COMPENSATING FOR MISSING WAVE DATA IN THE SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)

Todd R. Williams and Leroy Bailey, Bureau of the Census*
Todd R. Williams, Washington, D.C. 20233

Key Words : Imputation, Missing wave, Carryover

I. Introduction

The Survey of Income and Program Participation (SIPP) is a national longitudinal survey, conducted by the U.S. Bureau of the Census, which collects detailed information on income and wealth. The information collected includes employment income and income received from government transfer programs at the person, family, and household levels. The survey uses a rotating panel design, with a new panel of sample households being introduced at the start of each calendar year. Each panel is divided into four approximately equal rotation groups. Each month households from a different rotation group are interviewed. During each interview, the respondent is asked to provide information for the preceding four months. The four-month cycle, in which all of the households of the panel are interviewed, is referred to as a wave. The number of waves in a panel is determined by the length of the panel. Through 1995, households in most of the SIPP panels had been interviewed once every four months over a period of 32 months. Starting with the 1996 panel, the length of the panel will increase to four years with a new panel being introduced once the old panel is completed.

A major problem in obtaining accurate estimates of income and program participation from the SIPP is nonresponse. In some cases, a person or an entire household does not respond for one or more interviews during the length of the panel, resulting in one or more waves of missing data. These persons are referred to as panel nonrespondents. Longitudinal weighting procedures adjust for this nonresponse by assigning a zero weight to the nonrespondents and multiplying the weights of the persons interviewed during the entire period by a nonresponse adjustment factor. Because of this adjustment, the information for panel nonrespondents from waves for which they are interviewed is not used in the estimates. In an effort to include more of the available data in the estimates, a longitudinal imputation procedure is performed for some of the missing data, thereby changing panel nonrespondents to respondents before the longitudinal weighting adjustment takes place. For each nonrespondent, this procedure imputes data only for missing waves bounded on both sides by an interviewed wave. This procedure does not impute data for two or more consecutive missing waves or for the first or last wave of the panel.

This research expands on the work done at the Census Bureau by Antoinette Tremblay (Tremblay 1994), where she compared alternative longitudinal imputation methods for single missing waves. We evaluated the same four longitudinal imputation methods, but added imputed wage and salary, social security and Aid to Families with Dependent Children (AFDC) amounts to the imputed food stamp amounts evaluated by Tremblay.

II. Imputation Methods

The current longitudinal imputation procedure is referred to as the **random carryover** method. This method (Census 1994) imputes each nonrespondent's missing data for waves that are preceded and followed by interviewed waves. A value r is randomly assigned to each nonrespondent's household for each missing wave, where $r = 0,1,2,3,\text{or } 4$. The first r reference months within the missing wave receive their imputed amounts from the last reference month of the preceding wave and the remaining $4\text{-}r$ reference months receive their imputed amounts from the first reference month of the subsequent wave. A major advantage of this method is that it is simple to implement in terms of computer programming and execution time, thus making it the easiest procedure to automate. In addition, the procedure produces data conducive to multiple analytic purposes. On the other hand, random carryover forces stability in responses for wave nonrespondents, and could therefore lead to the underestimation of between wave changes.

A variation of the random carryover method is referred to as the **population carryover** method. Like the random carryover method, this method (Tremblay 1994) takes the imputed amounts for the reference months of the missing wave from the last reference month of the preceding interviewed wave and the first reference month of the subsequent interviewed wave. Unlike the random carryover method, the interviewed reference month used to donate the imputed amount is determined by a probability mass function defined by the probabilities associated with patterns found in the interviewed population. The patterns are defined by the occurrences of change (difference greater than zero) in the amounts between months within the wave. The interviewed reference month used to donate the imputed amount alternates between months when a change occurs. The advantage of this method over the current method is that the imputed data will more accurately reflect the patterns of within wave changes in amounts found in the interviewed data.

An expansion of Tremblay's research is to include imputation for missing first or last waves of the panel when using the two carryover methods. For this study, the method for doing this is simply to use the first reference month of the second wave to supply information to the reference months of the first wave when it is missing. Similarly, the last reference month of the preceding wave is used to supply information to the reference months of the last wave when it is missing. Since information is being supplied by only one interviewed reference month, the procedure for imputing these missing waves is identical for the two methods.

The third longitudinal imputation method, developed by Little and Su (Little 1989) and referred to as the **Little & Su** method, uses a multiplicative model based on row (person) and column (period) effects to determine the imputed amounts. The model is of the form

imputation = (row effect) x (column effect) x (residual).

When imputing for the i^{th} nonrespondent, the imputed amount $\hat{a}_{ij}$ for month j is

$$\hat{a}_{ij} = [r_i][c_j][a_{kj} / (r_k c_j)] = r_i \, a_{kj} / r_k ,$$

and the row effect r_i and the column effect c_j are

$$r_i = \frac{1}{m_R} \sum_{h=1}^{m_R} \left(\frac{a_{ih}}{c_h} \right) , \qquad c_j = \frac{m\bar{a}_j}{\sum_{h=1}^{m} \bar{a}_h} ,$$

where m_R is the number of interviewed reference months, m is the total number of reference months, and $\bar{a}_j$ is the mean amount for month j over all interviewed persons. The a_{kj} and the r_k in the residual effect of the initial equation are the donor amount and row effect of interviewed person k, whose row effect is closest in value to the row effect of the nonrespondent. A possible advantage in using this method is that information about both trend and individual levels is incorporated into the imputed amounts. Moreover, Little and Su argue that the procedure does not require separate modeling for different missing data patterns, and that it is comparatively easy to implement.

The fourth longitudinal imputation method finds a set of matching variables which are used to match the nonrespondent to an interviewed person and is referred to as the **flexible matching** method. This method (Census 1995) uses a forward stepwise multivariate linear regression procedure to determine the set of matching variables associated with the missing amounts for each wave. The matching variables are ranked by order of importance and a match is attempted between the nonrespondent and an interviewed person using all of the matching variables. If a match is not found, the least important matching variable is dropped and a match is attempted on the remaining variables. This procedure is continued until a match is found or all matching variables are dropped. A user-determined mechanism is built into the procedure to ensure that a match is found for all nonrespondents. Once a match is found, the nonrespondent's missing amounts are replaced with the corresponding amounts of the matched interviewed person. One possible advantage is that this method may provide a vehicle to more accurately impute amounts for persons who display uncommon variability, by obtaining imputed amounts from interviewed person with like characteristics.

III. Empirical Methodology

This research used data from the SIPP 1992 panel. To simplify the analysis, only the first four waves of the ten wave panel were examined, with the fourth wave being treated as the last wave. A person level research file was created for each of the items being imputed using interviewed persons and nonrespondents who were 15 years old or older at the beginning of the panel and received the item in at least one of the sixteen reference months. Since single wave imputation was being examined, only those nonrespondents who were missing data in one wave were on the files.

For the current procedure, the random selection of the donor reference month applies only to nonrespondents who do not move during the noninterview wave. For movers, the selection of which reference month to use for the imputed amount is based on when the move takes place. Because the intention of this research was to compare the other three longitudinal imputation methods with the random selection of the random carryover method, movers were removed from the research file.

To examine the accuracy of the imputation methods, data files containing simulated missing data were created. To create the data files we initially regressed both household and person response status onto selected survey variables for each item being studied using the data from the research file. For household response, the selected survey variables were taken from the household reference person.

From the research file, frequency distributions of five response patterns were developed for each level of the resulting indicator variables. The five possible response patterns were defined as 1) response in all four waves, 2) nonresponse in wave one, 3) nonresponse in wave two, 4) nonresponse in wave three, and 5) nonresponse in wave four. Since only those nonrespondents who were missing data in a single wave were kept on the files, any person who did not respond in one wave had to respond in the remaining waves. Using only the persons who responded in all four waves, a value was randomly generated for each household using the frequency distribution determined by the level of the predictor variable of the household reference person. If the value represented a noninterviewed wave, the amounts for the four reference months within the wave were coded to missing for each person in the household. Once the household nonresponse simulation was performed, person nonresponse simulation was done in the same manner for each individual person in the remaining interviewed households. This procedure was performed ten separate times in order to produce ten data files, for each item being studied, containing simulated missing data.

Once the simulated missing data files were created, the data were imputed using each imputation method. For the flexible matching method, the user defines the set of variables to be considered for matching. In this study, the set of variables included person and income characteristics. It also included the amount from the last reference month of the previous wave and the amount from the first reference month of the subsequent wave for the item being imputed. These amounts were determined by the procedure to be the most important matching variables.

IV. Evaluation Methodology

The original data file used for this research contains monthly estimates for food stamps, AFDC, wage and salary and social security income. For each item we computed wave means, the average of the estimate over the four reference months in the wave, for each person. These computations were made for the original data file and for the ten simulated data sets, after the application of the four imputation procedures. Secondly the ten imputed data sets were aggregated to form a single set of "combined measures" for the selected survey items. Therefore, for each item and imputation alternative, we derived one value per wave for each unit of analysis.

As was the case with Tremblay's research, our principal evaluation criteria for the alternative longitudinal imputation methodologies were estimates of the accuracy of the imputations derived from differences between the actual and imputed data. In addition, we compared values for selected descriptive statistics and measures of total error.

Three of the four imputation alternatives entail cross-wave imputation or assumptions about cross-wave relationships. Therefore the utility of these procedures is affected by the magnitude of the inter-wave correlation for the various survey items. In order to assess the relationship between waves, and facilitate the analysis of its effect on the quality of the imputation, between-wave correlation coefficients were computed for the four items under study, for both the actual values from the original data set and the imputed data sets.

We also computed means over the units of analysis and standard deviations of the sample units' wave means for the four survey items used for the study. The means for the individual imputation methods were obtained by dividing the applicable combined total for the specific item by the combined number of observations for which there were data entries for the item. To estimate the bias of the imputation procedures, we computed the means of the respective differences between item imputes and the corresponding actual data values. In addition, we derived corresponding estimates of relative bias and the mean absolute deviation between the imputed and the actual data. Finally, we compared estimates by individual reference month between the imputed and actual data by comparing the average absolute deviations between the reference month estimates and their corresponding wave means.

V. Results

Table 1 provides between-wave correlation coefficients for wage and salary, social security, food stamps, and AFDC, respectively, for the actual and imputed data sets. For each between-wave analysis, only those persons who have imputed data for one of the applicable waves that is being analyzed are used in the calculations. The between wave correlation measure for waves 1 and 2 actually reflects the correlation between months 4 and 5. Similarly the correlation estimate for waves 2 and 3 and for waves 3 and 4 reflect the correlation between months 8 and 9 and between months 12 and 13, respectively. The bold table entries indicate which correlation coefficients derived from the imputed data are closest to the actual correlation coefficients.

There seems to be considerable dispersion in the correlation coefficients across "wave pairs", ranging from moderate to high values. The results for the actual data are most varied, while, as expected, the coefficients for the carryover procedures are consistently high and the most stable. In general, the Little and Su method does the best job in maintaining the between-wave correlation exhibited by the actual data. In comparing the between-wave correlation coefficients for the actual and the imputed data, the correlations between waves 1 and 2 are the most variable.

Averages and standard deviations of the wave means, for those persons for whom imputes were derived, are presented in Table 2. Bold table entries indicate which imputation method produces the smallest estimate of relative bias. The check marks indicate significant deviations at the 10% level when testing the hypothesis that the population mean of the original or actual data is the same as the population mean after the application of each imputation method. For wage and salary and social security, mean estimates for the imputed data are reasonably close to the actual data . This occurs for the overall estimates as well as the individual wave estimates.

The magnitude of the overall estimates of relative bias ranges from 0.06 to 1.38 percent, while the corresponding range of the magnitude for the wave estimates is from 0.00 to 5.10 percent. Similarly the overall estimates for Food Stamps and AFDC were relatively close to the actual data; however, there is considerable variation in the wave estimates as a result of the small number of observation applicable to the given category. Note that for each item several of the table entries indicate a significant difference. For the carryover procedures, significant differences are found for wave 1 for wage and salary, over all waves and wave 2 for social security, and for wave 1 for AFDC. For the Little and Su procedure significant differences occur for wave 1 for the wage and salary, social security and food stamps items. Significant differences are detected for the flexible matching approach only for the overall and wave 1 cells for the food stamps item.

An assessment of the total error of the four imputation alternatives selected for this research is provided by Figure 1 which gives measures of the average absolute deviation for wage and salary and social security amounts. The food stamp and AFDC amounts showed similar results. The value of each measure is listed at the end of the horizontal bar and the standard errors are listed in the right margin. Initially we note that for each survey item there is considerable variation in the evaluation measures over the four waves, especially with the Little and Su and Flexible Matching procedures. While the relative bias estimates for these methods, which are operationally more complex, compared favorably with the carryover procedures, relative to total error, the results of this table suggest that they are less desirable. We also note that the quality of the imputation for wave 1 is generally worse than that of waves 2 and 3. Moreover, the quality of the wave 4 imputation tends to be closer to that of wave 1.

Figure 2 displays the average absolute deviation between the estimate for each reference month and the corresponding wave mean for the actual and imputed values of wage and salary and social security amounts. Similar results are seen with the food stamp and AFDC amounts. Those reference months that are in the same wave are connected by a line. Because the two carryover procedures obtained their imputed values from only one reference month for the first and last waves, the imputed values for months 1 through 4 are identical and the imputed values for months 13 through 16 are identical; therefore, the deviation for these months is zero. This is in large contrast to the deviations shown in the actual data and highlights the problem associated with imputing data for the four reference months within a noninterviewed wave using only one reference month as the donor. It also appears that the current random carryover procedure is, by far, the worst of the imputation methods in regards to accurately reflecting the actual monthly deviations. In general, the Little & Su and the flexible matching procedures produced monthly deviations closer to those displayed by the actual data.

VI. Summary and Recommendations

An assessment of the performance of the current SIPP wave nonresponse imputation procedure and three alternatives was conducted using data from the first four waves of the 1992 panel. This study represented an extension of a previous evaluation in which the effectiveness of the methods in estimating amounts for food stamps recipiency was evaluated. It included, in part, an investigation of two issues cited for future research in the previous evaluation - a plausible modification of the current procedure, and an examination of the effects of the alternative procedures on an expanded set of survey items.

We used several measures of data association and quality to evaluate the procedures that were considered for the research. As was the case with Tremblay's research, our work did not identify a uniformly "best" imputation procedure for compensating for wave nonresponse. Relative to total error, the data from the first four waves of the 1992 panel would clearly favor the carryover procedures. The Little and Su method and the flexible matching performed somewhat better than the carryover procedures in maintaining cross-wave relationships and imputation for which relative large changes have occurred between waves. In addition we obtained reasonably good estimates of bias for these procedures. However, the additional computational burden and the relative size of their total error are unfavorable aspects of the Little and Su and flexible matching methods in comparison to the carryover approaches.

The current carryover method performed compared favorably with the other imputation alternatives for the items selected for this study, as did the population carryover method. Lepkowski (1989) indicated that when the amount of wave nonresponse is substantial, imputation procedures such as the carryover method have some advantages. However, the extent to which it can lead to an attenuation in longitudinal relationship remains a matter of some concern. Its total error measures, simplicity, and flexibility are certainly among the desirable features of a compensatory procedure designed for a complex and multipurpose survey like the SIPP.

It is recommended that for SIPP we continue to pursue a combined (imputation and weighting) compensation strategy for wave nonresponse, while attempting to ensure that estimates of transition and longitudinal cumulation are not hampered by the inappropriate application of the carryover method to nonconducive nonresponse patterns. The population carryover method has shown some promise in its ability to compensate for some of the deficiencies of the random carryover. As we move further into the implementation of the new SIPP design, the problems associated with wave nonresponse may generate greater interest, since the duration of the panels will be extended. Therefore it is important that we continue to pursue plausible approaches to reduce nonresponse in the survey and to compensate effectively for its associated biases.

REFERENCES

Lepkowski, J.M. (1989), "Treatment of Wave Nonresponse in Panel Surveys," in *Panel Surveys* (D. Kasprzyk, G. Duncan, G. Kalton, M. P. Singh, ed.) pp. 348-374.

Little, R. J. A., and Su, Hong-Lin (1989), "Item Nonresponse in Panel Surveys," in *Panel Surveys* (D. Kasprzyk, G. Duncan, G. Kalton, M. P. Singh, ed.) pp. 400-425.

Tremblay, A. (1994), "Longitudinal Imputation of SIPP Food Stamp Benefits," in *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 809-814.

U.S. Bureau of the Census (1994), "SIPP 1990+: Specifications for Panel File Longitudinal Imputation - REVISION," memorandum from Preston Jay Waite to Sherry Courtland.

U.S. Bureau of the Census (1995), "Methodology Used for the Modeling of Missing Variables in the Flexible Matching Imputation Procedure," memorandum from Todd R. Williams.

U.S. Bureau of the Census (1995), "Using the Flexible Matching Imputation Software," memorandum from Todd R. Williams.

* This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.

Table 1. Between-Wave Correlation Coefficients

	Wage & salary	Social security	Food stamps	AFDC
waves 1 & 2				
Actual	0.66	0.84	0.71	0.88
Carryover-random	0.97	0.99	0.96	0.99
Carryover-pop.	0.88	0.91	0.87	0.80
Little & Su	**0.77**	**0.84**	**0.71**	0.77
Flexible matching	0.83	0.85	0.81	**0.86**
waves 2 & 3				
Actual	0.83	0.91	0.82	0.82
Carryover-random	0.97	0.96	0.94	0.94
Carryover-pop.	0.88	0.93	0.84	0.92
Little & Su	0.85	**0.91**	**0.81**	**0.84**
Flexible matching	**0.83**	0.93	0.83	0.79
waves 3 & 4				
Actual	0.86	0.93	0.86	0.87
Carryover-random	0.98	0.98	0.97	0.97
Carryover-pop.	0.97	0.97	0.96	**0.96**
Little & Su	**0.83**	0.89	0.73	0.72
Flexible matching	0.90	**0.91**	**0.76**	0.76

Figure 1. Average Absolute Deviation of the Imputed Amounts from the Actual Amounts ($)

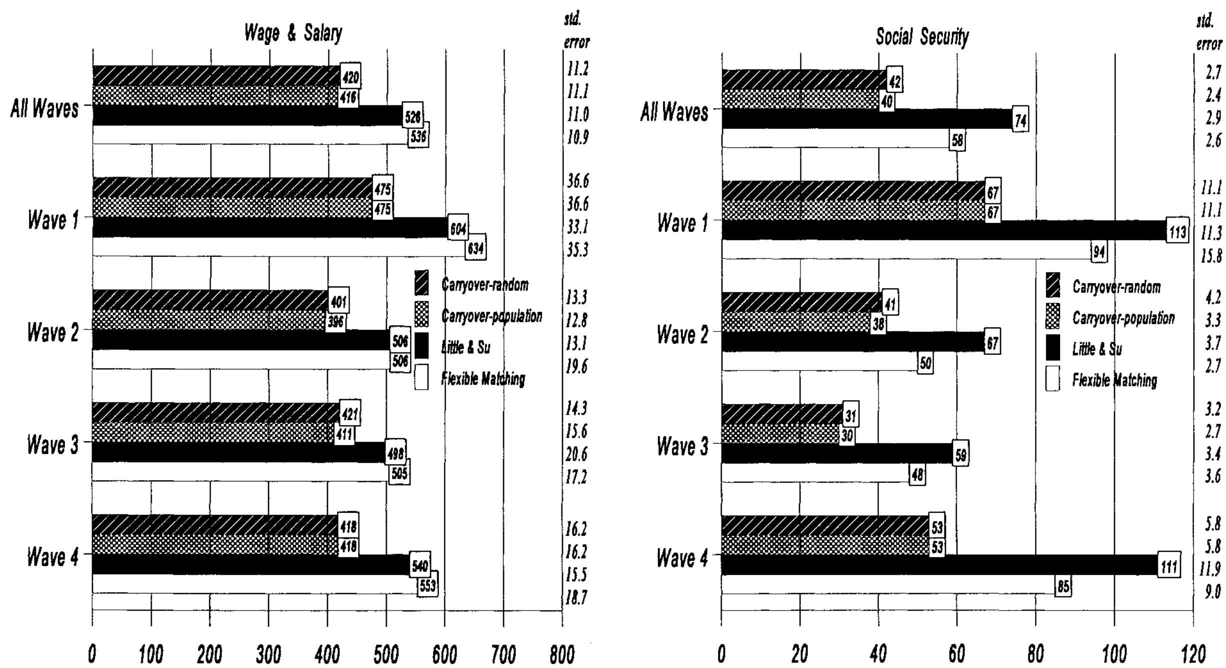

Table 2. Means and Standard Deviations for Imputed Units ($)

√ denotes significant difference between actual and imputed at the 10% level.		Wage & salary					Social security				
		Over all waves	Wave 1	Wave 2	Wave 3	Wave 4	Over all waves	Wave 1	Wave 2	Wave 3	Wave 4
Actual	Mean	1724.27	1876.18	1729.29	1694.78	1664.16	546.16	518.02	545.13	563.87	526.01
	St. Dev.	1651.42	1783.34	1657.94	1602.59	1609.51	277.50	342.52	270.03	264.31	279.55
Carryover-random	Mean	1722.89	√1793.56	1725.28	1716.23	1688.40	√553.72	**531.25**	√559.06	567.78	516.32
	St. Dev.	1730.44	1893.31	1693.00	1655.76	1758.86	282.40	**346.87**	271.21	269.24	296.24
	Rel. Bias	-0.08%	-4.40%	-0.23%	1.27%	1.46%	1.38%	**2.55%**	2.56%	0.69%	-1.84%
Carryover-population	Mean	1726.79	√1793.56	1731.59	√1722.96	1688.40	√550.75	**531.25**	√555.67	**563.36**	516.32
	St. Dev.	1730.90	1893.31	1694.72	1655.10	1758.86	285.23	**346.87**	278.48	**268.51**	296.24
	Rel. Bias	0.15%	-4.40%	0.13%	1.66%	1.46%	0.84%	**2.55%**	1.93%	**-0.09%**	-1.84%
Little & Su	Mean	1712.75	√1798.28	1732.27	1693.05	**1658.45**	**546.80**	√544.42	539.44	567.69	521.39
	St. Dev.	1690.96	1726.86	1752.88	1613.33	**1651.48**	**281.28**	328.52	269.46	267.75	312.94
	Rel. Bias	-0.67%	-4.15%	0.17%	-0.10%	**-0.34%**	**0.12%**	5.10%	-1.04%	0.68%	-0.88%
Flexible Matching	Mean	**1725.33**	**1876.22**	**1728.68**	**1693.39**	1670.32	549.13	536.80	**545.58**	565.57	**529.09**
	St. Dev.	**1791.44**	**1938.03**	**1817.12**	**1619.60**	1820.96	286.12	358.94	**270.96**	274.49	**303.72**
	Rel. Bias	**0.06%**	**0.00%**	**-0.04%**	**-0.08%**	0.37%	0.54%	3.63%	**0.08%**	0.30%	**0.58%**
		Food stamps					AFDC				
Actual	Mean	115.43	86.08	114.66	119.49	132.46	292.16	236.53	316.37	259.96	296.35
	St. Dev.	122.82	121.42	121.35	122.15	123.91	252.74	218.05	245.69	268.70	258.84
Carryover-random	Mean	119.64	**102.03**	115.52	√132.78	125.19	287.58	√274.07	**319.90**	236.62	**281.56**
	St. Dev.	126.32	**127.21**	122.83	135.79	120.60	247.06	207.74	**238.40**	237.34	**272.59**
	Rel. Bias	3.65%	**18.53%**	0.75%	11.12%	-5.49%	-1.57%	15.87%	**1.11%**	-8.98%	**-4.99%**
Carryover-population	Mean	**118.74**	**102.03**	116.28	**√127.62**	125.19	**293.57**	√274.07	323.89	**256.82**	**281.56**
	St. Dev.	**125.64**	**127.21**	125.16	**129.94**	120.60	**257.37**	207.74	257.25	**252.24**	**272.59**
	Rel. Bias	**2.86%**	**18.53%**	1.42%	**6.80%**	-5.49%	**0.48%**	15.87%	2.38%	**-1.21%**	**-4.99%**
Little & Su	Mean	120.04	√104.48	**114.23**	√128.25	**131.63**	303.42	298.28	323.61	247.89	316.53
	St. Dev.	122.41	118.65	**118.28**	124.99	**127.76**	249.54	247.26	245.04	221.32	273.45
	Rel. Bias	3.99%	21.38%	**-0.37%**	7.33%	**-0.62%**	3.85%	26.11%	2.29%	-4.64%	6.81%
Flexible Matching	Mean	√121.29	√109.44	110.98	128.58	138.32	296.86	**268.34**	321.19	284.60	278.01
	St. Dev.	129.29	126.00	119.62	148.28	124.76	258.70	**209.74**	251.24	278.84	267.53
	Rel. Bias	5.07%	27.13%	-3.21%	7.61%	4.42%	1.61%	**13.45%**	1.52%	9.48%	-6.19%

Figure 2. Average Absolute Deviations between Individual Months and the Wave Means ($)