

The TIAA-CREF/Global Financial Systems Project Survey

Contents of Survey Package and Condition of Provision

The materials provided with this survey include the following items:

- 1) Overview, general discussion, and selected tabulations from the survey. The information and tabulations in this section are intended to be used by researchers to verify that the survey data has been loaded into their software packages correctly.
- 2) Additional memorandum on methodology and weighting procedures used in the survey.
- 3) A twelve-page "codebook" describing the contents and coding of the electronic data files.
- 4) A copy of the survey questionnaire.
- 5) Two 3.5" IBM formatted floppy disks, one containing the "Premium-Paying" Panel survey data (GFS-PREM.DAT) and one containing the "Retired" Panel survey data (GFS-RET.DAT). The data are stored as tab-separated ASCII files. The first line of each file contains the variable names.

TIAA-CREF is happy to provide these data to researchers; however we have an important concern regarding the use of the data. Weights are included on the data file so that researchers may use the survey data to make inferences about the underlying TIAA-CREF population. We recognize that many researchers are not necessarily interested in the underlying TIAA-CREF population as such and would rather use the survey data to make inferences about the entire U.S. population as a whole, for example. While the researchers themselves will undoubtedly be mindful of the fact that the sample is not representative of the TIAA-CREF population -- or any other population other than the population of those who responded to the survey -- we are concerned that the consumers of research are often not as careful or discriminating in their interpretation of research results and statistics.

We therefore request that researchers using this data make an explicit and prominent statement that the sample is not representative of the TIAA-CREF population (unless the sample weights provided with the data are used throughout the analysis of the data).

We also request an opportunity to review and comment upon any articles or research results obtained using this data.

This data is provided to researchers on the condition that they comply with these requests.

TIAA-CREF/Global Financial Systems Survey

Overview

This document contains a summary and general description of the TIAA-CREF/Global Financial Systems ("GFS") Survey. Researchers may find some of the tabulations and discussion contained here to be helpful as they attempt to use the data provided.

This overview is divided into six sections. The first is a discussion of the background of the TIAA-CREF/GFS Survey and some summary information regarding the population from which the survey sample was drawn. The next section provides some summary data on the sample itself and discusses the content of the survey. The third, fourth, and fifth sections discuss the details of the major weakness of this survey: the amount of missing data. The last section concludes.

I. Description of Survey Composition

The TIAA-CREF/GFS Survey is composed of two parts: The Premium-Paying Survey and the Retiree Survey. The electronic data files containing each of these two parts are included with this documentation, each on a separate floppy disk. The basic distinction between the two samples is that the premium-paying sample members were, at the time of the survey, remitting contributions to their TIAA-CREF deferred annuity contracts, while members of the retiree sample had begun to draw annuity income from at least one of their TIAA-CREF contracts. Both groups were asked to respond to essentially the same survey questions.

It is important to note that these two sub-samples of individuals, which compose the TIAA-CREF/GFS Survey, were actually both drawn from a larger sample of TIAA-CREF participants, called the TIAA-CREF Research Panel (described below). As a result, there are multiple layers of sample selection that each observation in the enclosed data sets has passed through: All respondents (1) had to be chosen as prospective participants for the TIAA-CREF Research Panel Project (2) had to elect to participate in the Research Panel Project, (3) had to be chosen to participate in the TIAA-CREF/GFS Survey, and finally (4) had to elect to participate in the TIAA-CREF/GFS Survey.

The next section briefly describes some of the details of the Research Panel Project which may be of relevance to users of the TIAA-CREF/GFS Survey.

The TIAA-CREF Research Panel

The TIAA-CREF Research Panel was created in 1993 and was originally composed of three randomly selected and representative samples of TIAA-CREF participants:

1. Premium-Paying Panel - A sample of individuals currently making regular contributions to their deferred annuity contracts. (30,600 surveys mailed, 4,851

- completed = 15.9% Response Rate)
2. Paid-Up Panel - A sample of individuals who are no longer regularly contributing to their deferred annuities.¹ (19,200 surveys mailed, 2,678 completed = 13.9% Response Rate)
 3. Retired Panel - A sample of TIAA-CREF retirees, i.e. individuals drawing some annuity income from TIAA-CREF. (10,200 mailed, 2,318 completed = 22.7% Response Rate)

The members of the 1993 Research Panel constitute the “population” from which the TIAA-CREF/GFS Survey sample was drawn; i.e. the 1995/96 TIAA-CREF/GFS Survey was mailed to members of the Premium-Paying and Retired Research Panel who completed the initial 1993 Research Panel questionnaire, and who were still active participants at the time of the 1995/96 survey.

It is important to note that although the selection of the individuals to whom the original 1993 Research Panel questionnaires was mailed was random, the group of individuals who responded to the 1993 Research Panel questionnaire do not constitute a random sample of the TIAA-CREF population. In fact, the respondents to the initial 1993 Research Panel survey were much more likely to be male and were generally older than was the case for the relevant TIAA-CREF populations (across all three of the sub-panels listed above).

The point here is simply that the 1995/96 TIAA-CREF/GFS Survey sample was drawn from a population that is not representative of the TIAA-CREF population.

II. The TIAA-CREF/GFS Survey

The TIAA-CREF/GFS Survey was undertaken by TIAA-CREF in cooperation with professors Dwight B. Crane² and Zvi Bodie³ (both affiliated with Harvard University's Global Financial Systems project) in an effort to learn more about the financial structure of household asset portfolios, focusing on the allocation of assets held both inside and outside tax-privileged retirement accounts.

As described in the preceding section, the TIAA-CREF/GFS Survey was sent to a subset of Research Panel members in the winter of 1995-1996.⁴ 3,457 Premium-Paying Panel members were mailed the TIAA-CREF/GFS questionnaires; of these 974 were completed, yielding a response rate of around 28%. 1,200 Retiree Panel members were mailed questionnaires; of these 519 were completed, yielding a response rate of around 43%. Processing of the returned questionnaires and preliminary tabulations were done by Matthew Greenwald & Associates, Inc.

¹ A person is considered paid-up if at least six premium cycles (usually six months, but sometimes longer) have passed since TIAA-CREF has received a premium remission to be applied on any of that person's deferred annuity contracts.

² Gund Professor of Banking, Harvard University Graduate School of Business.

³ Professor of Finance, Boston University School of Management.

⁴ Roughly 1/5 of the surveys were sent out in December of 1995, the rest were mailed in February of 1996.

The survey is a household survey in as much as it requests respondents to provide data on financial assets owned by any members of their household. A copy of the actual questionnaire is included in this documentation package; it is the best source of information on how survey questions were asked.

The available information on age and sex of the TIAA-CREF/GFS Survey sample is presented in the table that follows:⁵

Table 1
Age and Sex Composition of the TIAA-CREF/GFS Survey Samples

<u>Count</u>	Premium Paying		Retired	
		<u>Percent</u>	<u>Count</u>	<u>Percent</u>
Gender				
Men	606	63%	346	67%
Women	360	37%	160	31%
Not Indicated	8	1%	13	3%
Age				
Under 35	57	5%	0	0%
35-44	208	21%	0	0%
45-54	337	35%	0	0%
55-64	301	31%	41	8%
65-70	49	5%	203	39%
71-75	7	1%	226	44%
76+	0	0%	33	6%
Not Indicated	15	2%	16	3%
Total	974	100%	519	100%

Age in the table above is as of 1995 birthdays; i.e. age = 1995 - year of birth as reported in the survey.

As the table shows, there are individuals in the data sets that did not report an age or sex. Rather than deleting these observations, we have left them in the data sets and encourage the researcher to make up his or her own mind about whether to include these cases in any analysis.

The respondents to the 1995/96 TIAA-CREF/GFS Survey were more likely to be both

⁵ Two records in the premium paying panel have year-of-birth responses that would make them over 100 years old as of 1996, and one has a year of birth that would make the respondent under 15. These three observations are treated as missing in the table.

male and older⁶ than the respondents to the 1993 TIAA-CREF Research Panel. And as was noted above, the TIAA-CREF Research Panel respondents were in turn older and more likely to be male than the TIAA-CREF population as a whole. The point here is that by the time we end up with TIAA-CREF/GFS Survey respondents, we have gotten quite far away from typical TIAA-CREF participants (who are actually slightly more likely to be female than male) at least in terms of age and sex, and quite possibly in terms of much more. Caution in the interpretation of the responses of the sample members is obviously in order.

III. Dealing with Missing Data

Perhaps the greatest weakness of the data -- and a reason for great care in the treatment of the survey responses -- is the amount of data that is missing in the survey due to incomplete responses. As can readily be seen from the survey questionnaire, most of the questions in the survey were two part questions. The first part asked the respondent to indicate if the household had or did not have a specific type of liability or asset. The second part asked the respondent to indicate the value of the asset or liability.

The completeness of responses to these questions fell into four general classes:

1. Respondent left the entire section blank; not indicating ownership or non-ownership and not indicating a value of the asset or liability.
2. Respondent indicated ownership only; he or she left the value part of the question blank.
3. Respondent indicated a value only; he or she left the ownership section blank
4. Respondent answered both parts of the question.

In cases 1 and 2, it is not necessarily clear how to interpret the information on the questionnaire.⁷ For example, some researchers may want to assume that in case 1, the respondent does not own the asset or liability in question. Some researchers may not want to make this assumption.

Given that this is the case, and that TIAA-CREF/GFS did not undertake a major imputation exercise, we have elected to allow the researcher to decide what to do about missing data by simply attempting to make clear what parts of the answers were reported on the survey and what was left blank or missing. The missing data is generally treated according to one rule: if the data was not reported on the questionnaire, it is simply recorded as missing in the computer files. The only exception to that rule are cases which fall into category 3 above, in which the yes/no part of the question has automatically been re-coded as yes.

⁶ The increase in age is of course partly a function of the panel nature of the questionnaire; i.e. because it is a subset of the 1993 Research Panel, the entire 1995/96 TIAA-CREF/GFS Survey sample is necessarily roughly 2 years older than the 1993 Research Panel. This aging effect does not explain all of the difference in the age distribution, however.

⁷ Case 4 does not present a problem, and it seems obvious that in case 3, the yes/no part of the questionnaire should be coded with a yes if the amount reported is non-zero, and a no otherwise.

The following tables summarize the missing data problem:

Table 2
Non-Response in the TIAA-CREF/GFS Survey
Premium-Paying Participant Sample

Survey Question Number and Asset Type		Yes (N)	No (N)	No Response (N)	Total Valid Value Responses*	
					(N)	(%)
3	Has Savings Account	678	279	17	638	94%
4	Has Checking	907	45	22	843	93%
5	Has CD's	292	658	24	249	85%
6	Has Other Stock Funds	433	490	51	375	87%
6	Has Tax-Free Bonds	163	620	191	139	85%
6	Has Other Bonds	100	638	236	83	83%
6	Has Other Funds	74	628	272	59	80%
7	Has Money Market	296	644	34	258	87%
8	Has Traded Stock	410	540	24	366	89%
9	Has U.S. Savings Bonds	324	586	64	294	91%
9	Has Other U.S.G. Bonds	68	802	104	56	82%
9	Has Mort. Backed Bonds	30	825	119	23	77%
9	Has State/Muni/Tax-Free Bonds	108	762	104	93	86%
9	Has Corp/Foreign Bonds	37	808	129	33	89%
10	Has Whole Life Insurance	473	471	30	392	83%
12/13	Owns Home	869	87	18	842	97%
14	Has Home Mortgage Loan	678	206	90	640	94%
15	Owns Non-Resid. Real Estate	359	569	46	338	94%
16	Has Other Mortgage Loan	149	617	208	132	89%
17	Has Auto Loan	294	670	10	255	87%
18	Has Educational Loan	101	847	26	89	88%
19	Has Other Personal Loan	142	797	35	115	81%
20	Carries a Credit Card Balance	355	615	4	331	93%

*The count in this column is the number of non-missing responses to the second part of the two part question that inquired about a given asset or liability. The base of the percentage reported is the count of "yes" responses. For example, the first row shows that 94% of those who said they had savings accounts also reported a value for that/those account(s).

Table 3
Missing Data in the TIAA-CREF/GFS Survey
Retiree Participant Sample

Survey Question Number and Asset Type		Yes	No	No	Total Valid Value	
		(N)	(N)	Response (N)	Responses* (N)	(%)
3	Has Savings Account	287	215	17	254	89%
4	Has Checking	479	26	14	446	93%
5	Has CD's	247	257	15	210	85%
6	Has Other Stock Funds	252	204	63	206	82%
6	Has Tax-Free Bonds	157	232	130	130	83%
6	Has Other Bonds	100	255	164	76	76%
6	Has Other Funds	75	244	200	57	76%
7	Has Money Market	156	338	25	136	87%
8	Has Traded Stock	253	245	21	223	88%
9	Has U.S. Savings Bonds	110	333	76	93	85%
9	Has Other U.S.G. Bonds	72	351	96	66	92%
9	Has Mort. Backed Bonds	35	372	112	27	77%
9	Has State/Muni/Tax-Free Bonds	127	309	83	110	87%
9	Has Corp/Foreign Bonds	41	361	117	31	76%
10	Has Whole Life Insurance	277	222	20	233	84%
12/13	Owns Home	468	35	16	446	95%
14	Has Home Mortgage Loan	138	331	50	122	88%
15	Owns Non-Resid. Real Estate	206	284	29	192	93%
16	Has Other Mortgage Loan	44	401	74	40	91%
17	Has Auto Loan	71	441	7	53	75%
18	Has Educational Loan	10	493	16	4	40%
19	Has Other Personal Loan	23	480	16	14	61%
20	Carries a Credit Card Balance	119	394	6	106	89%

*The count in this column is the number of non-missing responses to the second part of the two part question that inquired about a given asset or liability. The base of the percentage reported is the count of "yes" responses. For example, the first row shows that 89% of those who said they had savings accounts also reported a value for that/those account(s).

The two tables show that there are two significant types of non-response: First, there were significant numbers of sample members who left entire sections of the questionnaire blank (counts of these individuals are in the “no response” column). However, there are also significant numbers of individuals who reported ownership of an asset or liability, but then failed to report the value of the asset or liability.

IV. Data in the "Retirement Savings Plans and Accounts" Section

The first section of the questionnaire presents some unique difficulties. In this section individuals were asked:

Please indicate which types of retirement savings plans you, or anyone in your household, currently have. For each type your household has, please indicate the amount invested in each type of fund within those accounts, as well as the total amount your household has invested in that type of account.

Spaces were available for the individual to report a yes or no ownership response, a total amount invested, and an amount in each of three asset classes: Cash Funds, and Fixed Income Funds, and Stock Funds.

As with the rest of the survey, many respondents did not fill out this section fully. Also, some respondents reported total amounts that were not equal to the sum of the amounts in each of the three asset classes.

Here we believed that the missing data and incomplete responses required some corrective recoding. The following rules were adopted:

- In Question 1 (retirement plan assets), several respondents indicated they have a particular type of plan and filled in a dollar amount for at least one type of investment within that plan, but left blank the amounts under all other types of plan investments. In those cases, the blank columns were changed to reflect an investment value of \$0.
- Note also that in the data files, the original total invested amounts have all been replaced by totals calculated from the sum of reported components. (Provided that the components were reported. If components were not reported and the total value was reported, that total has been unaltered. All of the original total value responses have been preserved in the data files as additional variables.)
- Also, in order to provide the most accurate responses, the total value and allocation of the TIAA-CREF retirement accounts owned (item 1e) by the respondent were substituted for the values provided by the respondent. (As mentioned above, the original responses have been preserved on the data files in variables that appear at the end of the data sets. See the attached codebook.)

Table 4
Non-Response in the TIAA-CREF/GFS Survey
Premium-Paying Sample Retirement Accounts Section

Account Type	Has Account	Does Not Have Account	No Response	Did Not Indicate Value	Did Not Indicate Allocation
IRA	587	219	168	65	107
Keogh	54	511	409	13	15
401(k)	171	429	374	29	34
403(b)	252	361	361	46	60
TIAA-CREF	931	12	31	21*	21*
TDAs	249	365	360	40	57
Thrift	104	432	438	20	33

*These cells indicate the number of individuals who did not provide the information required to retrieve value and allocation of amounts in TIAA-CREF accounts from data in TIAA-CREF accounting databases.

Table 5
Non-Response in the TIAA-CREF/GFS Survey
Retiree Sample Retirement Accounts Section

Account Type	Has Account	Does Not Have Account	No Response	Did Not Indicate Value	Did Not Indicate Allocation
IRA	340	70	109	54	76
Keogh	22	183	314	2	3
401(k)	28	174	317	11	13
403(b)	36	168	315	13	13
TIAA-CREF	413	26	80	5*	5*
TDAs	163	114	242	44	68
Thrift	61	144	314	13	20

*These cells indicate the number of individuals who did not provide the information required to retrieve value and allocation of amounts in TIAA-CREF accounts from data in TIAA-CREF accounting databases.

In the above tables, “No Response” means that the individual did not answer the yes/no question as to whether they had the type of account in question (and also left the amounts section blank, as anyone who reported an amount was imputed to own the asset). It is probably true that most individuals who did not fill in this yes/no part of the questionnaire do not have the type of account in question, but there is obviously no way of knowing whether they do or do not. In addition to this problem, a large proportion of the individuals who responded that they did in fact own the type of account in question did not indicate the value of the account or the asset allocations within the account (for example, the table above shows that 32% of those with "Other Thrift Plans" did not indicate an asset allocation for these accounts). Again, this poses significant problems for interpretation of the data.

V. Problems of Aggregation

Ultimately, users of this data may be most interested in generating aggregates from the data (i.e. measures of total indebtedness, total assets, or total net worth). There is some question as to whether such aggregates can be accurately constructed from the data due to the many holes in the database.

For example, despite the fact that upwards of 80% (roughly) of those who claimed to have any individual asset reported a valid value for that asset (see the table above), once one tries to add up across the assets, the occurrence of at least one missing asset becomes more likely. Thus the number of missing aggregate totals is substantially higher. In fact, when (1) all completely blank sections of questionnaires are treated as indicating that the household does not have the asset inquired about in the section -- i.e. treating completely blank sections as indicating no ownership, as opposed to missing data, (2) retirement asset data is considered missing when any individual allocation amounts are not supplied in cases where ownership is indicated, and (3) non-retirement asset data and liability data is considered missing only when someone indicates ownership but does not supply an amount, we have been able to calculate total wealth for only 609 of the 974 sample members. Under the same approach, total net worth (total wealth minus total liabilities) is calculable for only 566 individuals. This emphasizes the need for great care in aggregation of variables in which there are substantial lacunae, and great caution in interpreting the results of aggregates over these amounts.

VI. Conclusion

We have spent many hours looking for problems in the data, and have tried to correct as many errors and resolve as many inconsistencies as possible. There is no doubt that problems and inconsistencies remain in the data, and we urge users of the data to contact TIAA-CREF when such problems are discovered. Of all of the difficulties, the problems relating to sample selection and missing data/non-response strike us as most severe; it is our hope that all users of the data will understand and develop strategies to deal with these problems in the course of their analyses. Despite the problems, we believe that this survey data is unique, and we hope that it will help to stimulate greater study of issues of retirement and wealth accumulation.

We are glad to provide you with this data, and ask in return to be provided the opportunity to comment on draft papers and to receive copies of any publication in which data from the TIAA-CREF/GFS dataset are used.