

Vanguard Research Initiative
December 2014

4. Response Analysis

1. Response rates

Table 4-1A gives the response rate by age and client type for the production survey. The response rates are calculated as ratio between the number of completes and non-bounced emailed invitations. The overall response rate 7.38% is very close to 7.7% from the pilot survey which we used to determine the overall size of invitations for production survey. The response rate does not have a monotonic pattern in age. The response rate for the employer sample is noticeably lower than for the individual client sample.

Table 4-1B shows the response rate by assets in the Vanguard administrative data. Interestingly, the response rate increases with wealth for both client types with the effect much more powerful for the individual account respondents. This higher response rate for the high-wealth individuals is consistent with Vanguard's experience in earlier survey efforts. It runs against the conventional wisdom that high-wealth individuals are hard to survey both because they are hard to reach and because they have complex finances. Evidently, engagement with Vanguard overcomes these issues.

Table 4-2 shows the distribution of administrative assets conditional on criteria used for being invited to participate in the survey and conditional on responding. To be Web-survey eligible, the account holder must be Web registered, have logged on in the past six months, not opted out of being contacted, and not have been contacted recently (see Vanguard Research Initiative: Documentation and Supporting Analysis, "2. Sample Design"). Survey eligible account holders have higher wealth than the broader Vanguard population. Given Web-survey eligibility, most accounts satisfy our asset cut-off (\$10K), so the additional effect of the

>\$10,000 condition on mean assets is minimal. We stratified the sample in various ways, i.e., by line of business, age, and single status (see “2. Sample Design,” Table 2-2). The consequence of this stratification—shown in the “invited” row—reduces assets, mainly in the upper tail.

Administratively-single account holders also tend to have lower Vanguard assets.

The last line shows the asset distribution conditional on completing the survey. Higher wealth account holders are more likely to respond, so the completed sample has higher wealth both than the universe and the invited account holders. To further explore selectivity in response, Table 4-3 reports the estimates from the linear probability model where the dependent variable is one for completing the survey and zero otherwise. The independent variables include various demographic information for the account holders from the administrative data. The result reconfirms all the findings from the previous tables. The wealth variable has the largest quantitative effect: an account holder from the lowest quintile is 4.4 percentage points less likely to complete the survey compared to one from the highest quintile. The oldest group is about 1.5 percentage points less likely to complete the survey, while among the remaining age groups the likelihood is flat. Employer-sponsored sample is 1.5 percentage points less likely to complete the survey compared to individual client sample.

2. Effect of reminder.

To improve response rates, we sent a reminder two weeks after the initial invitation. Figure 4-1 shows the response rate over time.¹ The vertical line indicates the date of the reminder. Before the reminder date the increase in the response rate slowed noticeably. The reminder leads to a

¹ To avoid having Vanguard’s help desk overwhelmed with questions, the initial invitations were sent to two groups. Group A were invited on August 15 and Group B on August 20 and 22. The figure is for Group A.

very large jump in the response rate of about 2 percentage points (from 5.3% to 7.3%). In terms of completed surveys, the reminder yields about 2,400 more observations.

3. Break-off analysis.

Some of the non-completion comes from invitees who start the survey, but break off at some point. Table 4-4 shows the number of respondents who completed the survey up to a selection of questions. After the initial large drop prior to the first question (age), that is respondents who clicked the survey link but did not participate at all, there are significant drops are observed in the income section (between Q8 and Q45) and also in the financial wealth section (between Q56 and Q65). These sections are indeed the most challenging ones on the survey. Especially for the financial wealth section, respondents face a battery of account-level questions. The wealth battery is the heart of the survey. It takes the most time,² it has the most complex structure, and is perhaps the most intrusive set of questions. Within the wealth modules, we strove for a high level of item response by asking again for any response left blank and only leaving implicit that questions could be skipped after a second item-non-response. This strategy yielded very low item non-response. For example, among those who said they have IRA accounts, only 0.2% of them did not give any information on the balance or answered that the balance is \$1, the minimum value allowed to type in. Hence, for the final 8,950 sample, the quality of the data is extremely high.³ But the survey's insistence on responses for each item might well account for the high level of breakoffs in this section.

Table 4-5 shows the proportions of groups of respondents at various points in the survey. By sample design, the shares of the age groups shown in the first row of Panel A are roughly

² The median length of the survey is approximately 40 minutes. The median time spent on the financial wealth section is about a third of it.

³ In the HRS or SCF, imputations are necessary to get better picture of overall wealth since the fraction of non-response or bracketed answers is not negligible.

equal.⁴ Panel A shows that older groups are more likely to accept the invitation, but have higher with-in survey attrition. The net effect is that the completion rate for the survey declines with age. Panel B does the same analysis across wealth groups. Wealthier households are more likely to respond to the invitation and also more likely to complete the survey when they have started it. Given the conventional wisdom that wealthier households are hard to survey, this finding may be surprising, though it is consistent with what Vanguard has found in previous surveys. The higher-asset individuals are likely more engaged with Vanguard and are potentially more interested in the topic of the survey. Panel C shows the break-offs by for the employer-sponsored versus individual client sectors. There are no noticeable patterns by sector.

⁴ The shares of each aged group invited are exactly equal to 0.2. The shares of invitations received are slightly different owing to different rates of bounced emails.

Table 4-1. Response Rates

A. Age and Client Type

Age	All	Client Type	
		Employer-sponsored	Individual client
All	7.38%	6.64%	7.67%
55-59	6.93%	7.09%	6.76%
60-64	7.27%	6.49%	8.09%
65-69	7.85%	6.24%	8.34%
70-74	8.08%	6.30%	8.27%
75-	6.80%	5.37%	6.84%

B. Administrative Assets and Client Type

Asset quintiles	Total	Client Type	
		Employer-sponsored	Individual client
All	7.38%	6.64%	7.67%
1st (Lowest)	5.80%	5.96%	5.74%
2 nd	6.24%	5.88%	6.36%
3 rd	6.86%	6.11%	7.16%
4 th	7.85%	6.93%	8.28%
5th (Highest)	10.23%	8.06%	11.14%

Note: Denominator of response rate is non-bounced invitations. 8950 observations (production sample).

Table 4-2. Administration Assets: Population to Sample

	Mean	Percentiles				
		10	25	50	75	90
Total Universe	272,264	6,892	24,570	86,611	261,995	635,017
Survey eligibility	309,715	10,366	35,672	117,501	319,447	688,349
Asset cut-off	287,873	19,972	42,727	114,455	307,339	700,283
Survey eligibility and Asset cut-off	310,967	23,311	53,470	142,518	352,837	727,007
Invited	277,172	22,868	52,150	138,327	331,887	655,287
Completed	355,329	27,151	68,174	188,191	445,497	826,482

Note: Age cut-off (higher than or equal to 55 and less than 100) is applied to all the rows

Table 4-3. Completed Survey (linear probability model)

Variable/Category	coefficient	standard error
Intercept	0.090	0.004
Asset quintile (1st, lowest)	-0.044	0.002
Asset quintile (2nd)	-0.040	0.002
Asset quintile (3rd)	-0.033	0.002
Asset quintile (4th)	-0.023	0.002
Sector - Employer-sponsored	-0.015	0.002
Age (55-59)	0.012	0.003
Age (60-64)	0.015	0.003
Age (65-69)	0.016	0.002
Age (70-74)	0.015	0.002
Admin non-single	0.002	0.002
Gender – F	0.000	0.003
Gender – M	0.006	0.003

Note: N = 121,201. Dependent variable is dummy variable for completion (1=completed). Omitted categories for independent variables are the highest asset quintile, individual client sector, the oldest, admin single and unknown gender.

Table 4-4. Progress through survey
(Respondents reaching various questions)

Question	Respondents
Start	19,477
Age (Q1)	14,091
Income (Q8)	13,829
Housing wealth (Q45)	11,495
Wealth (Q56)	11,438
Other assets (Q65)	9,163
Demographic (Q77)	9,090
End	8,950

Table 4-5. Break-off analysis

A. Age

Question	Age				
	55-59	60-64	65-69	70-74	75-
Invitation received	0.20	0.20	0.20	0.20	0.20
Start point	0.16	0.18	0.21	0.23	0.22
Age (Q1)	0.18	0.19	0.20	0.22	0.21
Income (Q8)	0.18	0.19	0.20	0.22	0.21
Housing wealth (Q45)	0.18	0.19	0.21	0.22	0.20
Wealth (Q56)	0.18	0.19	0.21	0.22	0.20
Other assets (Q65)	0.19	0.20	0.21	0.22	0.18
Demographic (Q77)	0.19	0.20	0.21	0.22	0.18
End	0.19	0.20	0.21	0.22	0.18

Note: Table show fraction of each age group (shares by row) at various points in the survey.

Table 4-5. Break-off Analysis (continued).

B. Administrative Assets

Question	Asset quintile				
	1st	2 nd	3 rd	4th	5th
Invitation received	0.20	0.20	0.20	0.20	0.19
Start point	0.17	0.18	0.20	0.22	0.24
Age (Q1)	0.16	0.17	0.19	0.22	0.25
Income (Q8)	0.16	0.17	0.19	0.22	0.25
Housing wealth (Q45)	0.16	0.17	0.19	0.22	0.26
Wealth (Q56)	0.16	0.17	0.19	0.22	0.26
Other assets (Q65)	0.16	0.17	0.19	0.22	0.27
Demographic (Q77)	0.16	0.17	0.19	0.22	0.27
End	0.15	0.17	0.19	0.22	0.27

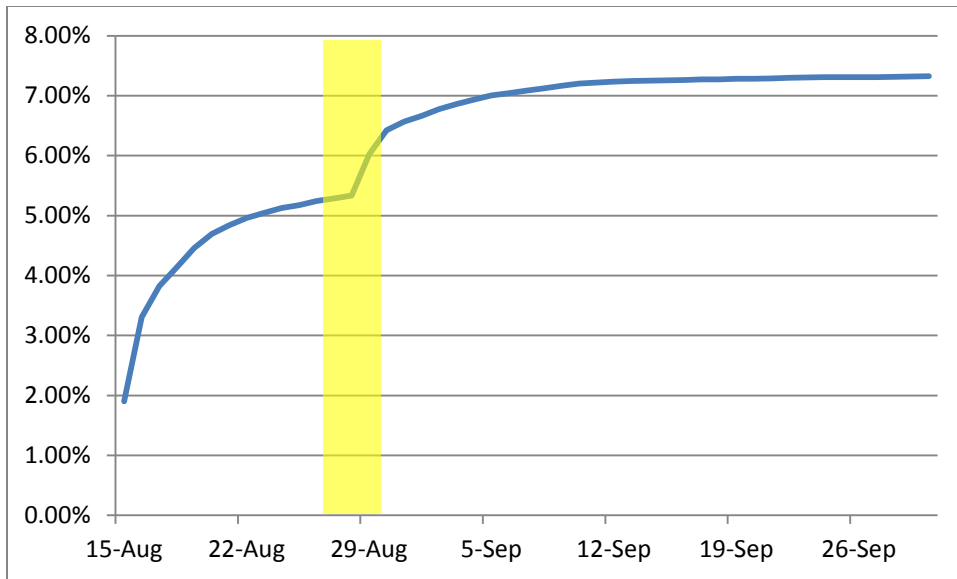
Note: Table show fraction of each wealth group (shares by row) at various points in the survey.

C. Client type

Question	Employer-sponsored	Individual client
Invitation received	0.28	0.72
Start point	0.24	0.76
Age (Q1)	0.24	0.76
Income (Q8)	0.24	0.76
Housing wealth (Q45)	0.24	0.76
Wealth (Q56)	0.24	0.76
Other assets (Q65)	0.25	0.75
Demographic (Q77)	0.25	0.75
End	0.25	0.75

Note: Table show fraction of each client type (shares by row) at various points in the survey.

Figure 4-1. Response rate



Note: Response rate of group A over time. The yellow vertical line indicates the date of the reminder.