GOVERNMENT DIVISION REPORT SERIES
(Research Report #2011-1)

Bridge Study of the Imputation Methods for the Quarterly Survey of the Finances of Selected State and Local Government Employee Retirement Systems

Andrea Roberson

U.S. Census Bureau
Washington, DC 20233

CITATION: Roberson, Andrea. 2011. Bridge Study of the Imputation Methods for the Quarterly Survey of the Finances of Selected State and Local Government Employee Retirement Systems. Governments Division Report Series, Research Report #2011-1

Report Completed: June 29, 2011
Report Issued: June 30, 2011

Acknowledgments: The author would like to thank Lisa Blumerman, Kimberly Moore, Carma Hogue, Terri Craig, and Aneesah Williams for their comments on this paper.

Bridge Study of the Imputation Methods for the Quarterly Survey of the Finances of Selected State and Local Government Employee Retirement Systems

Andrea Roberson

U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Introduction

The quarterly survey of the Finances of Selected State and Local Government Employee Retirement Systems (QRET) provides a national summary of the revenues and expenditures of public employee retirement systems. As authorized by Title 13, United States Code, Section 182, the U.S. Census Bureau collects the data from the 100 largest retirement systems based on Total Cash and Security Holdings as identified in the most recent Census of Governments: Retirement Component.

In 2010, a study was conducted to review imputation procedures, and identify and implement a more robust imputation process. A bridge report is required once the desired imputation methodology is determined. This report contains the bridge between the old and new imputation methods from the second quarter of 2010 to the fourth quarter of 2010. We consider the impact of these modifications on the estimates.

The QRET Imputation Review began with an evaluation of the imputation processes that were being used at the time. Prior to the evaluation, the data elements in Part 1 of the survey were imputed using the most recent annual retirement survey data multiplied by a factor of .25. Data elements in Part 2 of the questionnaire were imputed as a percent of the most recent annual Z81 (total holdings and investments) multiplied by [the most recent annual Z81 + (Revenue – Expenses)]×a factor weight. The factor weight depended on what the current quarter was and when the system's fiscal year ended. It was, therefore, a top priority of the QRET team to develop more appropriate ways of mitigating nonresponse.

Methodology

A common preliminary step to imputation research is the formation of imputation cells for grouping like units. An attempt was made to ensure the units would be as similar as possible, and the imputation cells would be large enough to adequately impute within each cell. The Statistical Methods Branch (SMB) in the Governments Division quantified the strength of the association between the variables from each part of the survey and the summation (Z77+Z78). The sum of the variables Z77 and Z78 (Corporate Bonds and Corporate Stocks) was considered for groupings since the two items are inversely correlated and this sum over time is relatively stable. Correlation analysis was conducted with these variables. Correlations were calculated

using data from the third quarter of 2007 to the fourth quarter of 2009. SMB considered only those systems that had respondent data in the first part of the survey, and either respondent or compiled data in the second part of the survey. A number of items had a strong overall correlation with (Z77+Z78). The strongest was X11 (Benefits Paid). Three imputation cells were chosen that contained enough respondents to adequately impute for the nonrespondents in a cell. The data was stratified into three strata. The three strata arrived at were:

Stratum 1: Units where $\frac{Z77 + Z78}{Z81} \leq 40\%$

Stratum 2: Units where $40\% < \frac{Z77 + Z78}{Z81} \leq 60\%$

Stratum 3: Units where $\frac{Z77 + Z78}{Z81} > 60\%$

Various imputation methods were tested on datasets which contained simulated nonresponse in order to determine the best method of imputation for each variable in the questionnaire. SMB created a dataset of all original, reported data for 8 consecutive quarters, ending with 2010 quarter 3. From that base data set, 40 simulated data sets were created with simulated random nonresponse of 30 percent. Once the data sets were created, seven methods of imputing the missing data were applied. The imputation methods considered, along with the methodology selected for each variable, can be found in Attachment A. The methods were separated into two categories, those that required historical data and those that did not. Methods requiring historical data considered both the prior quarter and the same quarter prior year. Two measures of performance were calculated for each imputation method. The average mean deviation was used as a measure of the bias of the estimate, aggregated over the simulated data sets. The ratio of the sum of predicted values to the sum of the actual values measured the effect that the estimated values had on the final aggregated estimates. No single method was the single best performer across all measures of performance. SMB evaluated scenarios where the best method was the mean or median growth rate, and chose that as the best option. For methods where both the same quarter prior year and prior quarter had the same best method, SMB considered which method was better with the greatest frequency.

This bridge study highlights the differences between the new and old imputation methods. In this report, Table 1 shows Cash and Security Holdings of the retirement systems using the old and the new imputation methodology. Additionally, Table 2 provides Total Receipts data while Table 3 displays Total Payments data. In Tables 1, 2, and 3 the change and percentage change between the imputation methods for each variable are provided. Charts 1, 2, and 3 show the old

and new distributions for three quarters for Cash and Security Holdings, Total Receipts, and Total Payments with the new and old methods.

Discussion of Results

The data in Table 1 show that with the exception of Other Securities, all other holdings changed less than 1.0 percent when the old imputation methods were replaced with the new methods for the second quarter of 2010. For the third quarter of 2010 with the exception of Corporate Stocks and Other Securities, all other holdings changed less than 2.0 percent. For the fourth quarter of 2010, we have percent changes under 2.0 percent for all holdings. Other Securities demonstrated the largest percent change overall for quarter two. Corporate Stocks exhibited the largest percentage drop for quarter three. State and Local Securities however dropped only below 2.0 percent in the fourth quarter.

Table 2 shows the comparison of the old and new methods for Total Receipts. There were differences for all three quarters; the largest increase was in the second quarter. The results for Earnings on Investments showed a large increase in the first quarter but smaller increases in the third and fourth quarters. Employee Contributions showed the largest increase for the third quarter.

The percentage changes in Table 3 vary widely for each quarter. The percentage increase for Benefits in the second quarter was 2.6 percent. Total Payments and Benefits displayed increases of 8.8 and 9.0 percent, respectively in the third quarter.

The charts show that the distributions of Cash and Security Holdings are very similar with the exception of Corporate Stocks and Federal Government Securities. There is a substantial difference with the new method in quarter three for Corporate Stocks, and in quarter four for Federal Government Securities. Chart 2 illustrates how the Total Receipts distribution varies for the second quarter, with more similar results in the third and fourth quarters. Chart 3 shows that the distribution of Total Payments is nearly identical for both old and new methods.

Conclusions

The modification of imputation methodology has had an effect on the data. For Total Holdings the percent changes are all under 1 percent. The largest percent increase can be observed for Total Receipts, and there were substantial increases for Total Payments. As can be seen in the charts, the distributions of Holdings, Receipts, and Payments are mostly unchanged from the second quarter of 2010 to the fourth quarter. Holdings and Receipts each demonstrated a quarter of significant spikes in the distribution.

Table 1: Three Quarter Comparison of Old and New QRET Imputation Methods for Cash and Security Holdings
Source: U.S. Census Bureau, Finances of Selected State and Local Government Employee Retirement Systems. Data are not subject to sampling error, but are subject to nonsampling error.

	Second Quarter 2010				Third Quarter 2010			
	New	**Old**	**Difference**	**Percent Difference**	**New**	**Old**	**Difference**	**Percent Difference**
Total Holdings	2,359,053,094	2,359,624,245	-571,151	-0.02%	2,509,379,539	2,498,370,902	11,008,637	0.44%
Corporate Stocks	736,266,314	735,963,470	302,844	0.04%	801,189,569	2,496,483,382	-1,695,293,813	-67.91%
Corporate Bonds	415,070,977	415,040,051	30,926	0.01%	429,514,295	427,574,809	1,939,486	0.45%
Federal Government Securities	168,083,615	167,555,113	528,502	0.32%	165,458,128	164,049,683	1,408,445	0.86%
International Securities	398,705,752	398,745,191	-39,439	-0.01%	456,142,540	450,921,075	5,221,465	1.16%
Mortgages	10,203,118	10,251,432	-48,314	-0.47%	10,537,446	10,537,597	-151	0.00%
State & Local Securities	1,565,273	1,566,723	-1,450	-0.09%	1,833,207	1,834,638	-1,431	-0.08%
Cash & Deposits	83,037,717	83,668,871	-631,154	-0.75%	79,558,921	79,507,323	51,598	0.06%
Other Securities	121,082,106	112,832,291	8,249,815	7.31%	108,712,473	117,788,231	-9,075,758	-7.71%

Table 1: Three Quarter Comparison of Old and New QRET Imputation Methods for Cash and Security Holdings (Continued)

	Fourth Quarter 2010			
	New	**Old**	**Difference**	**Percent Difference**
Total Holdings	2,632,701,575	2,637,219,160	-4,517,585	-0.17%
Corporate Stocks	849,508,032	855,170,406	-5,662,374	-0.66%
Corporate Bonds	429,438,003	430,544,079	-1,106,076	-0.26%
Federal Government Securities	159,850,086	160,079,317	-229,231	-0.14%
International Securities	482,589,683	482,516,010	73,673	0.02%
Mortgages	9,925,474	9,880,835	44,639	0.45%
State & Local Securities	1,484,138	1,507,472	-23,334	-1.55%
Cash & Deposits	89,708,268	90,083,896	-375,628	-0.42%
Other Securities	109,396,053	110,036,545	-640,492	-0.58%

Table 2: Three Quarter Comparison of Old and New QRET Imputation Methods for Total Receipts

Source: U.S. Census Bureau, Finances of Selected State and Local Government Employee Retirement Systems. Data are not subject to sampling error, but are subject to nonsampling error.

	Second Quarter 2010				Third Quarter 2010			
	New	**Old**	**Difference**	**Percent Difference**	**New**	**Old**	**Difference**	**Percent Difference**
Total Receipts	-27,501,120	-24,565,079	-2,936,041	11.95%	160,423,931	154,644,330	5,779,601	3.74%
Government Contributions	17,845,361	17,658,534	186,827	1.06%	18,594,412	17,147,235	1,447,177	8.44%
Earnings on Investments	-55,775,273	-52,574,360	-3,200,913	6.09%	134,658,545	130,505,761	4,152,784	3.18%
Employee Contributions	10,428,792	10,350,747	78,045	0.75%	7,170,974	6,991,334	179,640	2.57%

	Fourth Quarter 2010			
	New	**Old**	**Difference**	**Percent Difference**
Total Receipts	156,801,355	156,032,682	768,673	0.49%
Government Contributions	21,612,448	21,563,738	48,710	0.23%
Earnings on Investments	125,589,315	124,944,684	644,631	0.52%
Employee Contributions	9,599,592	9,524,260	75,332	0.79%

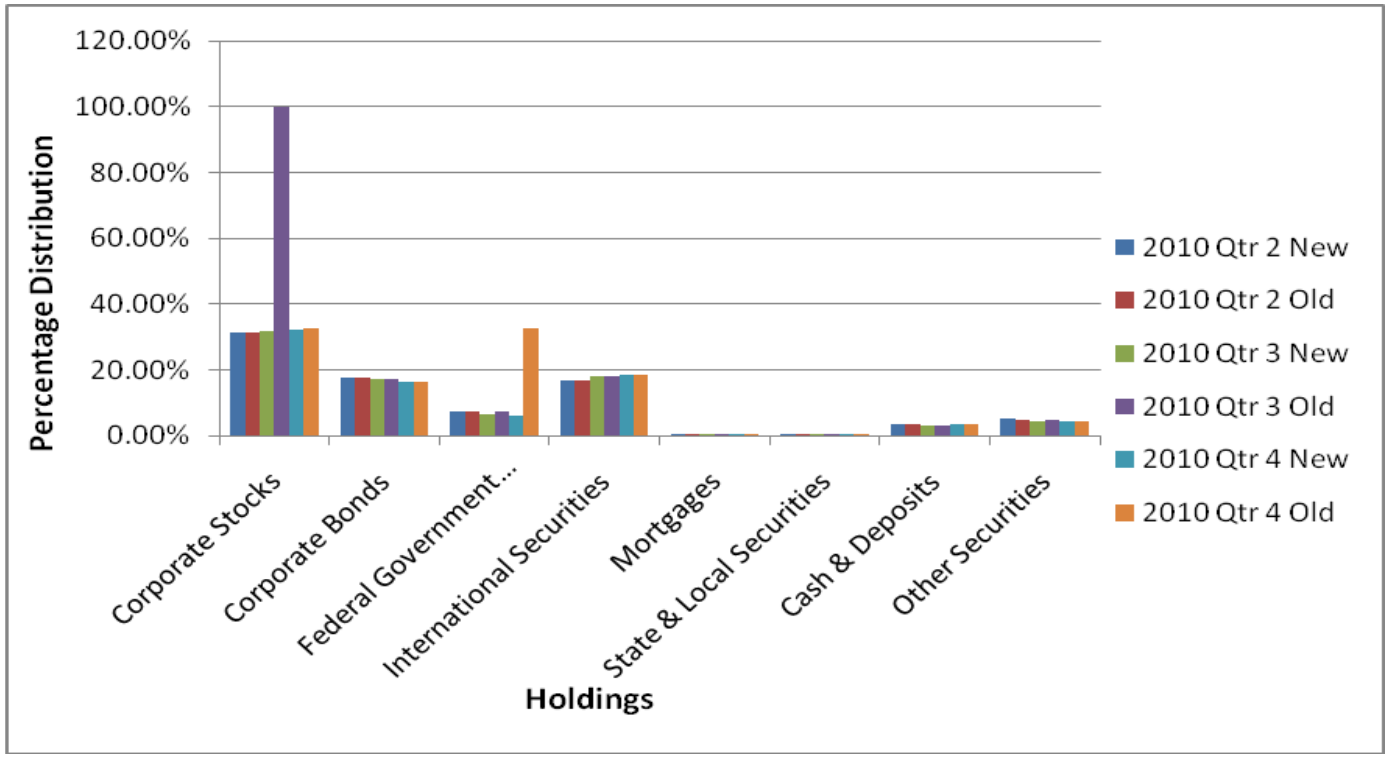
Table 3: Three Quarter Comparison of Old and New QRET Imputation Methods for Total Payments

Source: U.S. Census Bureau, Finances of Selected State and Local Government Employee Retirement Systems. Data are not subject to sampling error, but are subject to nonsampling error.

	Second Quarter 2010				Third Quarter 2010			
	New	**Old**	**Difference**	**Percent Difference**	**New**	**Old**	**Difference**	**Percent Difference**
Total Payments	47,875,748	46,680,004	1,195,744	2.56%	51,126,033	47,004,257	4,121,776	8.77%
Benefits	46,775,041	45,580,036	1,195,005	2.62%	49,994,085	45,878,672	4,115,413	8.97%
Withdrawals	1,100,707	1,099,968	739	0.07%	1,131,948	1,125,585	6,363	0.57%

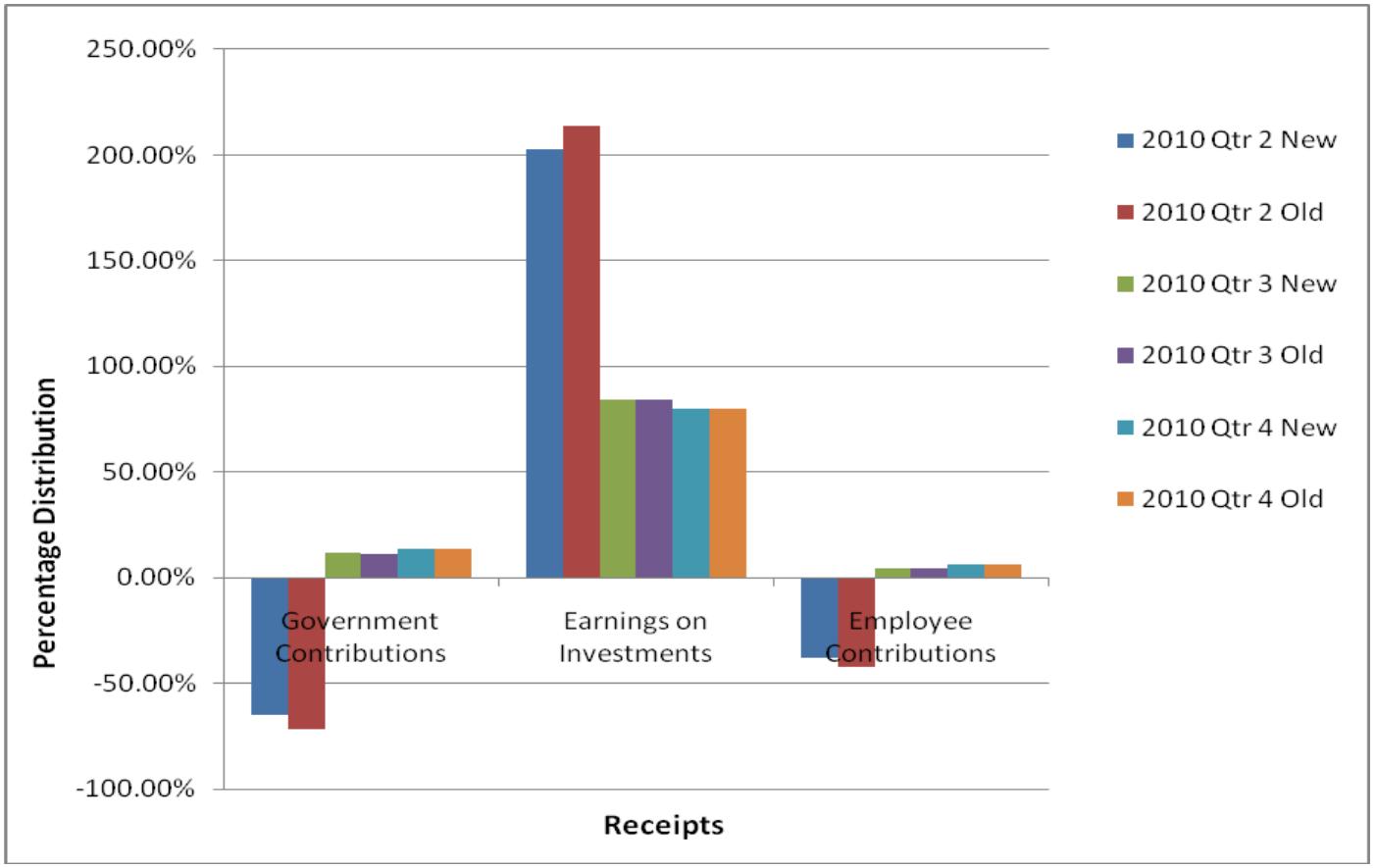
	Fourth Quarter 2010			
	New	**Old**	**Difference**	**Percent Difference**
Total Payments	50,370,913	50,396,341	-25,428	-0.05%
Benefits	49,227,196	49,250,311	-23,115	-0.05%
Withdrawals	1,143,717	1,146,030	-2,313	-0.20%

Chart 1: Cash & Security Holdings Distribution



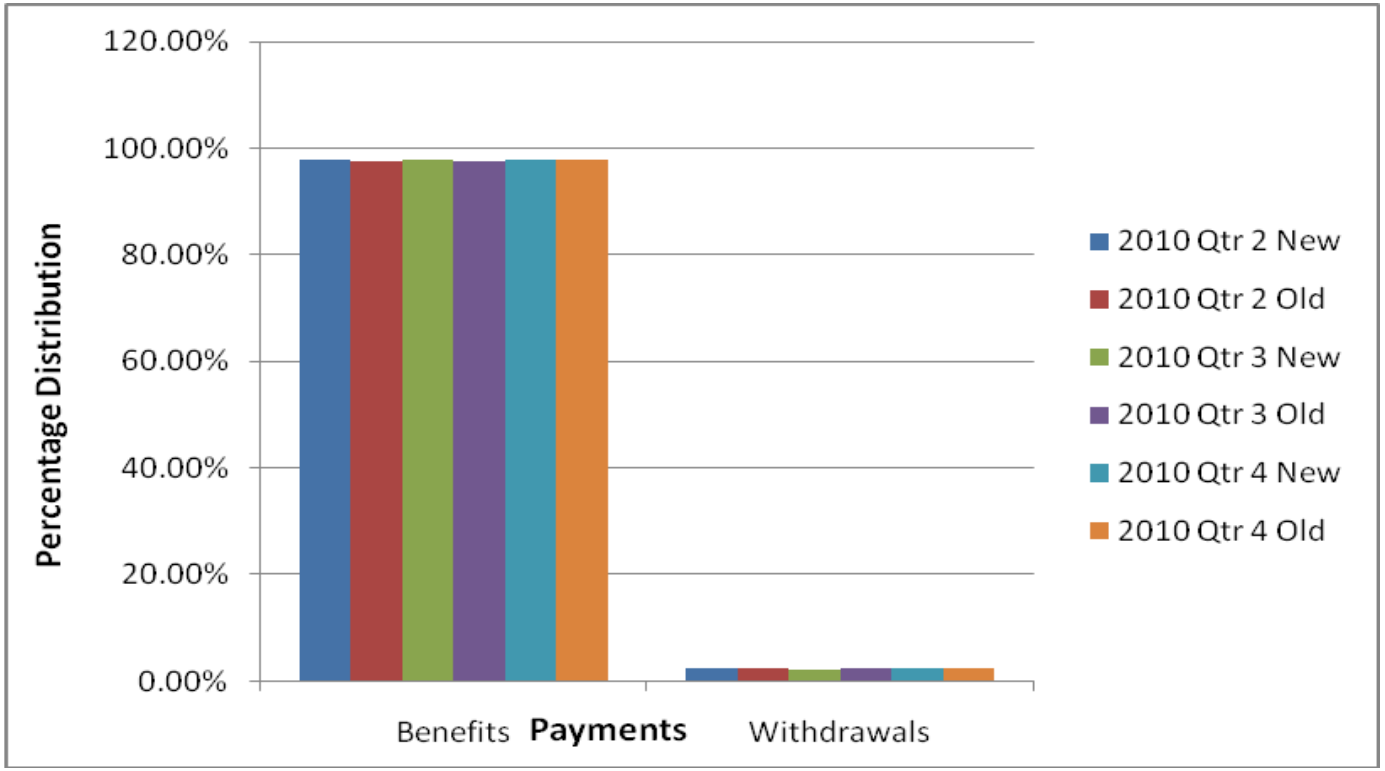
Source: U.S. Census Bureau, Finances of Selected State and Local Government Employee Retirement Systems

Chart 2: Total Receipts Distribution



Source: U.S. Census Bureau, Finances of Selected State and Local Government Employee Retirement Systems

Chart 3: Total Payments Distribution



Source: U.S. Census Bureau, Finances of Selected State and Local Government Employee Retirement Systems

Attachment A

Imputation Methods

Method 1: Pull forward prior year (quarter) data:

$$y_i = y_p$$

where y_i represents the variable to be imputed for the i^{th} nonrespondent, and y_p represents the reported prior year (quarter) value for nonrespondent, *i*.

Method 2: Prior year (quarter) data with a cell mean growth rate:

$$y_{h,i,t} = y_{i,t-1} * \left(\frac{\sum_{j=1}^{n_{hR}} \left(\frac{y_{h,j,t} - y_{h,j,t-1}}{y_{h,j,t-1}} \right)}{n_{hR}} \right)$$

where the expression in parentheses is the mean growth rate of respondents in an imputation cell *h*; n_{hR} denotes total number of respondents R in cell *h*; *i* represents the i^{th} nonrespondent in imputation cell *h*; *j* represents the j^{th} respondent in imputation cell *h*; and *t* is the current year (quarter).

Method 3: Prior year (quarter) data with a cell median growth rate:

$$y_{h,i,t} = y_{i,t-1} * \left(\sum_{j=1}^{n_{hR}} \left(\frac{y_{h,j,t} - y_{h,j,t-1}}{y_{h,j,t-1}} \right) \right)_{Median}$$

where the expression in parentheses is the median growth rate of respondents in an imputation cell *h*; i represents the i^{th} nonrespondent in imputation cell *h*; *j* represents the j^{th} respondent in imputation cell *h*; and *t* is the current year (quarter).

Method 4: The cell mean:

$$y_{h,i} = \bar{y}_{h,R}$$

where *h* represents an imputation cell and R denotes the group of respondents in the cell.

Method 5: The cell median:

$$y_{h,i} = y_{h,Median,R}$$

where *Median* is calculated from all respondents R in a cell *h*.

Method 6: An adjusted cell mean:

$$y_{h,i} = \frac{Z81_{h,i}}{\overline{Z81}_h} * \overline{y}_{h,R}$$

where *h* represents an imputation cell; *R* denotes the group of respondents in an imputation cell; *i* represents the i^{th} nonrespondent in imputation cell *h*; and *z81* is Total Holdings.

Method 7: An adjusted cell median:

$$y_{h,i} = \frac{Z81_{h,i}}{Z81_{h,Median}} * y_{h,Median,R}$$

where *h* represents an imputation cell; *R* denotes the group of respondents in an imputation cell; *i* represents the i^{th} nonrespondent in imputation cell *h*; and *Z81* is Total Holdings.

New Imputation Methods by QRET Variable

Imputation Methodology			
Prior Quarter		Same Quarter Prior Year	
Mean Growth Rate	Median Growth Rate	Mean Growth Rate	Median Growth Rate
Z91	X01, X04, Z72, X12, Z93, X21, X30, Z70, X42, Z83_Z84, Z82, Z96	X05	Z71, X11, X35, Z81