for many products, although discounting the high nonresponse rate would be very optimistic as the possibility of a low donor-to-recipient ratio for hot deck is quite high. The variance estimation challenges that are a direct consequence of the sample design need to be addressed. In the 2012 Economic Census, sample design varied by sector. It is not unreasonable to expect to find a variance estimation method that produces estimates with good statistical properties in terms of bias and stability for the well-reported products. It may be unreasonable to hold similar hopes for the less frequently reported products.

Earlier research broke down the challenge of producing product variance estimates into two constituent parts: sampling variance and variance due to imputation. Several replication methods were tested under the assumption of complete response with calibration in order to evaluate their performance in estimating sampling variance. Ultimately, the finite population Bayesian bootstrap (FPBB) was selected as a means of estimating this component of variance (Thompson, Thompson, Kurec 2016). In addition, the statistical properties of the nonresponse variance (imputation) component were studied under a design-based (model-assisted) framework and under a model-based framework, with the approximate Bayesian bootstrap (ABB) showing the most promise (Thompson and Thompson 2016).

In this paper, we combine the results from both studies in a simulation "cook-off" that integrates the FPBB and ABB using the variance estimator proposed in Zhou et al (2012). We use the simulation study in order to determine the number of implicates needed for each component (FPBB and ABB) and to assess the utility and reliability of the fully-assembled variance estimator. Note that the production estimates are singly imputed, whereas the variance estimate computations use a multiple imputation procedure. This hybrid approach is not justified theoretically, especially since the bias of the single and multiply imputed estimates can differ and is often not negligible. However, this approach is useful in a production setting, as it allows the analysts to review the tabulated product estimates on a flow basis during the data review period.

Section 2 provides some background on the 2012 Economic Census sample designs by sector, as well as some details about hot deck imputation. Section 3 describes the variance estimation methodology. Section discusses the simulation study design and presents the results of the variance estimation evaluation. Section 5 concludes the paper with some final thoughts about this research project, and its future impact on the Economic Census data.

2. Product Data Estimation in the Economic Census

2.1 Sample Design and Calibration

The Economic Census is a sample survey in all sectors except for Wholesale trade. A small proportion of the frame is sampled with probability less than 1 (noncertainty). All multi-unit establishments are included with certainty, as are the largest single-unit establishments within a stratum (representing 80% of the total receipts), so that only a small proportion of the SU establishments are sampled. For the 2017 Economic Census, strata are defined by 6- or 8-digit NAICS industry and state depending on the sector. The remaining single-unit establishments are a systematic sample. In general, the sample design is not intended or well-suited for direct variance estimation: sample sizes within strata are often very small, consisting of one or two establishments.

Hot deck imputation is used to account for product nonresponse from sampled units. These procedures are described in Section 2.2. After imputation, the sample-weighted product estimates are calibrated by controlling the sample-based estimates of receipts (aggregated across the products) to the census values of receipts within industry-by-state. This final adjustment accounts for product nonresponse and post-stratification.

2.2 Hot Deck Imputation from the Donor Distribution

Hot deck imputation selects a 'similar' unit from a donor pool (indicated by superscript *d*) and uses its data to impute a group of missing values, thus preserving existing relationships between items. For product data imputation, we select a single donor for each recipient and impute values for all products as $\tilde{y}_{ijk} = x_{ik}(y_{ijl}^{d}/x_{il}^{d})$, where y_{ijk} is the value of product j in industry i for *recipient* establishment k ($k \neq l$) and x_{ik} is the unit's value of total receipts which is always available. The imputation ratios are obtained from the donor record's corresponding product and total receipts value, thus preserving the establishment level multivariate distributions and ensuring additivity.

In random hot deck imputation, the donor record is randomly selected, usually with replacement (Brick and Kalton 1996). This particular application of hot deck is optimal when both the response propensities and the expected value(s) of the variable(s) of interest are homogenous within an imputation cell (Andridge and Little 2010). Due to the random selection, this method "preserves the distributional properties of the imputed dataset; that is, the distribution function for imputed data within a cell differs from the distribution function for the respondents in the cell only because of the randomness of imputation" (Kim and Fuller 2004).

Nearest neighbor hot deck imputation uses a distance measure to select the donor. The distance measure can be a function of one or more auxiliary variables and can have several functional forms. Distance measures based on a variable (or variables) available for all units are calculated for all donors compared to all recipients. The donor that is closest to a particular recipient within an imputation cell is selected, with a donor randomly selected in the event of a tie. This hot deck method is optimal when the variable(s) used for the distance measure is highly correlated with the variable(s) of interest and the response propensities are homogenous within an imputation cell. Nearest neighbor imputation is deterministic imputation and does not have the same asymptotic properties as random imputation. That said, Chen and Shao (2001) describe many advantages of the nearest neighbor imputes including: (1) reasonable values with little or no chance of "nonsensical" imputes; (2) asymptotically unbiased and consistent estimators of population means and quantiles; and (3) employment of a robust nonparametric model that relates outcome (to be predicted) to matching variables. Nearest neighbor hot deck imputation is especially attractive for business surveys with skewed populations, as it guards against selecting a donor record with a very different variable distribution assuming that reporting patterns are correlated with size of business. We use the absolute difference between donor receipts value and recipient receipts value. In the event of a donor tie, we randomly choose a donor.

Cell collapsing is necessary when there are insufficient donors in an imputation cell. For this research, our cell minimum was five. This differs from the 2017 Economic Census practice, which only collapses when there are no donors in an imputation cell. We used the same imputation cells as the 2017 Economic Census. The finest level was industry /state code/ unit type, where the industry code could incorporate a further sub-classification by legal form of operation (LFO) or type of operation, depending on the sector. If fewer than five donors are available in a cell, we dropped the unit type classification, then the region. Collapsing occurs very rarely in our applications.

3. Variance Estimation Methodology

3.1 Finite Population Bayesian Bootstrap (FPBB)

The Finite Population Bayesian Bootstrap (FPBB) described in Zhou et al (2012) is a non-parametric multiple imputation method that accounts for complex sampling procedures and post-stratification. With the FPBB, the idea is to expand the sample of size n into several FPBB populations, or implicates, each of size N, where N is the original population size.

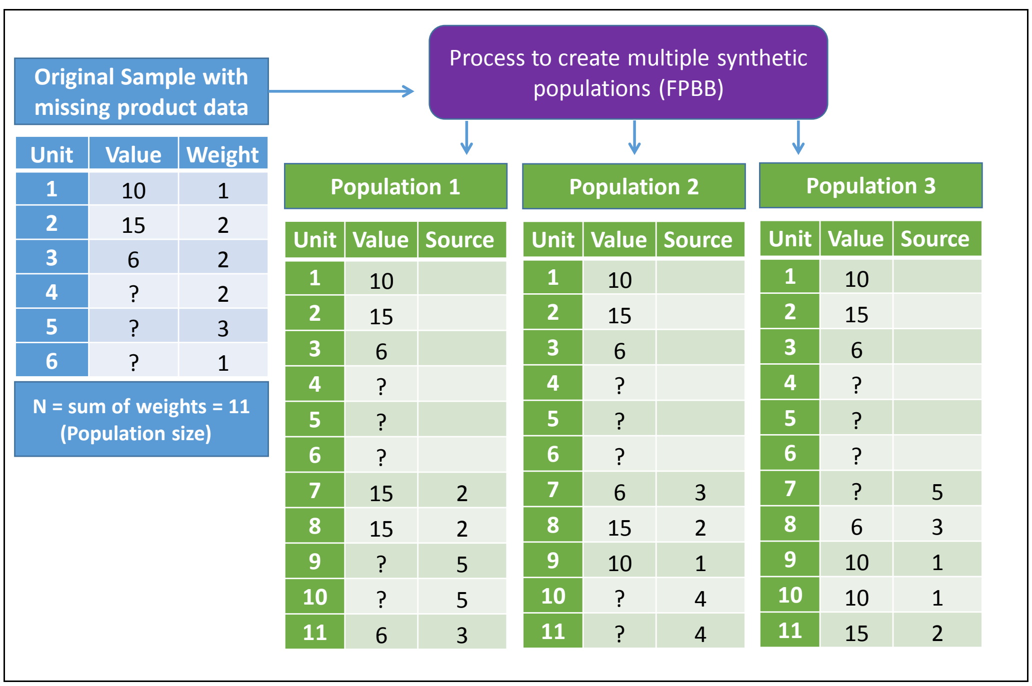
These FPBB implicates are created by drawing $(N_h - n_h)$ units from stratum h from the original sample with probability for the k^{th} selection,

$$p_{h,k} = \frac{w_i - 1 + l_{i,k-1}\dfrac{(N_h - n_h)}{n_h}}{N_h - n_h + \dfrac{(k_h - 1)(N_h - n_h)}{n_h}}$$

where w_i is the post-stratified sampling weight of unit i, $l_{i,k-1}$ is the number of times unit i has been selected up to the $(k-1)^{\text{th}}$ selection, and k_h is the number of selections that have been made.

The $(N_h - n_h)$ resampled units are added to the original sample to complete the FPBB implicate. As described by Zhou et al (2012), this is an application of a Pólya sample designed to "restore the existing complex survey sample back to some SRS-type/self-weighting data structure." This process, which Zhou refers to as "uncomplexing" the sample, is repeated several times to create a total of B implicates. Determining a sufficiently large B value is discussed in section 4.2.2.

Figure 1 illustrates the FPBB process with $B=3$ implicates for an unequal probability sample of size $n = 6$, sampled from a population of size $N = 11$. A value of "?" indicates a nonrespondent in the sample that is likewise included as a nonrespondent in the expanded population implicate.



Unit	Value	Weight
1	10	1
2	15	2
3	6	2
4	?	2
5	?	3
6	?	1

Population 1

Unit	Value	Source
1	10	
2	15	
3	6	
4	?	
5	?	
6	?	
7	15	2
8	15	2
9	?	5
10	?	5
11	6	3

Population 2

Unit	Value	Source
1	10	
2	15	
3	6	
4	?	
5	?	
6	?	
7	6	3
8	15	2
9	10	1
10	?	4
11	?	4

Population 3

Unit	Value	Source
1	10	
2	15	
3	6	
4	?	
5	?	
6	?	
7	?	5
8	6	3
9	10	1
10	10	1
11	15	2

Figure 1. Illustration of creating three pseudo-populations from a sample by FPBB

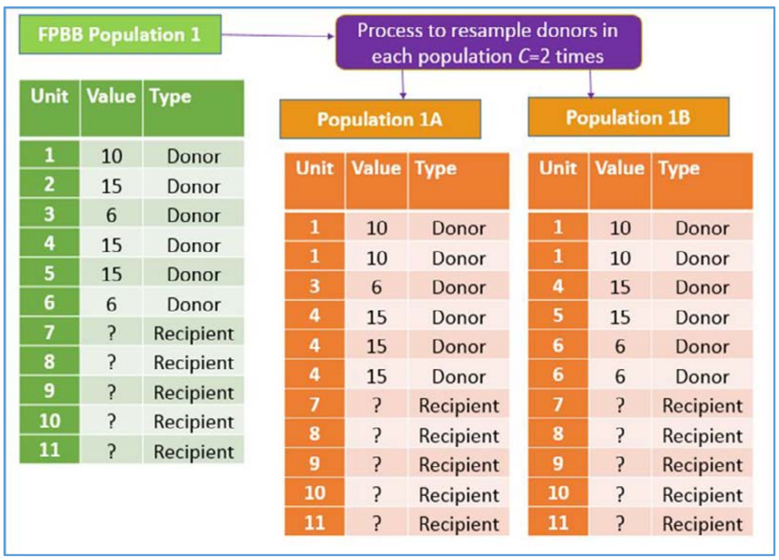
Recall that the Economic Census product estimates must account for the calibration error as well as the sampling error. We account for this by using the post-stratified sampling weight instead of the design weight in the Pólya sampling procedure, so that the expanded population sizes may differ from the original sampling frame population sizes.

3.2 Approximate Bayesian Bootstrap (ABB)

The next step is to incorporate product nonresponse and estimate the nonresponse variance. To do this, we employ the Approximate Bayesian Bootstrap (ABB) within each FPBB implicate. The ABB is a natural and straightforward way to implement multiple imputation for the hot deck methodology. Rubin and Schenker (1986) and Rubin (1987) propose the ABB as a tool for introducing appropriate variability into a multiple imputation procedure. ABB is a non-Bayesian method that approximates a Bayesian procedure and adjusts for the uncertainty in the distribution parameters resulting in a proper imputation procedure. ABB involves:

1. Drawing a simple random sample (SRS) of respondents with replacement (see Figure 2), and
2. Imputing values for missing data using the sample of respondents drawn in the first step as the implicate donors (See Figure 3).

Each round of the ABB procedure results in one complete dataset. This procedure is then repeated C times to obtain multiple imputed datasets. Ultimately, each of the B FPBB implicates will have C ABB implicates.



FPBB Population 1

Unit	Value	Type
1	10	Donor
2	15	Donor
3	6	Donor
4	15	Donor
5	15	Donor
6	6	Donor
7	?	Recipient
8	?	Recipient
9	?	Recipient
10	?	Recipient
11	?	Recipient

Population 1A

Unit	Value	Type
1	10	Donor
1	10	Donor
3	6	Donor
4	15	Donor
4	15	Donor
4	15	Donor
7	?	Recipient
8	?	Recipient
9	?	Recipient
10	?	Recipient
11	?	Recipient

Population 1B

Unit	Value	Type
1	10	Donor
1	10	Donor
4	15	Donor
5	15	Donor
6	6	Donor
6	6	Donor
7	?	Recipient
8	?	Recipient
9	?	Recipient
10	?	Recipient
11	?	Recipient

Figure 2: Illustration of the first step of ABB resampling within a given FPBB implicate

Use hot deck imputation with resampled donors

Population 1A

Unit	Imputed Value	Type
1	10	Donor
1	10	Donor
3	6	Donor
4	15	Donor
4	15	Donor
4	15	Donor
7	10	Recipient
8	6	Recipient
9	15	Recipient
10	15	Recipient
11	6	Recipient

Population 1B

Unit	Imputed Value	Type
1	10	Donor
1	10	Donor
4	15	Donor
5	15	Donor
6	6	Donor
6	6	Donor
7	6	Recipient
8	15	Recipient
9	10	Recipient
10	6	Recipient
11	6	Recipient

Figure 3: Multiply Imputed ABB Implicates for Population 1

There is some discussion on how to appropriately incorporate unequal probability sampling into the ABB application (Andridge and Little 2009). However, the resampling procedure used by the FPBB already accounts for unequal probability sampling, so that the usage of the unrestricted SRS is appropriate.

3.3 Variance Estimation

Once the FPBB implicates have been created, nonresponse induced, and the ABB employed to multiply impute, then variance estimates can be calculated in the following manner. First, calculate a within-FPBB implicate variance estimate as

$$\hat{V}_{imp} = \left(1 + \frac{1}{C}\right)\left(\frac{1}{C-1}\right)\sum_{b=1}^{B}\sum_{c=1}^{C}\left[TOT_{b,c} - FPBBAVG_b\right]^2 \quad (3.1)$$

where $TOT_{b,c}$ is the estimate for ABB replicate c within implicate b and $FPBBAVG_b$ is the average of these C estimates within FPBB implicate b. Compute a between-implicate variance as

$$\hat{V}_{samp} = \left(1 + \frac{1}{B}\right)\left(\frac{1}{B-1}\right)\sum_{b=1}^{B}[FPBBAVG_b - AVG]^2 \quad (3.2)$$

where $AVG = \frac{1}{B}\sum_{b=1}^{B} FPBBAVG_b$. Figure 4 presents a holistic picture of the resampling and averaging processes for B=3 and C=2.

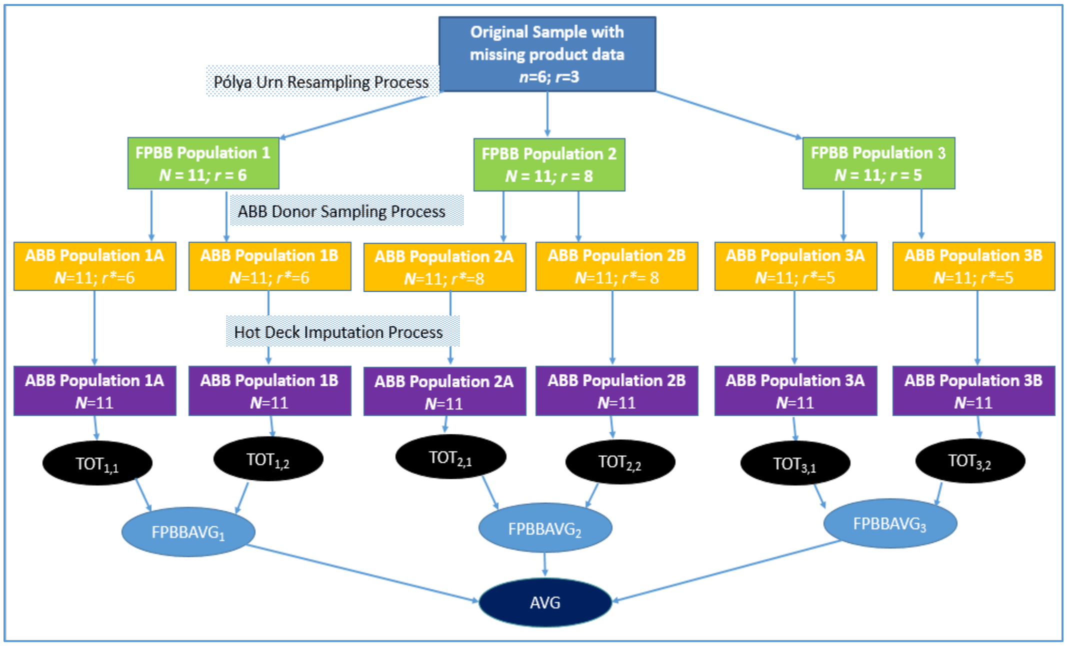


Figure 4: Illustration of the Complete FPBB and ABB Combined Procedure for B = 3 and C = 2

Finally, use the revised multiple imputation (MI) combining rule (Zhou, Raghunathan, and Elliot, M. 2012) to calculate the total variance as

$$\hat{V}_{final} = \hat{V}_{samp} + \frac{1}{B}\hat{V}_{imp} \quad (3.3)$$

4. Study Data and Simulation Approach

4.1 Simulation Design

We used a simulation approach to assess the performance of the combined FPBB/ABB variance estimator with varying numbers of *B* and *C* over repeated samples. This simulation study used the modeled population data developed for the initial evaluations of variance estimation methods (Thompson and Thompson 2016; Thompson, Thompson, and Kurec 2016). The 30 studied industries were provided by subject matter and classification experts, covering 12 North American Industry Classification (NAICS) sectors. These industries are not meant to be representative of the entire Economic Census. The selected industries' eligible products were expected to remain consistent under the introduction of NAPCS.

The populations were modeled from 2012 Economic Census microdata, which was subject to complete product nonresponse and contained only sampled units. To create the finite industry populations, we filled in missing product values using nearest neighbor hot deck imputation for the sampled cases using the donor records identified by the subject matter experts. Then, we used the SIMDAT algorithm to create completed records for the unsampled SU establishments in each industry (Thompson 2000). This nonparametric “nearest neighbors” simulation technique creates simulated data with the same correlation structure as the sample survey (training) data and similar quantile values. To implement the algorithm, we had to limit the number of simulated products in each industry to the four best-reported products in each industry in terms of number of establishments that reported the product (Products 1 through 4) plus an “all other product values” item containing the balance of the difference between the establishment total receipts and summed top four products (Product 5). This final “catch all” product does not resemble the collected data and is excluded from our analyses.

After establishing the complete population, we independently selected 5,000 stratified SRS-WOR samples in each industry and then randomly induced non-response, using the sampling parameters and observed product response rates from the 2012 Economic Census. We used single imputation twice -- once for random hot deck and once for nearest neighbor hot deck -- to obtain two sets of singly-imputed estimates per sample. We obtained MSE values (truth) from the 5,000 samples as

$$MSE\left(\hat{Y}_{ijg}^{m}\right) = \left[\sum_{s=1}^{5,000}\left(\hat{Y}_{ijg}^{ms} - Y_{ijg}\right)^2 \Big/ 5,000\right]$$

where *m* indexes the hot deck imputation method, *i* indexes the industry, *j* indexes the product, *g* indexes the state, and *s* indexes the sample.

For 1,000 of the 5,000 samples, we applied our variance estimation methodology. For each of these 1,000 samples, the sample was expanded 20 times in order to create 20 distinct FPBB populations. Then for each of the FPBB populations, donors were resampled 20 times to create 20 different donor pools from which to impute the FPBB population, resulting in a total of 400 implicate populations per sample. Processing considerations determined the upper bound on the allowable number of implicates. There were storage space concerns due to the size of the FPBB populations, and the hot deck imputation procedures can be slow in real time.

Each of these populations was then imputed using each of the production imputation methods available for use for the 2017 Economic Census – nearest neighbor hot deck and random hot deck. From these 400 imputed populations, variance estimates were then created using different subsets of populations for different quantities of *B* and *C* (i.e. for *B* = 5 and *C* = 5, only the first five FPBB populations and the first five ABB donor pools were used in the calculation of the variance estimate).

We used these 1,000 samples to assess the relative bias and coverage of each considered combination of *B* and *C* as

Relative Bias $$RBV\left(v^{bc}\left(\hat{Y}_{ijg}^{bcm} \right)\right) = \left[\sum_{s=1}^{1,000} v^{s}\left(\hat{Y}_{ijg}^{bcsm}\right)/1,000 \Big/ MSE\left(\hat{Y}_{ijg}^{m}\right)\right] - 1$$

Coverage Percentage of 90% confidence intervals constructed with v^{bcs} and $t_{b-1,0.95}$ containing the true population value of the studied product j in industry i. The critical values for the confidence intervals were 2.132 (B = 5), 1.833 (B = 10), 1.761 (B = 15), and 1.729 (B = 20).

where v^{bcs} indexes the variance estimate obtained using b FPBB populations and c ABB-selected donor pools in sample s.

4.2 Analysis

The table below presents the combinations of FPBB and ABB implicates used in the analysis. Our objective was to find a variance estimate with low bias and nearly nominal coverage for the majority of studied products. If all measures were comparable, then we selected the variation that used the smaller number of ABB implicates due to processing considerations.

Table 1: Total number of implicates for each studied FPBB(B)/ABB(C) Combination

	ABB (*C*)			
FPBB (*B*)	***C* = 5**	***C* = 10**	***C* = 15**	***C* = 20**
***B* = 5**	25	50	75	100
***B* = 10**	50	100	150	200
***B* = 15**	75	150	225	300
***B* = 20**	100	200	300	400

Initially, we planned to conduct our evaluation on the industry-by-state level estimates. However, the sample sizes in these categories were often very small (especially for rare products), and the collective sets of variance estimates were too noisy to detect patterns. Instead, we use the industry level estimates in the following analyses.

4.2.1 Choosing minimum number of ABB replicates

Hot deck imputation is the most time and resource-consuming process in the variance estimation procedure. Consequently, we wanted to select the minimum number of ABB replicates that yielded stable imputed estimates within each FPBB population. Within each FPBB population, we computed the median imputation variance estimates (the $\hat{V}_{imp}$ defined by (3.1)) over the 1000 samples for C = 5, 10, 15, 20. The convergence of this median variance with increasing C was considered as the criteria for choosing a minimum recommended value for C.

Figure 5 summarizes results for all thirty studied industries. Each chart contains results for the four studied products with five FPBB populations (B = 5), in the form of three variance estimate ratios. Appendices 1 and 2 provide the complete set of results for all industries and number of studied FPBB implicates. Although the pattern is consistent throughout the industries, the detail is difficult to see. To illustrate the pattern, Figure 6 presents the ratios for a single manufacturing sector industry. When the variance estimate approaches convergence, this ratio will approach 1. Figures 5 and 6 show a consistent variance estimation pattern, regardless of products: a very large estimate of $\hat{V}_{imp}$with five ABB implicates (C=5), followed by a steep drop in the $\hat{V}_{imp}$ with additional five implicates (C = 10), followed by another increase in the magnitude of the variance estimate as additional implicates are added. Notice that these ratios are approaching 1 very slowly, but they do appear to be stabilizing.

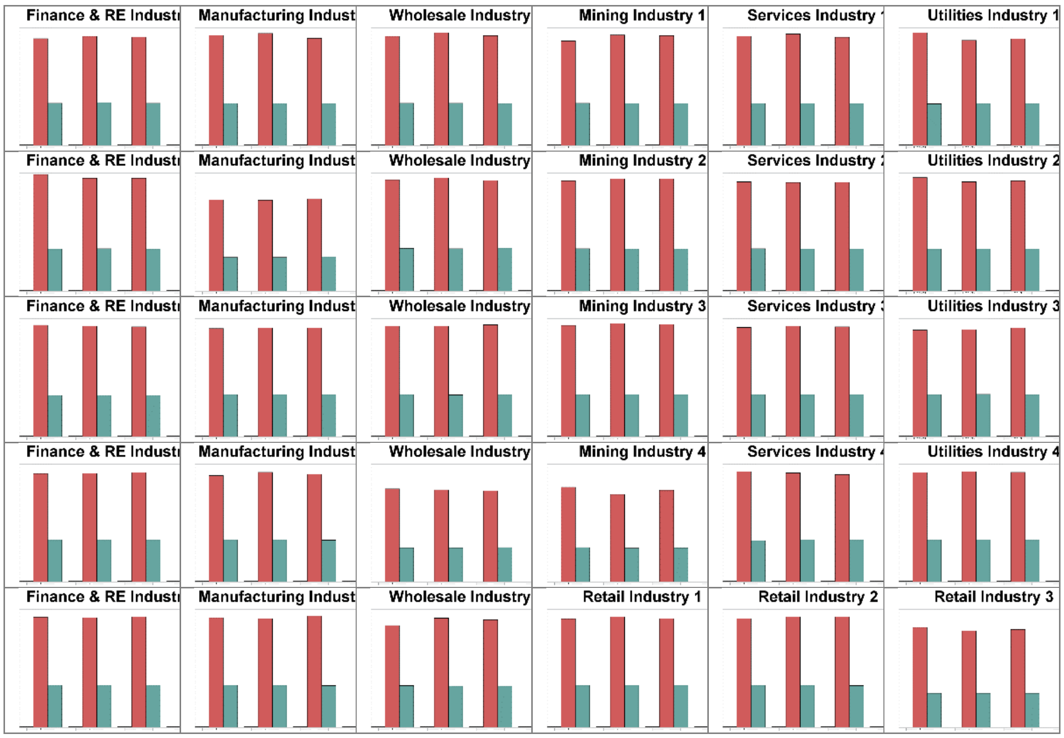


Figure 5: Median Ratio of Variances Estimates Imputed with Nearest Neighbor Hot Deck for B = 5, all industries. Each group of three bars within industry show one of the median variance ratios for 1 of the 4 products

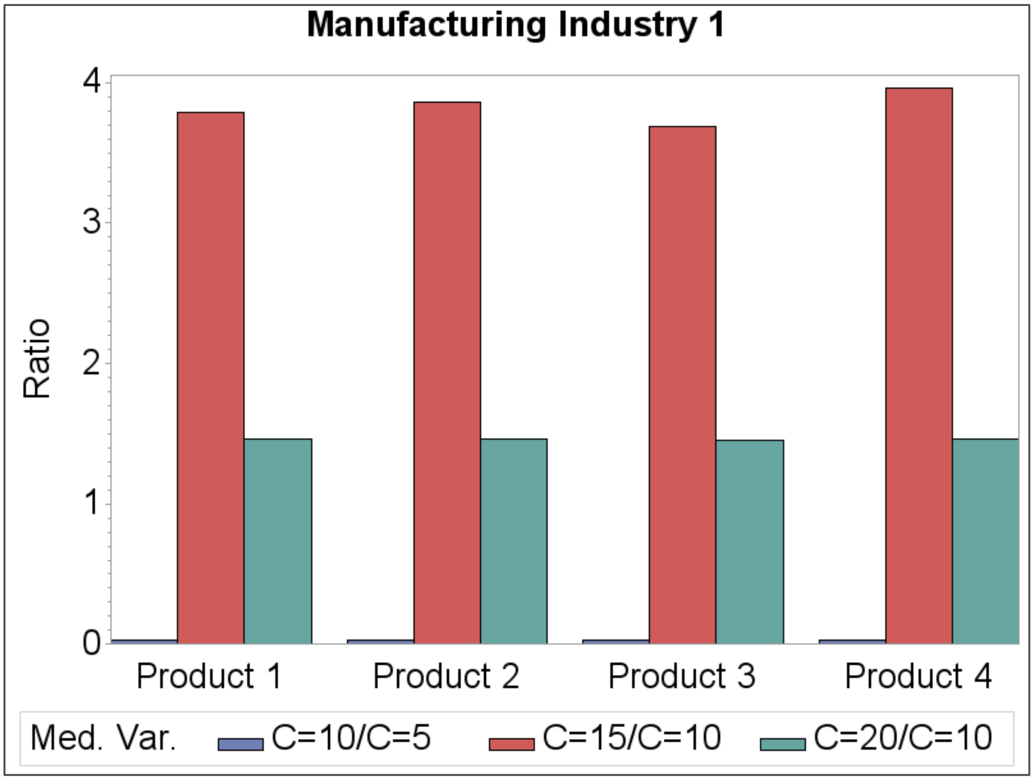


Figure 6: Median Ratio of Within-FPBB Variances Estimates from a Manufacturing Industry Imputed With Nearest Neighbor Hot Deck for B = 5

With median variance$_{C=20}$/ median variance$_{C=15}$ being a little less than 1.5 for all industries, we treat C = 15 as the absolute minimum number of ABB replicates that could be used in our application. However, we decided to use C = 20 since the ratio of within-FPBB variance estimates did not converge to near unity. Moreover, Thompson and Thompson (2016) found that the ABB variance estimates for both random and nearest neighbor hot deck were generally biased with these populations (sampling was not considered), but the bias was reduced when using C=20 over C=10. Ideally, we would have preferred to consider more ABB implicates, given this lack of convergence. However, processing considerations precluded this option.

4.2.2 Choosing minimum number of FPBB implicates

Having determined the number of ABB implicates for our proposed production implementation (C = 20), we turned to the FPBB portion. Even at the industry level, finding a clear pattern for the differing number of FPBB implicates proved to be difficult. First, there are many rarely reported products – in several industries even the second most frequently reported product was reported by only a small percentage of units (less than 30%). Furthermore, for many industries and products, both hot deck imputed estimates are biased. Thus, we restricted our attention to estimates with percent non-zero reported greater than or equal to 30% and relative bias less than or equal to 0.20.

For this analysis, we studied the relative bias of the variance estimator and 90% confidence interval coverage. Appendices 3 and 4 present the relative biases of the variance estimates (all with C = 20) for nearest neighbor and random imputed estimates, respectively. Appendices 5 and 6 present the corresponding coverage rates.

Given C = 20, the magnitude of the relative bias of the variance estimates is minimized with B = 5, although the variance estimates are rarely unbiased. It seems counterintuitive that the most accurate variance estimates would be obtained using the fewest amount of FPBB implicates. Recall that the final variance estimator (3.3) is the sum of two variance components: $\hat{V}_{samp}$ and $\hat{V}_{imp}$. This second component is a constant value in all computations for a given product and industry and is likely an underestimate. The first component scales the squared difference between the individual FPBB estimates ($FPBBAVG_b$) and the across-FPBB average (AVG). Let $SS_{samp(B)} = \left(\frac{B+1}{B}\right)(B-1)\hat{V}_{samp} = \sum_{b=1}^{B}[FPBBAVG_b - AVG]^2$. Within each industry by product, we computed the median industry-level ratios of $SS_{samp(B)}$ provided in Table 2 below. These ratios were consistent to within the third decimal point for all industries and products, so that Table 2 provides the median value across all industries and samples. This allows us to compare the effect of adding additional FPBB implicates into the between-Population (sampling) variance estimates.

Table 2: Median $SS_{samp(B)}$ Ratios by Product Across all Industries.

Product	$SS_{samp(10)}$ $/SS_{samp(5)}$	$SS_{samp(10)}/SS_{samp(15)}$	$SS_{samp(15)}$ $/SS_{samp(10)}$	$SS_{samp(20)}$ $/SS_{samp(5)}$
1	0.056	6.768	1.996	0.809
2	0.053	6.914	1.999	0.810
3	0.056	6.760	2.003	0.810
4	0.054	6.687	1.999	0.810

This shows a very similar pattern as shown with $\hat{V}_{imp}$. The sums of squares for B = 5 are approximately 20 times larger than those for B = 10. As the number of FPBB implicates increases, the sums of squares likewise increase, so that the values for B = 20 are approaching those of B = 5. However, these sums of squares are rescaled by $\left(1+\frac{1}{B}\right)\left(\frac{1}{B-1}\right)$ or values of 0.30 (B=*5),* 0.122 (B=10), 0.076 (B=15), and 0.055 (B=*20).* In short, as more implicates are added, the differences between the multiply-imputed FPBB estimates tend to approach a single value, about 20-percent smaller than the value obtained with B = 5. However, as $B \rightarrow \infty$, the rescaling factor converges to 0.05, so that $\hat{V}_{samp}(5) > \hat{V}_{samp}(B > 20)$. Thompson, Thompson, and Kurec (2016) reported that the FPBB estimates of $\hat{V}_{samp}$ (without imputation variance) tended to underestimate the true variance. Thus, using the smaller number of FPBB implicates may compensate for this underestimation.

The overestimation in the total variance using B=5 and $C=20$ yields the best overall coverage, although these coverage rates are rarely nominal. It is difficult to disentangle the combined effects of the biased estimates, biased variance estimates, and differing critical values on coverage. However, we note that the addition of more FPBB implicates should not improve the coverage over B = 5, as the critical value will eventually converge to z = 1.645 when B = 50 and the estimates of $\hat{V}_{samp}$ (given C = 20) are likewise slowly approaching a common smaller value as the number of implicates increase.

5. Conclusion

The research presented in this paper was done collectively by a team of methodologists. As a team, we found that the simulation process and evaluation were more interesting than the actual results. Indeed, for the rarely reported items especially, we were a bit discouraged by the statistical properties of their variance estimates using the recommended method. In this case, we are using sample-based estimators when small area estimation techniques would likely be more appropriate. On the other hand, our recommended method is easy to implement in the existing production environment and does not provide entirely unreliable estimates of precision for the more frequently-reported items. This additional information should prove valuable for data users.

This simulation was an ambitious undertaking, even with the limited number of test industries and studied products. We found that our intuition was often not confirmed by the simulation results, especially with respect to the FPBB portion. This could be a function of the small sampling error component in the Economic Census. In hindsight, we might have been better served with alternative numbers of FPBB and ABB implicates, allocating the 400 implicates differently. On the other, production concerns were expressed throughout our research process, especially with respect to the replication of hot deck imputation. Respecting these concerns in the research process led to a feasible implementation, currently being tested for the 2017 Economic Census.

Acknowledgements

References

Andridge, R.R. and Little, R.J. (2009). The Use of Sample Weights in Hot Deck Imputation. *JOS*, **25**, pp. 21-36.

Andridge, R. R., & Little, R. J. (2010). A Review of Hot Deck Imputation for Survey Non-response. *International Statistical Review, 78*(1), 40-64. doi:10.1111/j.1751-5823.2010.00103.x

Bechtel, L., Morris, D.S., and Thompson, K.J. (2015). Using Classification Trees to Recommend Hot Deck Imputation Methods: A Case Study. *Proceedings of the FCSM Research Conference.*

Brick, J., & Kalton, G. (1996). Handling missing data in survey research. *Statistical Methods in Medical Research, 5*(3), 215-238. doi:10.1177/096228029600500302

Chen, J., & Shao, J. (2001). Jackknife Variance Estimation for Nearest-Neighbor Imputation. *Journal of the American Statistical Association, 96*(453), 260-269. doi:10.1198/016214501750332839

Ellis, Y. and Thompson, K.J. (2015). Exploratory Data Analysis of Economic Census Products: Methods and Results. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Kim, J. K., & Fuller, W. (2004). Fractional hot deck imputation. *Biometrika, 91*(3), 559-578. doi:10.1093/biomet/91.3.559

Knutson, J. and Martin, J. (2015). Evaluation of Alternative Imputation Methods for Economic Census Products: The Cook-Off. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Thompson, J. R. (2000). Simulation: A Modeler's Approach. New York: Wiley.

Thompson, K.J. and Liu, X. (2015). On Recommending a Single Imputation Method for Economic Census Products. *Proceedings of the Section on Government Statistics*, American Statistical Association.

Thompson, K.J. and Thompson, M. (2016). Estimating the Variance Due to Hot Deck Imputation for Product Value Estimates in the 2017 Economic Census. . *Proceedings of the Governments Statistics Section*, American Statistical Association.

Thompson, M. Thompson, K.J., and Kurec. R. (2016). Variance Estimation for Product Value Estimates in the 2017 Economic Census Under the Assumption of Complete Response. *Proceedings of the Governments Statistics Section*, American Statistical Association.

Tolliver, K. and Bechtel, L. (2015). Implementation of Hot Deck Imputation on Economic Census Products. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Zhou, H., Raghunathan, T., and Elliot, M. (2012). A Semi-Parametric Approach to Account for Complex Designs in Multiple Imputation. *Proceedings of the Proceedings of the FCSM Research Conference.*

Appendix 1. Ratios of median variance *C*=10 / *C*=5*, C*=15 / *C*=10*,* and *C*=20 / *C*=15*,* for *B*= 5, imputed by HDN, by sector, industry and product

Sector	Industry	Product	$MedVar_{C=10}$/ $MedVar_{C=5}$	$MedVar_{C=15}$/ $MedVar_{C=10}$	$MedVar_{C=20}$/ $MedVar_{C=15}$
FIR	1	1	0.027	3.683	1.460
FIR	1	2	0.026	3.757	1.482
FIR	1	3	0.027	3.736	1.465
FIR	1	4	0.028	3.654	1.449
FIR	2	1	0.026	3.996	1.446
FIR	2	2	0.026	3.870	1.466
FIR	2	3	0.026	3.872	1.451
FIR	2	4	0.026	3.810	1.464
FIR	3	1	0.027	3.831	1.427
FIR	3	2	0.027	3.787	1.425
FIR	3	3	0.027	3.769	1.425
FIR	3	4	0.027	3.704	1.441
FIR	4	1	0.027	3.719	1.446
FIR	4	2	0.027	3.730	1.449
FIR	4	3	0.027	3.765	1.448
FIR	4	4	0.027	3.750	1.447
FIR	5	1	0.027	3.776	1.456
FIR	5	2	0.027	3.761	1.451
FIR	5	3	0.027	3.799	1.454
FIR	5	4	0.027	3.811	1.464
MAN	1	1	0.027	3.786	1.458
MAN	1	2	0.026	3.860	1.459
MAN	1	3	0.027	3.692	1.453
MAN	1	4	0.025	3.964	1.458
MAN	2	1	0.025	3.910	1.462
MAN	2	2	0.026	3.896	1.456
MAN	2	3	0.025	3.959	1.478
MAN	2	4	0.025	4.008	1.466
MAN	3	1	0.027	3.712	1.455
MAN	3	2	0.027	3.729	1.454
MAN	3	3	0.027	3.727	1.456
MAN	3	4	0.027	3.736	1.452
MAN	4	1	0.028	3.656	1.445
MAN	4	2	0.027	3.773	1.447
MAN	4	3	0.027	3.707	1.440
MAN	4	4	0.027	3.730	1.454
MAN	5	1	0.027	3.756	1.452
MAN	5	2	0.027	3.727	1.448
MAN	5	3	0.027	3.819	1.442
MAN	5	4	0.027	3.733	1.450
MIN	1	1	0.028	3.598	1.463
MIN	1	2	0.027	3.802	1.453
MIN	1	3	0.027	3.782	1.453

Sector	Industry	Product	$MedVar_{C=10}$/ $MedVar_{C=5}$	$MedVar_{C=15}$/ $MedVar_{C=10}$	$MedVar_{C=20}$/ $MedVar_{C=15}$
MIN	1	4	0.026	3.889	1.447
MIN	2	1	0.026	3.782	1.464
MIN	2	2	0.026	3.848	1.456
MIN	2	3	0.026	3.854	1.454
MIN	2	4	0.025	3.932	1.460
MIN	3	1	0.026	3.813	1.445
MIN	3	2	0.026	3.891	1.457
MIN	3	3	0.026	3.851	1.444
MIN	3	4	0.027	3.816	1.452
MIN	4	1	0.024	4.079	1.487
MIN	4	2	0.027	3.760	1.456
MIN	4	3	0.025	3.937	1.468
MIN	4	4	0.025	3.918	1.461
RET	1	1	0.027	3.723	1.449
RET	1	2	0.027	3.790	1.456
RET	1	3	0.027	3.728	1.451
RET	1	4	0.027	3.802	1.457
RET	2	1	0.027	3.736	1.444
RET	2	2	0.027	3.787	1.447
RET	2	3	0.027	3.790	1.441
RET	2	4	0.027	3.801	1.447
RET	3	1	0.023	4.286	1.472
RET	3	2	0.024	4.139	1.469
RET	3	3	0.023	4.204	1.466
RET	3	4	0.024	4.139	1.466
SER	1	1	0.027	3.763	1.456
SER	1	2	0.026	3.837	1.454
SER	1	3	0.027	3.718	1.448
SER	1	4	0.026	3.819	1.453
SER	2	1	0.027	3.749	1.467
SER	2	2	0.027	3.723	1.452
SER	2	3	0.027	3.727	1.452
SER	2	4	0.027	3.757	1.445
SER	3	1	0.027	3.749	1.449
SER	3	2	0.027	3.786	1.448
SER	3	3	0.027	3.775	1.453
SER	3	4	0.027	3.772	1.448
SER	4	1	0.027	3.787	1.432
SER	4	2	0.027	3.750	1.453
SER	4	3	0.027	3.689	1.449
SER	4	4	0.028	3.659	1.445
UTL	1	1	0.026	3.882	1.437
UTL	1	2	0.028	3.623	1.449
UTL	1	3	0.028	3.675	1.450
UTL	1	4	0.026	3.927	1.440
UTL	2	1	0.026	3.899	1.448
UTL	2	2	0.027	3.747	1.451

Sector	Industry	Product	$MedVar_{C=10}$/ $MedVar_{C=5}$	$MedVar_{C=15}$/ $MedVar_{C=10}$	$MedVar_{C=20}$/ $MedVar_{C=15}$
UTL	2	3	0.027	3.777	1.453
UTL	2	4	0.027	3.776	1.449
UTL	3	1	0.028	3.653	1.444
UTL	3	2	0.027	3.674	1.467
UTL	3	3	0.028	3.727	1.445
UTL	3	4	0.027	3.820	1.454
UTL	4	1	0.027	3.765	1.448
UTL	4	2	0.027	3.798	1.452
UTL	4	3	0.027	3.769	1.448
UTL	4	4	0.027	3.804	1.448
WHO	1	1	0.027	3.751	1.462
WHO	1	2	0.026	3.881	1.463
WHO	1	3	0.027	3.782	1.450
WHO	1	4	0.026	3.816	1.448
WHO	2	1	0.026	3.815	1.468
WHO	2	2	0.026	3.886	1.462
WHO	2	3	0.026	3.794	1.483
WHO	2	4	0.026	3.830	1.468
WHO	3	1	0.027	3.787	1.447
WHO	3	2	0.027	3.794	1.442
WHO	3	3	0.027	3.841	1.447
WHO	3	4	0.026	3.846	1.448
WHO	4	1	0.024	4.003	1.473
WHO	4	2	0.024	3.965	1.473
WHO	4	3	0.025	3.906	1.489
WHO	4	4	0.025	3.799	1.470
WHO	5	1	0.029	3.500	1.441
WHO	5	2	0.027	3.748	1.434
WHO	5	3	0.027	3.693	1.427
WHO	5	4	0.027	3.684	1.431

Appendix 2. Ratios of median variance *C*=10 / *C*=5*, C*=15 / *C*=10*,* and *C*=20 / *C*=15*,* for *B*= 5, imputed by HDR, by sector, industry and product

Sector	Industry	Product	MedVarC=10/ MedVarC=5	MedVarC=15/ MedVarC=10	MedVarC=20/ MedVarC=15
FIR	1	1	0.027	3.682	1.458
FIR	1	2	0.026	3.757	1.483
FIR	1	3	0.027	3.727	1.464
FIR	1	4	0.027	3.693	1.452
FIR	2	1	0.026	3.972	1.442
FIR	2	2	0.027	3.728	1.452
FIR	2	3	0.027	3.773	1.443
FIR	2	4	0.027	3.732	1.440
FIR	3	1	0.027	3.827	1.427
FIR	3	2	0.027	3.744	1.432
FIR	3	3	0.027	3.813	1.428
FIR	3	4	0.027	3.759	1.450
FIR	4	1	0.027	3.729	1.449
FIR	4	2	0.027	3.785	1.447
FIR	4	3	0.027	3.779	1.449
FIR	4	4	0.027	3.778	1.449
FIR	5	1	0.027	3.752	1.452
FIR	5	2	0.027	3.724	1.453
FIR	5	3	0.027	3.727	1.450
FIR	5	4	0.027	3.719	1.452
MAN	1	1	0.027	3.804	1.450
MAN	1	2	0.027	3.788	1.450
MAN	1	3	0.027	3.793	1.450
MAN	1	4	0.027	3.799	1.446
MAN	2	1	0.027	3.766	1.455
MAN	2	2	0.027	3.760	1.454
MAN	2	3	0.026	3.804	1.471
MAN	2	4	0.027	3.780	1.457
MAN	3	1	0.027	3.745	1.453
MAN	3	2	0.027	3.797	1.449
MAN	3	3	0.027	3.796	1.449
MAN	3	4	0.027	3.786	1.448
MAN	4	1	0.027	3.772	1.450
MAN	4	2	0.027	3.793	1.451
MAN	4	3	0.027	3.788	1.452
MAN	4	4	0.027	3.797	1.452
MAN	5	1	0.027	3.751	1.449
MAN	5	2	0.027	3.721	1.449
MAN	5	3	0.027	3.789	1.444
MAN	5	4	0.027	3.747	1.450
MIN	1	1	0.028	3.596	1.463
MIN	1	2	0.027	3.794	1.452
MIN	1	3	0.027	3.765	1.452

Sector	Industry	Product	MedVarC=10/ MedVarC=5	MedVarC=15/ MedVarC=10	MedVarC=20/ MedVarC=15
MIN	1	4	0.027	3.769	1.452
MIN	2	1	0.027	3.763	1.448
MIN	2	2	0.027	3.748	1.449
MIN	2	3	0.027	3.747	1.450
MIN	2	4	0.027	3.756	1.449
MIN	3	1	0.027	3.864	1.448
MIN	3	2	0.026	3.844	1.455
MIN	3	3	0.027	3.765	1.445
MIN	3	4	0.027	3.760	1.450
MIN	4	1	0.024	4.042	1.479
MIN	4	2	0.027	3.755	1.459
MIN	4	3	0.027	3.796	1.454
MIN	4	4	0.027	3.771	1.454
RET	1	1	0.027	3.718	1.449
RET	1	2	0.027	3.769	1.454
RET	1	3	0.027	3.767	1.454
RET	1	4	0.027	3.762	1.453
RET	2	1	0.027	3.737	1.446
RET	2	2	0.027	3.735	1.447
RET	2	3	0.027	3.767	1.450
RET	2	4	0.027	3.762	1.450
RET	3	1	0.023	4.273	1.472
RET	3	2	0.023	4.260	1.474
RET	3	3	0.023	4.253	1.473
RET	3	4	0.023	4.259	1.474
SER	1	1	0.027	3.795	1.452
SER	1	2	0.027	3.795	1.451
SER	1	3	0.027	3.759	1.451
SER	1	4	0.027	3.800	1.449
SER	2	1	0.027	3.746	1.462
SER	2	2	0.027	3.751	1.455
SER	2	3	0.027	3.784	1.454
SER	2	4	0.027	3.776	1.453
SER	3	1	0.026	3.814	1.445
SER	3	2	0.026	3.860	1.444
SER	3	3	0.027	3.827	1.445
SER	3	4	0.026	3.876	1.445
SER	4	1	0.027	3.779	1.433
SER	4	2	0.027	3.768	1.456
SER	4	3	0.027	3.772	1.456
SER	4	4	0.027	3.768	1.457
UTL	1	1	0.026	3.890	1.436
UTL	1	2	0.028	3.701	1.449
UTL	1	3	0.027	3.735	1.450
UTL	1	4	0.027	3.800	1.447
UTL	2	1	0.026	3.870	1.449
UTL	2	2	0.027	3.738	1.452

Sector	Industry	Product	MedVarC=10/ MedVarC=5	MedVarC=15/ MedVarC=10	MedVarC=20/ MedVarC=15
UTL	2	3	0.027	3.750	1.448
UTL	2	4	0.027	3.782	1.448
UTL	3	1	0.028	3.694	1.450
UTL	3	2	0.027	3.690	1.462
UTL	3	3	0.027	3.755	1.454
UTL	3	4	0.027	3.765	1.452
UTL	4	1	0.027	3.778	1.447
UTL	4	2	0.027	3.769	1.451
UTL	4	3	0.027	3.813	1.449
UTL	4	4	0.027	3.787	1.448
WHO	1	1	0.027	3.766	1.452
WHO	1	2	0.027	3.771	1.453
WHO	1	3	0.027	3.759	1.450
WHO	1	4	0.027	3.756	1.451
WHO	2	1	0.026	3.842	1.476
WHO	2	2	0.026	3.804	1.472
WHO	2	3	0.026	3.775	1.474
WHO	2	4	0.026	3.811	1.472
WHO	3	1	0.027	3.756	1.447
WHO	3	2	0.027	3.784	1.447
WHO	3	3	0.027	3.747	1.452
WHO	3	4	0.027	3.776	1.455
WHO	4	1	0.024	4.030	1.469
WHO	4	2	0.025	3.881	1.469
WHO	4	3	0.026	3.744	1.459
WHO	4	4	0.027	3.757	1.460
WHO	5	1	0.029	3.525	1.449
WHO	5	2	0.027	3.740	1.450
WHO	5	3	0.027	3.744	1.450
WHO	5	4	0.027	3.743	1.450

Appendix 3. Relative Bias of the variance for *B*= 5, 10, 15, and 20, for *C*= 20, imputed by HDN, by sector, industry and product

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Variance Rel. Bias, *B*=5	Variance Rel. Bias, *B*=10	Variance Rel. Bias, *B*=15	Variance Rel. Bias, *B*=20
FIR	1	1	98.6%	-0.03	4.81	-0.66	-0.18	0.13
FIR	1	2	16.1%	0.04	29.00	0.74	3.21	4.89
FIR	1	3	6.2%	-0.08	17.35	0.07	1.59	2.58
FIR	1	4	43.0%	0.04	-0.86	-0.99	-0.98	-0.97
FIR	2	1	99.0%	-0.02	13.80	1.05	1.05	1.88
FIR	2	2	19.6%	-0.23	0.09	-0.85	-0.85	-0.79
FIR	2	3	5.0%	0.23	0.88	-0.74	-0.74	-0.64
FIR	2	4	73.8%	-0.23	0.98	-1.00	-1.00	-1.00
FIR	3	1	93.4%	-0.16	3.99	-0.71	-0.03	-0.02
FIR	3	2	59.7%	-0.13	5.34	-0.63	-0.13	0.23
FIR	3	3	29.3%	-0.24	4.73	-0.66	-0.21	0.12
FIR	3	4	17.5%	-0.13	-0.96	-1.00	-0.99	-0.99
FIR	4	1	99.9%	-0.10	-0.09	-0.99	-0.98	-0.98
FIR	4	2	15.4%	-0.02	0.87	-0.89	-0.74	-0.64
FIR	4	3	18.0%	-0.10	-0.29	-0.96	-0.90	-0.86
FIR	4	4	9.0%	-0.02	-0.51	-0.97	-0.93	-0.90
FIR	5	1	99.9%	0.04	-0.67	-0.98	-0.95	-0.94
FIR	5	2	15.4%	-0.15	-0.92	-1.00	-0.99	-0.98
FIR	5	3	18.0%	-0.27	-0.95	-1.00	-0.99	-0.99
FIR	5	4	9.0%	-0.15	-1.00	-1.00	-1.00	-1.00
MAN	1	1	100.0%	-0.53	-0.98	-1.00	-1.00	-1.00
MAN	1	2	23.4%	-0.72	-0.99	-1.00	-1.00	-1.00
MAN	1	3	16.3%	-0.78	-0.98	-1.00	-1.00	-1.00
MAN	1	4	1.3%	-0.72	-0.98	-1.00	-1.00	-1.00
MAN	2	1	48.4%	-0.17	-0.68	-0.98	-0.96	-0.94
MAN	2	2	45.0%	0.14	-0.33	-0.96	-0.91	-0.87
MAN	2	3	52.3%	-0.11	0.49	-0.92	-0.79	-0.71
MAN	2	4	48.2%	0.14	-0.33	-0.96	-0.91	-0.87
MAN	3	1	99.9%	-0.06	-0.92	-1.00	-0.99	-0.99
MAN	3	2	19.7%	0.33	-0.79	-0.99	-0.97	-0.96
MAN	3	3	8.8%	0.58	-0.88	-0.99	-0.98	-0.98
MAN	3	4	6.1%	0.33	-0.72	-0.98	-0.96	-0.95
MAN	4	1	46.4%	0.16	0.16	-0.93	-0.84	-0.77
MAN	4	2	10.8%	-0.48	-0.77	-0.99	-0.97	-0.96
MAN	4	3	6.4%	-0.39	-0.42	-0.97	-0.92	-0.89
MAN	4	4	15.6%	-0.48	-0.89	-0.99	-0.99	-0.98
MAN	5	1	64.7%	-0.22	-0.87	-0.99	-0.98	-0.97
MAN	5	2	48.0%	-0.19	-0.76	-0.99	-0.97	-0.95
MAN	5	3	44.6%	-0.19	-0.44	-0.97	-0.92	-0.89
MAN	5	4	28.3%	-0.19	-0.99	-1.00	-1.00	-1.00
MIN	1	1	82.5%	-0.06	3.15	-0.75	-0.41	-0.19
MIN	1	2	70.4%	0.01	4.67	-0.67	-0.20	0.11
MIN	1	3	2.1%	-0.12	2.97	-0.77	-0.44	-0.23

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Variance Rel. Bias, *B*=5	Variance Rel. Bias, *B*=10	Variance Rel. Bias, *B*=15	Variance Rel. Bias, *B*=20
MIN	1	4	2.2%	0.01	-0.93	-1.00	-0.99	-0.99
MIN	2	1	53.4%	-0.28	-0.52	-0.97	-0.93	-0.91
MIN	2	2	54.7%	-0.49	-0.93	-1.00	-0.99	-0.99
MIN	2	3	49.7%	-0.45	-0.88	-0.99	-0.98	-0.98
MIN	2	4	40.9%	-0.49	-0.70	-0.98	-0.96	-0.94
MIN	3	1	49.9%	-0.14	0.67	-0.91	-0.77	-0.68
MIN	3	2	3.5%	-0.27	-0.22	-0.95	-0.89	-0.85
MIN	3	3	15.2%	-0.13	0.14	-0.93	-0.84	-0.78
MIN	3	4	11.3%	-0.27	-0.98	-1.00	-1.00	-1.00
MIN	4	1	72.0%	0.00	10.29	-0.36	0.59	1.21
MIN	4	2	10.2%	-0.01	1.78	-0.84	-0.61	-0.46
MIN	4	3	17.8%	-0.02	2.36	-0.80	-0.52	-0.34
MIN	4	4	8.0%	-0.01	-0.85	-0.99	-0.98	-0.97
RET	1	1	100.0%	-0.01	0.96	-0.89	-0.73	-0.62
RET	1	2	69.2%	-0.01	0.71	-0.90	-0.76	-0.66
RET	1	3	66.3%	-0.05	-0.33	-0.96	-0.91	-0.87
RET	1	4	53.4%	-0.01	-0.33	-0.96	-0.91	-0.87
RET	2	1	100.0%	-0.02	-0.26	-0.96	-0.90	-0.86
RET	2	2	74.9%	0.00	1.58	-0.85	-0.64	-0.50
RET	2	3	41.4%	-0.03	0.05	-0.94	-0.85	-0.79
RET	2	4	47.5%	0.00	-0.86	-0.99	-0.98	-0.97
RET	3	1	100.0%	-0.01	0.28	-0.92	-0.82	-0.75
RET	3	2	100.0%	-0.09	-0.95	-1.00	-0.99	-0.99
RET	3	3	99.9%	-0.07	-0.93	-1.00	-0.99	-0.99
RET	3	4	95.1%	-0.09	-1.00	-1.00	-1.00	-1.00
SER	1	1	78.6%	-0.24	-0.54	-0.97	-0.94	-0.91
SER	1	2	29.4%	-0.17	-0.28	-0.96	-0.90	-0.86
SER	1	3	20.1%	-0.15	-0.12	-0.95	-0.88	-0.83
SER	1	4	16.5%	-0.17	0.21	-0.93	-0.83	-0.76
SER	2	1	79.5%	0.02	2.47	-0.81	-0.53	-0.33
SER	2	2	22.4%	-0.05	0.98	-0.89	-0.72	-0.61
SER	2	3	14.7%	-0.09	0.09	-0.94	-0.85	-0.79
SER	2	4	10.5%	-0.05	-0.67	-0.98	-0.95	-0.94
SER	3	1	94.6%	-0.26	-0.91	-0.99	-0.99	-0.98
SER	3	2	61.8%	-0.20	-0.70	-0.98	-0.96	-0.94
SER	3	3	55.4%	-0.31	-0.79	-0.99	-0.97	-0.96
SER	3	4	57.7%	-0.20	-0.80	-0.99	-0.97	-0.96
SER	4	1	100.0%	-0.01	3.02	-0.77	-0.44	-0.22
SER	4	2	41.2%	-0.09	-0.51	-0.97	-0.93	-0.90
SER	4	3	14.2%	0.16	-0.23	-0.96	-0.89	-0.85
SER	4	4	23.6%	-0.09	-0.96	-1.00	-0.99	-0.99
UTL	1	1	83.6%	-0.03	9.02	-0.45	0.38	0.94
UTL	1	2	8.8%	0.19	0.35	-0.93	-0.81	-0.74
UTL	1	3	12.8%	-0.17	1.65	-0.85	-0.63	-0.49
UTL	1	4	2.3%	0.19	-0.26	-0.96	-0.90	-0.86
UTL	2	1	21.3%	-0.32	-0.69	-0.98	-0.96	-0.94
UTL	2	2	36.2%	-0.15	-0.49	-0.97	-0.93	-0.90
UTL	2	3	43.6%	-0.15	-0.46	-0.97	-0.92	-0.89

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Variance Rel. Bias, *B*=5	Variance Rel. Bias, *B*=10	Variance Rel. Bias, *B*=15	Variance Rel. Bias, *B*=20
UTL	2	4	11.0%	-0.15	-0.65	-0.98	-0.95	-0.93
UTL	3	1	51.7%	0.00	2.60	-0.79	-0.50	-0.30
UTL	3	2	21.3%	-0.14	2.09	-0.83	-0.57	-0.40
UTL	3	3	39.7%	0.04	2.21	-0.81	-0.55	-0.37
UTL	3	4	10.5%	-0.14	-0.66	-0.98	-0.95	-0.93
UTL	4	1	100.0%	-0.07	-0.65	-0.98	-0.95	-0.93
UTL	4	2	8.0%	-0.14	0.62	-0.91	-0.77	-0.68
UTL	4	3	5.9%	-0.16	0.14	-0.93	-0.84	-0.78
UTL	4	4	1.6%	-0.14	0.15	-0.93	-0.84	-0.78
WHO	1	1	53.1%	-0.24	-0.84	-0.99	-0.98	-0.97
WHO	1	2	42.3%	0.03	0.67	-0.90	-0.77	-0.68
WHO	1	3	29.9%	-0.04	0.56	-0.91	-0.78	-0.70
WHO	1	4	34.1%	0.03	-0.70	-0.98	-0.96	-0.94
WHO	2	1	58.4%	-0.26	-0.76	-0.99	-0.97	-0.95
WHO	2	2	38.8%	-0.08	0.37	-0.92	-0.81	-0.73
WHO	2	3	17.1%	-0.58	-0.80	-0.99	-0.97	-0.96
WHO	2	4	32.4%	-0.08	-0.61	-0.98	-0.95	-0.92
WHO	3	1	100.0%	-0.02	3.34	-0.75	-0.39	-0.16
WHO	3	2	24.6%	-0.05	2.12	-0.82	-0.56	-0.39
WHO	3	3	9.0%	0.00	1.66	-0.84	-0.63	-0.48
WHO	3	4	7.9%	-0.05	-0.13	-0.95	-0.88	-0.83
WHO	4	1	100.0%	-0.18	1.36	-0.86	-0.67	-0.54
WHO	4	2	65.9%	-0.15	2.10	-0.81	-0.57	-0.39
WHO	4	3	6.8%	0.45	0.64	-0.91	-0.76	-0.68
WHO	4	4	22.7%	-0.15	-1.00	-1.00	-1.00	-1.00
WHO	5	1	100.0%	-0.41	-0.86	-0.99	-0.98	-0.97
WHO	5	2	60.2%	0.05	0.67	-0.90	-0.77	-0.67
WHO	5	3	22.1%	-0.07	2.51	-0.80	-0.51	-0.31
WHO	5	4	20.7%	0.05	-0.98	-1.00	-1.00	-1.00

Appendix 4. Relative Bias of the variance for *B*= 5, 10, 15, and 20, for *C*= 20, imputed by HDR, by sector, industry and product

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Variance Rel. Bias, *B*=5	Variance Rel. Bias, *B*=10	Variance Rel. Bias, *B*=15	Variance Rel. Bias, *B*=20
FIR	1	1	98.6%	0.09	-0.24	-0.96	-0.89	-0.85
FIR	1	2	16.1%	0.49	-0.28	-0.96	-0.90	-0.86
FIR	1	3	6.2%	2.85	-0.97	-1.00	-1.00	-0.99
FIR	1	4	43.0%	0.49	-0.99	-1.00	-1.00	-1.00
FIR	2	1	99.0%	0.00	65.62	2.95	8.23	11.95
FIR	2	2	19.6%	6.88	-1.00	-1.00	-1.00	-1.00
FIR	2	3	5.0%	23.80	-1.00	-1.00	-1.00	-1.00
FIR	2	4	73.8%	6.88	-1.00	-1.00	-1.00	-1.00
FIR	3	1	93.4%	0.08	7.52	-0.50	0.18	0.67
FIR	3	2	59.7%	0.16	1.84	-0.84	-0.61	-0.45
FIR	3	3	29.3%	0.61	1.67	-0.84	-0.63	-0.48
FIR	3	4	17.5%	0.16	0.07	-0.94	-0.85	-0.79
FIR	4	1	99.9%	-0.01	4.77	-0.67	-0.19	0.13
FIR	4	2	15.4%	8.32	-1.00	-1.00	-1.00	-1.00
FIR	4	3	18.0%	11.76	-1.00	-1.00	-1.00	-1.00
FIR	4	4	9.0%	8.32	-1.00	-1.00	-1.00	-1.00
FIR	5	1	99.9%	0.08	-0.90	-0.99	-0.99	-0.98
FIR	5	2	15.4%	3.28	-1.00	-1.00	-1.00	-1.00
FIR	5	3	18.0%	5.01	-1.00	-1.00	-1.00	-1.00
FIR	5	4	9.0%	3.28	-1.00	-1.00	-1.00	-1.00
MAN	1	1	100.0%	0.15	-0.88	-0.99	-0.98	-0.10
MAN	1	2	23.4%	6.13	-1.00	-1.00	-1.00	-1.00
MAN	1	3	16.3%	19.50	-1.00	-1.00	-1.00	-1.00
MAN	1	4	1.3%	6.13	-1.00	-1.00	-1.00	-1.00
MAN	2	1	48.4%	-0.47	-0.92	-1.00	-0.99	-0.98
MAN	2	2	45.0%	-0.38	-0.84	-0.99	-0.98	-0.97
MAN	2	3	52.3%	-0.09	1.34	-0.87	-0.67	-0.54
MAN	2	4	48.2%	-0.38	-0.84	-0.99	-0.98	-0.97
MAN	3	1	99.9%	-0.02	-0.35	-0.96	-0.91	-0.87
MAN	3	2	19.7%	63.22	-1.00	-1.00	-1.00	-1.00
MAN	3	3	8.8%	117.67	-1.00	-1.00	-1.00	-1.00
MAN	3	4	6.1%	63.22	-1.00	-1.00	-1.00	-1.00
MAN	4	1	46.4%	0.80	-0.77	-0.99	-0.97	-0.96
MAN	4	2	10.8%	0.75	-0.67	-0.98	-0.95	-0.94
MAN	4	3	6.4%	3.01	-0.88	-0.99	-0.98	-0.98
MAN	4	4	15.6%	0.75	-0.70	-0.98	-0.96	-0.94
MAN	5	1	64.7%	-0.10	-0.23	-0.96	-0.89	-0.85
MAN	5	2	48.0%	0.17	-0.58	-0.98	-0.94	-0.92
MAN	5	3	44.6%	0.33	-0.72	-0.98	-0.96	-0.94
MAN	5	4	28.3%	0.17	-0.71	-0.98	-0.96	-0.94
MIN	1	1	82.5%	-0.05	4.91	-0.64	-0.17	0.15
MIN	1	2	70.4%	0.61	-0.90	-0.99	-0.99	-0.98
MIN	1	3	2.1%	20.72	-0.97	-1.00	-1.00	-0.99
MIN	1	4	2.2%	0.61	-0.94	-1.00	-0.99	-0.99

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Variance Rel. Bias, B=5	Variance Rel. Bias, B=10	Variance Rel. Bias, B=15	Variance Rel. Bias, B=20
MIN	2	1	53.4%	0.66	-0.36	-0.96	-0.91	-0.88
MIN	2	2	54.7%	1.62	-0.68	-0.98	-0.95	-0.94
MIN	2	3	49.7%	1.93	-0.71	-0.98	-0.96	-0.94
MIN	2	4	40.9%	1.62	-0.68	-0.98	-0.95	-0.94
MIN	3	1	49.9%	-0.01	3.92	-0.73	-0.32	-0.04
MIN	3	2	3.5%	0.35	-0.28	-0.96	-0.90	-0.86
MIN	3	3	15.2%	2.98	-0.89	-0.99	-0.98	-0.98
MIN	3	4	11.3%	0.35	-0.55	-0.97	-0.94	-0.91
MIN	4	1	72.0%	-0.02	7.67	-0.51	0.21	0.69
MIN	4	2	10.2%	2.18	-0.93	-1.00	-0.99	-0.99
MIN	4	3	17.8%	7.55	-0.97	-1.00	-1.00	-0.99
MIN	4	4	8.0%	2.18	-0.95	-1.00	-0.99	-0.99
RET	1	1	100.0%	-0.02	-0.41	-0.97	-0.92	-0.89
RET	1	2	69.2%	32.03	-1.00	-1.00	-1.00	-1.00
RET	1	3	66.3%	38.73	-1.00	-1.00	-1.00	-1.00
RET	1	4	53.4%	32.03	-1.00	-1.00	-1.00	-1.00
RET	2	1	100.0%	0.01	0.65	-0.91	-0.77	-0.68
RET	2	2	74.9%	5.25	-1.00	-1.00	-1.00	-1.00
RET	2	3	41.4%	23.68	-1.00	-1.00	-1.00	-1.00
RET	2	4	47.5%	5.25	-1.00	-1.00	-1.00	-1.00
RET	3	1	100.0%	0.04	-0.80	-0.99	-0.97	-0.96
RET	3	2	100.0%	1.12	-1.00	-1.00	-1.00	-1.00
RET	3	3	99.9%	6.03	-1.00	-1.00	-1.00	-1.00
RET	3	4	95.1%	1.12	-1.00	-1.00	-1.00	-1.00
SER	1	1	78.6%	0.00	3.91	-0.72	-0.32	-0.04
SER	1	2	29.4%	0.62	-0.76	-0.99	-0.97	-0.95
SER	1	3	20.1%	1.89	-0.90	-0.99	-0.99	-0.98
SER	1	4	16.5%	0.62	-0.71	-0.98	-0.96	-0.94
SER	2	1	79.5%	0.04	1.08	-0.88	-0.71	-0.60
SER	2	2	22.4%	2.27	-0.99	-1.00	-1.00	-1.00
SER	2	3	14.7%	6.68	-0.99	-1.00	-1.00	-1.00
SER	2	4	10.5%	2.27	-0.99	-1.00	-1.00	-1.00
SER	3	1	94.6%	0.05	0.51	-0.91	-0.79	-0.71
SER	3	2	61.8%	1.36	-0.97	-1.00	-1.00	-0.99
SER	3	3	55.4%	1.62	-0.97	-1.00	-1.00	-0.99
SER	3	4	57.7%	1.36	-0.97	-1.00	-1.00	-0.99
SER	4	1	100.0%	-0.01	2.00	-0.83	-0.58	-0.42
SER	4	2	41.2%	8.35	-1.00	-1.00	-1.00	-1.00
SER	4	3	14.2%	31.12	-1.00	-1.00	-1.00	-1.00
SER	4	4	23.6%	8.35	-1.00	-1.00	-1.00	-1.00
UTL	1	1	83.6%	-0.02	21.87	0.26	2.14	3.42
UTL	1	2	8.8%	3.03	-0.98	-1.00	-1.00	-1.00
UTL	1	3	12.8%	7.78	-0.99	-1.00	-1.00	-1.00
UTL	1	4	2.3%	3.03	-0.99	-1.00	-1.00	-1.00
UTL	2	1	21.3%	-0.14	0.38	-0.92	-0.81	-0.73
UTL	2	2	36.2%	-0.14	-0.31	-0.96	-0.90	-0.86
UTL	2	3	43.6%	-0.17	-0.51	-0.97	-0.93	-0.90
UTL	2	4	11.0%	-0.14	-0.42	-0.97	-0.92	-0.89

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Variance Rel. Bias, B=5	Variance Rel. Bias, B=10	Variance Rel. Bias, B=15	Variance Rel. Bias, B=20
UTL	3	1	51.7%	-0.06	1.12	-0.88	-0.71	-0.59
UTL	3	2	21.3%	1.29	-0.88	-0.99	-0.98	-0.98
UTL	3	3	39.7%	0.93	-0.90	-0.99	-0.99	-0.98
UTL	3	4	10.5%	1.29	-0.94	-1.00	-0.99	-0.99
UTL	4	1	100.0%	0.02	1.03	-0.88	-0.72	-0.61
UTL	4	2	8.0%	25.41	-1.00	-1.00	-1.00	-1.00
UTL	4	3	5.9%	26.62	-1.00	-1.00	-1.00	-1.00
UTL	4	4	1.6%	25.41	-1.00	-1.00	-1.00	-1.00
WHO	1	1	53.1%	-0.22	-0.71	-0.98	-0.96	-0.94
WHO	1	2	42.3%	0.56	-0.84	-0.99	-0.98	-0.97
WHO	1	3	29.9%	1.76	-0.95	-1.00	-0.99	-0.99
WHO	1	4	34.1%	0.56	-0.85	-0.99	-0.98	-0.97
WHO	2	1	58.4%	-0.38	-0.85	-0.99	-0.98	-0.97
WHO	2	2	38.8%	0.96	-0.77	-0.99	-0.97	-0.95
WHO	2	3	17.1%	1.48	-0.82	-0.99	-0.97	-0.96
WHO	2	4	32.4%	0.96	-0.77	-0.99	-0.97	-0.96
WHO	3	1	100.0%	0.05	0.53	-0.91	-0.79	-0.70
WHO	3	2	24.6%	4.68	-1.00	-1.00	-1.00	-1.00
WHO	3	3	9.0%	12.71	-1.00	-1.00	-1.00	-1.00
WHO	3	4	7.9%	4.68	-1.00	-1.00	-1.00	-1.00
WHO	4	1	100.0%	-0.04	25.72	0.65	2.74	4.27
WHO	4	2	65.9%	1.71	-0.91	-0.99	-0.99	-0.98
WHO	4	3	6.8%	17.90	-0.97	-1.00	-1.00	-0.99
WHO	4	4	22.7%	1.71	-0.95	-1.00	-0.99	-0.99
WHO	5	1	100.0%	-0.13	0.44	-0.92	-0.80	-0.72
WHO	5	2	60.2%	9.14	-0.98	-1.00	-1.00	-1.00
WHO	5	3	22.1%	79.24	-0.98	-1.00	-1.00	-1.00
WHO	5	4	20.7%	9.14	-0.98	-1.00	-1.00	-1.00

Appendix 5. Coverage Rates for C= 20, imputed by HDN, by sector, industry and product

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Coverage, B=5	Coverage, B=10	Coverage, B=15	Coverage, B=20
FIR	1	1	98.6%	-0.03	100.0%	53.6%	85.0%	91.4%
FIR	1	2	16.1%	0.04	100.0%	97.8%	99.5%	99.8%
FIR	1	3	6.2%	-0.08	99.9%	96.8%	99.0%	99.2%
FIR	1	4	43.0%	0.04	18.5%	5.5%	8.5%	10.8%
FIR	2	1	99.0%	-0.02	99.6%	76.7%	97.6%	98.1%
FIR	2	2	19.6%	-0.23	86.0%	17.2%	33.1%	43.4%
FIR	2	3	5.0%	0.23	97.8%	33.2%	53.1%	62.5%
FIR	2	4	73.8%	-0.23	6.6%	1.6%	2.5%	3.2%
FIR	3	1	93.4%	-0.16	71.2%	71.2%	71.2%	71.2%
FIR	3	2	59.7%	-0.13	83.0%	75.4%	81.8%	82.4%
FIR	3	3	29.3%	-0.24	71.2%	71.2%	71.2%	71.2%
FIR	3	4	17.5%	-0.13	45.0%	13.9%	21.4%	24.0%
FIR	4	1	99.9%	-0.10	31.9%	11.4%	17.1%	19.6%
FIR	4	2	15.4%	-0.02	86.6%	27.7%	42.4%	49.2%
FIR	4	3	18.0%	-0.10	71.7%	17.1%	25.6%	29.7%
FIR	4	4	9.0%	-0.02	56.4%	13.9%	20.4%	25.4%
FIR	5	1	99.9%	0.04	44.3%	7.8%	12.4%	15.6%
FIR	5	2	15.4%	-0.15	15.7%	2.0%	4.0%	5.5%
FIR	5	3	18.0%	-0.27	6.4%	0.1%	0.9%	1.4%
FIR	5	4	9.0%	-0.15	0.2%	0.1%	0.1%	0.2%
MAN	1	1	100.0%	-0.53	4.5%	1.0%	1.9%	1.9%
MAN	1	2	23.4%	-0.72	2.4%	0.3%	0.8%	0.9%
MAN	1	3	16.3%	-0.78	2.7%	1.1%	1.5%	1.8%
MAN	1	4	1.3%	-0.72	1.1%	0.0%	0.2%	0.2%
MAN	2	1	48.4%	-0.17	46.2%	11.4%	18.3%	21.6%
MAN	2	2	45.0%	0.14	57.4%	15.6%	25.2%	2.9%
MAN	2	3	52.3%	-0.11	61.8%	17.0%	26.5%	31.7%
MAN	2	4	48.2%	0.14	51.9%	16.8%	24.8%	29.1%
MAN	3	1	99.9%	-0.06	7.1%	0.5%	0.9%	1.4%
MAN	3	2	19.7%	0.33	39.9%	8.7%	14.4%	16.8%
MAN	3	3	8.8%	0.58	21.6%	2.9%	5.6%	7.6%
MAN	3	4	6.1%	0.33	42.2%	12.0%	19.6%	22.7%
MAN	4	1	46.4%	0.16	81.2%	21.4%	33.2%	39.8%
MAN	4	2	10.8%	-0.48	26.4%	14.0%	18.1%	19.0%
MAN	4	3	6.4%	-0.39	49.5%	14.1%	22.3%	26.0%
MAN	4	4	15.6%	-0.48	17.8%	4.0%	6.2%	8.0%
MAN	5	1	64.7%	-0.22	18.9%	1.7%	3.2%	3.7%
MAN	5	2	48.0%	-0.19	28.0%	3.4%	7.4%	9.1%
MAN	5	3	44.6%	-0.19	75.5%	4.3%	9.8%	13.8%
MAN	5	4	28.3%	-0.19	2.7%	0.6%	1.2%	1.3%
MIN	1	1	82.5%	-0.06	92.2%	60.7%	74.0%	78.0%
MIN	1	2	70.4%	0.01	97.5%	59.0%	78.2%	84.4%
MIN	1	3	2.1%	-0.12	93.5%	47.8%	64.9%	71.0%
MIN	1	4	2.2%	0.01	27.9%	7.4%	11.2%	12.6%
MIN	2	1	53.4%	-0.28	51.0%	15.8%	22.4%	27.0%
MIN	2	2	54.7%	-0.49	11.4%	2.3%	3.5%	4.0%
MIN	2	3	49.7%	-0.45	22.2%	4.6%	7.2%	8.5%

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Coverage, B=5	Coverage, B=10	Coverage, B=15	Coverage, B=20
MIN	2	4	40.9%	-0.49	34.9%	3.1%	5.1%	7.1%
MIN	3	1	49.9%	-0.14	53.3%	39.5%	48.8%	49.7%
MIN	3	2	3.5%	-0.27	71.4%	14.2%	29.6%	39.8%
MIN	3	3	15.2%	-0.13	77.8%	23.7%	35.9%	41.7%
MIN	3	4	11.3%	-0.27	4.6%	0.8%	1.7%	1.7%
MIN	4	1	72.0%	0.00	100.0%	58.3%	80.8%	86.9%
MIN	4	2	10.2%	-0.01	87.4%	29.1%	46.7%	55.1%
MIN	4	3	17.8%	-0.02	85.4%	34.6%	50.6%	56.5%
MIN	4	4	8.0%	-0.01	18.1%	3.2%	5.2%	6.1%
RET	1	1	100.0%	-0.01	92.0%	36.1%	53.4%	60.5%
RET	1	2	69.2%	-0.01	60.1%	15.0%	24.6%	28.7%
RET	1	3	66.3%	-0.05	51.3%	14.9%	22.1%	25.3%
RET	1	4	53.4%	-0.01	40.6%	8.8%	14.5%	17.2%
RET	2	1	100.0%	-0.02	64.1%	20.7%	29.0%	33.8%
RET	2	2	74.9%	0.00	80.1%	25.6%	40.3%	45.5%
RET	2	3	41.4%	-0.03	67.8%	20.5%	29.7%	34.4%
RET	2	4	47.5%	0.00	24.7%	5.8%	8.4%	10.2%
RET	3	1	100.0%	-0.01	40.0%	10.7%	16.0%	18.8%
RET	3	2	100.0%	-0.09	13.9%	3.5%	5.3%	6.0%
RET	3	3	99.9%	-0.07	23.2%	5.7%	9.1%	10.7%
RET	3	4	95.1%	-0.09	3.6%	1.1%	1.3%	1.4%
SER	1	1	78.6%	-0.24	41.5%	7.3%	17.9%	23.1%
SER	1	2	29.4%	-0.17	71.1%	15.5%	25.9%	32.2%
SER	1	3	20.1%	-0.15	83.3%	12.5%	24.8%	31.1%
SER	1	4	16.5%	-0.17	48.1%	16.1%	26.8%	32.9%
SER	2	1	79.5%	0.02	95.3%	32.9%	49.2%	56.1%
SER	2	2	22.4%	-0.05	84.2%	30.5%	44.1%	50.5%
SER	2	3	14.7%	-0.09	57.1%	10.8%	18.8%	23.5%
SER	2	4	10.5%	-0.05	44.7%	11.0%	17.8%	20.2%
SER	3	1	94.6%	-0.26	14.3%	4.2%	6.6%	8.0%
SER	3	2	61.8%	-0.20	32.7%	4.6%	8.2%	10.1%
SER	3	3	55.4%	-0.31	28.4%	4.7%	7.4%	8.3%
SER	3	4	57.7%	-0.20	25.3%	3.8%	6.8%	9.1%
SER	4	1	100.0%	-0.01	98.5%	42.0%	64.3%	73.2%
SER	4	2	41.2%	-0.09	60.5%	16.0%	24.4%	28.8%
SER	4	3	14.2%	0.16	74.7%	20.1%	32.4%	38.0%
SER	4	4	23.6%	-0.09	18.2%	3.9%	6.5%	7.7%
UTL	1	1	83.6%	-0.03	98.8%	78.3%	79.1%	84.2%
UTL	1	2	8.8%	0.19	100.0%	4.8%	15.8%	26.5%
UTL	1	3	12.8%	-0.17	99.5%	32.1%	48.1%	60.5%
UTL	1	4	2.3%	0.19	91.1%	3.4%	7.8%	12.0%
UTL	2	1	21.3%	-0.32	44.8%	6.8%	10.8%	12.9%
UTL	2	2	36.2%	-0.15	61.5%	8.9%	16.2%	20.6%
UTL	2	3	43.6%	-0.15	62.7%	9.8%	17.7%	21.2%
UTL	2	4	11.0%	-0.15	42.9%	6.6%	11.2%	14.1%
UTL	3	1	51.7%	0.00	90.7%	38.0%	55.1%	61.8%
UTL	3	2	21.3%	-0.14	85.1%	44.5%	60.9%	66.5%
UTL	3	3	39.7%	0.04	91.0%	34.0%	50.7%	58.6%
UTL	3	4	10.5%	-0.14	51.3%	13.1%	21.0%	24.9%

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Coverage, B=5	Coverage, B=10	Coverage, B=15	Coverage, B=20
UTL	4	1	100.0%	-0.07	56.8%	28.7%	40.5%	43.9%
UTL	4	2	8.0%	-0.14	80.1%	30.4%	46.9%	52.1%
UTL	4	3	5.9%	-0.16	76.1%	27.4%	39.1%	44.1%
UTL	4	4	1.6%	-0.14	79.4%	23.7%	35.9%	42.3%
WHO	1	1	53.1%	-0.24	22.9%	6.5%	9.8%	11.5%
WHO	1	2	42.3%	0.03	68.8%	19.5%	29.0%	34.9%
WHO	1	3	29.9%	-0.04	64.8%	18.0%	27.5%	32.3%
WHO	1	4	34.1%	0.03	33.7%	8.1%	12.5%	15.4%
WHO	2	1	58.4%	-0.26	41.2%	11.4%	16.6%	20.3%
WHO	2	2	38.8%	-0.08	43.6%	11.4%	18.1%	20.9%
WHO	2	3	17.1%	-0.58	26.4%	6.9%	11.0%	13.0%
WHO	2	4	32.4%	-0.08	22.0%	5.5%	8.6%	9.8%
WHO	3	1	100.0%	-0.02	97.7%	60.1%	74.1%	79.5%
WHO	3	2	24.6%	-0.05	92.3%	41.8%	61.5%	68.6%
WHO	3	3	9.0%	0.00	87.3%	35.2%	50.4%	56.5%
WHO	3	4	7.9%	-0.05	74.4%	23.3%	33.6%	39.5%
WHO	4	1	100.0%	-0.18	73.5%	61.6%	63.6%	69.5%
WHO	4	2	65.9%	-0.15	74.0%	63.6%	69.8%	71.3%
WHO	4	3	6.8%	0.45	95.1%	9.8%	44.9%	53.9%
WHO	4	4	22.7%	-0.15	9.8%	2.1%	3.2%	4.0%
WHO	5	1	100.0%	-0.41	23.7%	1.2%	1.6%	1.6%
WHO	5	2	60.2%	0.05	61.2%	18.5%	27.9%	32.7%
WHO	5	3	22.1%	-0.07	85.4%	31.1%	48.9%	55.3%
WHO	5	4	20.7%	0.05	10.0%	2.9%	4.4%	5.0%

Appending 6. Coverage Rates for *C*= 20, imputed by HDR, by sector, industry and product

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Coverage, *B*=5	Coverage, *B*=10	Coverage, *B*=15	Coverage, *B*=20
FIR	1	1	98.6%	0.09	87.2%	17.7%	32.7%	39.6%
FIR	1	2	16.1%	0.49	99.0%	0.0%	0.0%	0.2%
FIR	1	3	6.2%	2.85	0.0%	0.0%	0.0%	0.0%
FIR	1	4	43.0%	0.49	0.0%	0.0%	0.0%	0.0%
FIR	2	1	99.0%	0.00	100.0%	98.6%	99.9%	100.0%
FIR	2	2	19.6%	6.88	0.0%	0.0%	0.0%	0.0%
FIR	2	3	5.0%	23.80	0.0%	0.0%	0.0%	0.0%
FIR	2	4	73.8%	6.88	0.0%	0.0%	0.0%	0.0%
FIR	3	1	93.4%	0.08	99.9%	77.4%	86.3%	89.2%
FIR	3	2	59.7%	0.16	99.7%	72.0%	79.3%	82.3%
FIR	3	3	29.3%	0.61	99.0%	52.7%	76.3%	82.7%
FIR	3	4	17.5%	0.16	79.4%	21.3%	32.2%	36.7%
FIR	4	1	99.9%	-0.01	99.9%	65.0%	84.8%	90.7%
FIR	4	2	15.4%	8.32	0.0%	0.0%	0.0%	0.0%
FIR	4	3	18.0%	11.76	0.0%	0.0%	0.0%	0.0%
FIR	4	4	9.0%	8.32	0.0%	0.0%	0.0%	0.0%
FIR	5	1	99.9%	0.08	0.1%	0.0%	0.0%	0.0%
FIR	5	2	15.4%	3.28	0.0%	0.0%	0.0%	0.0%
FIR	5	3	18.0%	5.01	0.0%	0.0%	0.0%	0.0%
FIR	5	4	9.0%	3.28	0.0%	0.0%	0.0%	0.0%
MAN	1	1	100.0%	0.15	-88.3%	-99.3%	-98.3%	-9.8%
MAN	1	2	23.4%	6.13	-99.9%	-100.0%	-100.0%	-100.0%
MAN	1	3	16.3%	19.50	-99.9%	-100.0%	-100.0%	-100.0%
MAN	1	4	1.3%	6.13	-99.9%	-100.0%	-100.0%	-100.0%
MAN	2	1	48.4%	-0.47	-92.1%	-99.6%	-98.9%	-98.5%
MAN	2	2	45.0%	-0.38	-83.7%	-99.1%	-97.7%	-96.8%
MAN	2	3	52.3%	-0.09	134.4%	-87.1%	-67.3%	-54.0%
MAN	2	4	48.2%	-0.38	-83.6%	-99.1%	-97.7%	-96.8%
MAN	3	1	99.9%	-0.02	-34.7%	-96.2%	-90.8%	-87.2%
MAN	3	2	19.7%	63.22	-100.0%	-100.0%	-100.0%	-100.0%
MAN	3	3	8.8%	117.67	-100.0%	-100.0%	-100.0%	-100.0%
MAN	3	4	6.1%	63.22	-100.0%	-100.0%	-100.0%	-100.0%
MAN	4	1	46.4%	0.80	-77.3%	-98.7%	-96.8%	-95.6%
MAN	4	2	10.8%	0.75	-67.0%	-98.1%	-95.4%	-93.6%
MAN	4	3	6.4%	3.01	-88.3%	-99.3%	-98.4%	-97.7%
MAN	4	4	15.6%	0.75	-70.4%	-98.3%	-95.9%	-94.2%
MAN	5	1	64.7%	-0.10	-23.4%	-95.6%	-89.2%	-85.0%
MAN	5	2	48.0%	0.17	-58.2%	-97.6%	-94.1%	-91.8%
MAN	5	3	44.6%	0.33	-71.9%	-98.4%	-96.1%	-94.5%
MAN	5	4	28.3%	0.17	-71.2%	-98.3%	-96.0%	-94.4%
MIN	1	1	82.5%	-0.05	100.0%	64.9%	81.0%	87.7%
MIN	1	2	70.4%	0.61	0.5%	0.0%	0.0%	0.0%
MIN	1	3	2.1%	20.72	0.0%	0.0%	0.0%	0.0%

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Coverage, *B*=5	Coverage, *B*=10	Coverage, *B*=15	Coverage, *B*=20
MIN	1	4	2.2%	0.61	0.0%	0.0%	0.0%	0.0%
MIN	2	1	53.4%	0.66	76.9%	15.8%	28.2%	34.2%
MIN	2	2	54.7%	1.62	51.6%	0.8%	4.4%	6.4%
MIN	2	3	49.7%	1.93	48.2%	0.4%	2.0%	4.6%
MIN	2	4	40.9%	1.62	51.7%	0.8%	4.5%	6.7%
MIN	3	1	49.9%	-0.01	99.7%	56.6%	73.2%	79.8%
MIN	3	2	3.5%	0.35	84.7%	9.4%	19.7%	29.0%
MIN	3	3	15.2%	2.98	0.9%	0.0%	0.0%	0.0%
MIN	3	4	11.3%	0.35	64.5%	7.8%	13.5%	17.7%
MIN	4	1	72.0%	-0.02	100.0%	74.8%	91.7%	95.7%
MIN	4	2	10.2%	2.18	0.1%	0.0%	0.0%	0.0%
MIN	4	3	17.8%	7.55	0.0%	0.0%	0.0%	0.0%
MIN	4	4	8.0%	2.18	0.0%	0.0%	0.0%	0.0%
RET	1	1	100.0%	-0.02	75.8%	0.9%	1.5%	2.7%
RET	1	2	69.2%	32.03	0.5%	0.1%	0.3%	0.3%
RET	1	3	66.3%	38.73	0.5%	0.4%	0.5%	0.5%
RET	1	4	53.4%	32.03	0.4%	0.1%	0.2%	0.3%
RET	2	1	100.0%	0.01	98.2%	9.9%	27.5%	40.1%
RET	2	2	74.9%	5.25	0.5%	0.1%	0.2%	0.3%
RET	2	3	41.4%	23.68	0.5%	0.0%	0.1%	0.2%
RET	2	4	47.5%	5.25	0.0%	0.0%	0.0%	0.0%
RET	3	1	100.0%	0.04	31.7%	7.3%	11.6%	12.6%
RET	3	2	100.0%	1.12	0.5%	0.2%	0.3%	0.3%
RET	3	3	99.9%	6.03	0.5%	0.4%	0.5%	0.5%
RET	3	4	95.1%	1.12	0.1%	0.0%	0.0%	0.0%
SER	1	1	78.6%	0.00	100.0%	64.9%	79.2%	83.5%
SER	1	2	29.4%	0.62	27.5%	0.0%	0.1%	0.3%
SER	1	3	20.1%	1.89	1.2%	0.0%	0.0%	0.0%
SER	1	4	16.5%	0.62	53.1%	0.0%	0.4%	0.8%
SER	2	1	79.5%	0.04	99.4%	35.1%	54.7%	65.5%
SER	2	2	22.4%	2.27	0.0%	0.0%	0.0%	0.0%
SER	2	3	14.7%	6.68	0.0%	0.0%	0.0%	0.0%
SER	2	4	10.5%	2.27	0.0%	0.0%	0.0%	0.0%
SER	3	1	94.6%	0.05	96.1%	38.0%	57.3%	65.0%
SER	3	2	61.8%	1.36	0.0%	0.0%	0.0%	0.0%
SER	3	3	55.4%	1.62	0.0%	0.0%	0.0%	0.0%
SER	3	4	57.7%	1.36	0.0%	0.0%	0.0%	0.0%
SER	4	1	100.0%	-0.01	99.9%	38.5%	62.2%	73.1%
SER	4	2	41.2%	8.35	0.0%	0.0%	0.0%	0.0%
SER	4	3	14.2%	31.12	0.0%	0.0%	0.0%	0.0%
SER	4	4	23.6%	8.35	0.0%	0.0%	0.0%	0.0%
UTL	1	1	-1.9%	-0.03	100.0%	92.0%	97.6%	98.9%
UTL	1	2	303.2%	0.19	0.0%	0.0%	0.0%	0.0%
UTL	1	3	778.2%	-0.17	0.0%	0.0%	0.0%	0.0%
UTL	1	4	303.2%	0.19	0.0%	0.0%	0.0%	0.0%
UTL	2	1	-14.1%	-0.32	89.1%	29.0%	45.1%	52.5%
UTL	2	2	-13.8%	-0.15	76.1%	17.7%	26.6%	32.4%

Sector	Industry	Product	Percent Positive	Estimate Rel. Bias	Coverage, *B*=5	Coverage, *B*=10	Coverage, *B*=15	Coverage, *B*=20
UTL	2	3	-16.8%	-0.15	62.6%	8.8%	15.9%	20.3%
UTL	2	4	-13.8%	-0.15	70.4%	13.9%	23.3%	27.7%
UTL	3	1	-5.5%	0.00	95.5%	37.5%	54.0%	62.3%
UTL	3	2	129.4%	-0.14	0.7%	0.0%	0.0%	0.0%
UTL	3	3	93.0%	0.04	0.5%	0.0%	0.0%	0.0%
UTL	3	4	129.4%	-0.14	0.4%	0.0%	0.0%	0.0%
UTL	4	1	2.5%	-0.07	99.8%	14.0%	42.4%	59.6%
UTL	4	2	2541.5%	-0.14	0.0%	0.0%	0.0%	0.0%
UTL	4	3	2662.4%	-0.16	0.0%	0.0%	0.0%	0.0%
UTL	4	4	2541.5%	-0.14	0.0%	0.0%	0.0%	0.0%
WHO	1	1	53.1%	-0.22	33.5%	2.5%	6.5%	9.0%
WHO	1	2	42.3%	0.56	11.7%	0.0%	0.1%	0.3%
WHO	1	3	29.9%	1.76	0.0%	0.0%	0.0%	0.0%
WHO	1	4	34.1%	0.56	9.6%	0.0%	0.1%	0.3%
WHO	2	1	58.4%	-0.38	23.6%	3.8%	6.5%	7.5%
WHO	2	2	38.8%	0.96	38.2%	4.5%	8.7%	10.8%
WHO	2	3	17.1%	1.48	25.8%	1.6%	3.4%	4.5%
WHO	2	4	32.4%	0.96	37.7%	4.5%	8.5%	10.7%
WHO	3	1	100.0%	0.05	99.3%	11.8%	31.6%	44.5%
WHO	3	2	24.6%	4.68	0.0%	0.0%	0.0%	0.0%
WHO	3	3	9.0%	12.71	0.0%	0.0%	0.0%	0.0%
WHO	3	4	7.9%	4.68	0.0%	0.0%	0.0%	0.0%
WHO	4	1	100.0%	-0.04	100.0%	86.5%	94.3%	96.5%
WHO	4	2	65.9%	1.71	4.7%	0.0%	0.0%	0.0%
WHO	4	3	6.8%	17.90	0.0%	0.0%	0.0%	0.0%
WHO	4	4	22.7%	1.71	0.0%	0.0%	0.0%	0.0%
WHO	5	1	100.0%	-0.13	91.0%	18.6%	37.4%	45.6%
WHO	5	2	60.2%	9.14	0.0%	0.0%	0.0%	0.0%
WHO	5	3	22.1%	79.24	0.0%	0.0%	0.0%	0.0%
WHO	5	4	20.7%	9.14	0.0%	0.0%	0.0%	0.0%