

Consumer Data Source Evaluation

The Performance Evaluation of:

Access Sphinx

March 25, 2009

By

Mktg.
Incorporated



Harvesting Quality Data...since 1979

TABLE OF CONTENTS

	Page
1. Data Source Metrics.....	3
2. Demographic Evaluation.....	5
2.1. Age Distribution.....	5
2.2. Income Distribution.....	6
2.3. Education Distribution.....	7
2.4. Martial Status Distribution.....	8
3. Structural Segment Distributions	9
3.1. Buyer Behavior Segments.....	10
3.2. Sociographic Segments.....	12
3.3. Media Segments.....	14
4. Source Performance	16
4.1. Performance	17
4.2. Characteristics (Frequent Survey Takers).....	18
4.3. Deviate Behavior (Speeders and Straight-Liners)	20

1. DATA SOURCE METRICS

Performance metrics focus on the structure, characteristics and performance of panels and sources of data. However, to be effective, the metrics must be standardized and independent. While there is always some degree of arbitrariness in the choice of metrics, their structure can not be arbitrary.

Introduction to Procedures

Metrics need to be standardized in such a way that analysis by any party would produce the same results. Furthermore, the procedures should be such that multiple executions will produce statistically common results. That is, that the error will always be within statistical limits.

Independence of the metrics is a key principle in their development. It makes little sense to have standards that are, in fact, tied in principle to specific types of data, methods, and instruments. An alternative way of looking at the issue of independence is that of objectivity. That is, that all usable standardized metrics must be objectively obtained. We believe that requires independence in several ways.

- ***Source*** – Metrics must be independent of the panels and sources. They have to be applicable across all or at least most sources of data. Furthermore, the metrics can not depend on the nature of the sources or in their construction.
- ***Methods*** – The methods for constructing the metrics must be independent. We are looking for different measures; performance measures can not be structurally interconnected.
- ***Instrument*** – The principles of metrics should be extendable to any number of compatible instruments (questionnaires).
- ***Evaluator*** - In order to assure objectivity and independence of metrics, it is critical that the evaluator needs to be also independent. This is required both as a means to assure objectivity and as a means of certifying that independence.

Methods Used

The testing procedures were based on the execution of a standard questionnaire. The questions were selected to allow a consistent standard and independent assessment of the panel or data source. In all cases, samples were made by the source supplier using their standard methods of management and incentives to reflect current operations.

For this assessment 400 completes were used in the test based on the execution of a standardized online (internet form) questionnaire. This questionnaire is available in multiple local languages for global execution. However, only one language is used in each test of a source to maintain consistency.

Evaluation of Access Sphinx Data Source

Typically for “random” and non-structured convenience sampling some form of quota corrections are used to help balance the resulting sample. These quotas may reflect the specific needs of a study, regional or preference requirements or to balance demographics. In the latter instance the objective is to correct for inherent incongruities against some standard, often the general population. In order to correctly test the data source, quotas were requested based on demographics.

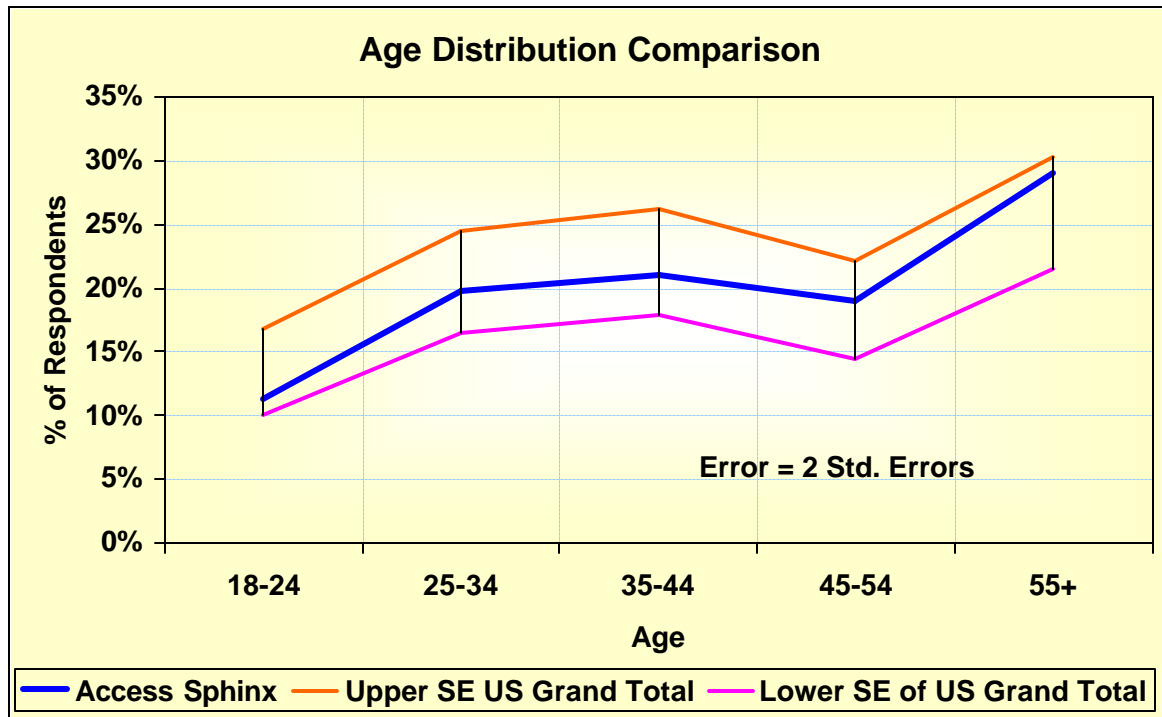
The reference or basis of comparison for this test of Access Sphinx data source is the US Grand Total. This is viewed as a “Grand Mean” value. However, note that we have used exclusively standard online data sources and panels. Alternative sources of data such as random phone dialing were not included in the formulation of this “Grand Mean”. It is important to also note that this grand mean is specifically regional. That is, it reflects the samples for a specified country and is not global.

2. DEMOGRAPHIC EVALUATION

Demographics reflect the nature of the respondents. One would expect the demographics of a sample to reflect that of the reference. As previously mentioned, quotas were requested based on the general population. As such, unless there was some error in the process, the resulting demographic distributions of these quoted variables should be in-line with those of the reference. Both age of the respondent and income group were quoted. However, there were several other demographic characteristics measured that should be in-line with the reference though they were not specifically in the quota.

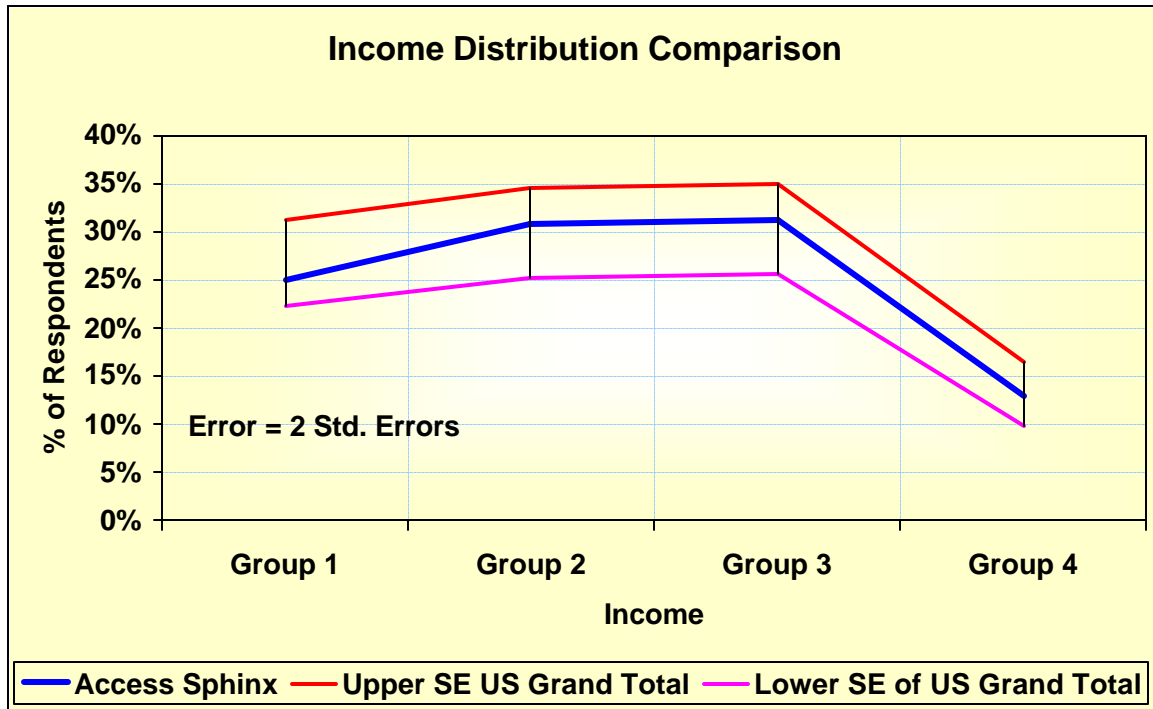
2.1. AGE DISTRIBUTION

A comparison of the age distribution for Access Sphinx against the US Grand Total is shown below. The two red and magenta lines represent the confidence interval (two standard errors) around the reference. The space between represents a 95% probability interval. As would be expected the age distribution remains within the specified limits. A χ^2 test is used to estimate the likelihood that the series is the same as the standard. Based on that test there is greater than a 50% change that the panel age distribution for Access Sphinx is the same as for the US Grand Total.



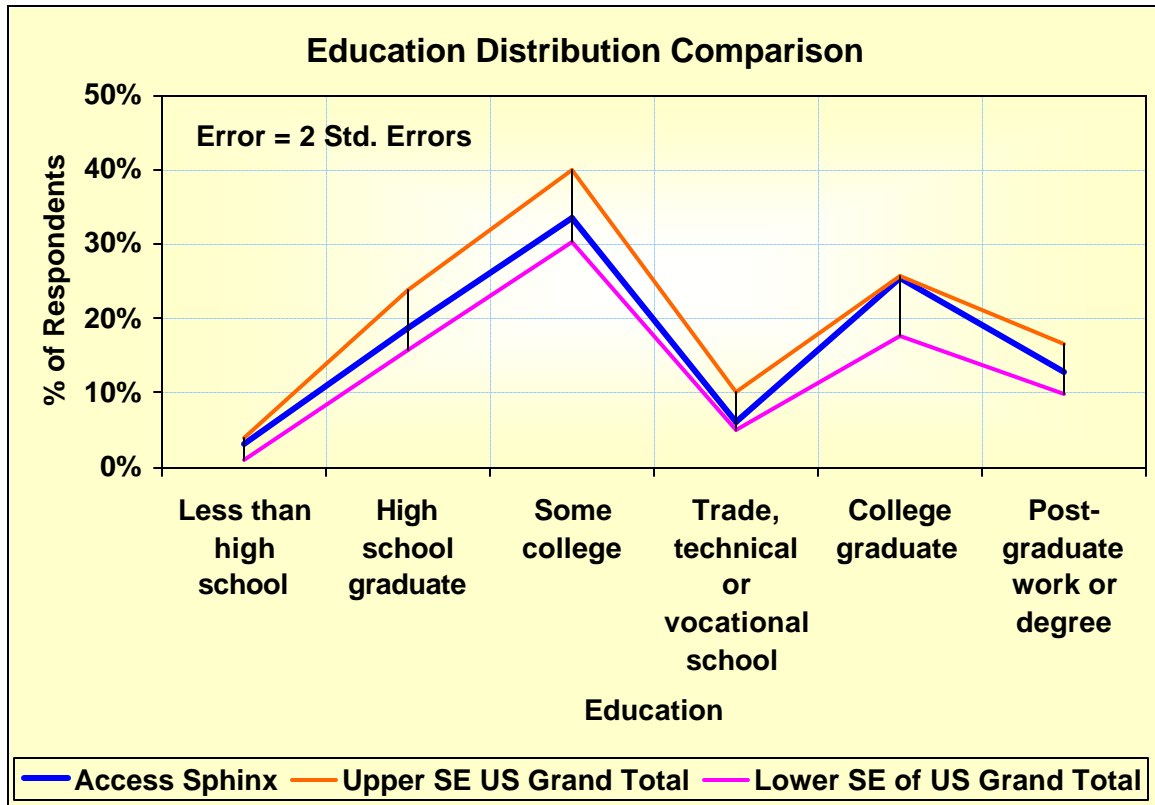
2.2. INCOME DISTRIBUTION

Similarly the income distribution for Access Sphinx was quoted and as such it would be expected to fall well within the confidence interval of the reference, US Grand Total. Note that we are using grouped income levels. This allows for a similar analysis to be done on global panels with different monetary system and ranges. As would be expected the income distribution remains within the specified limits. Based on the χ^2 test there is greater than a 87% chance that the panel income distribution for Access Sphinx is the same as for the US Grand Total.



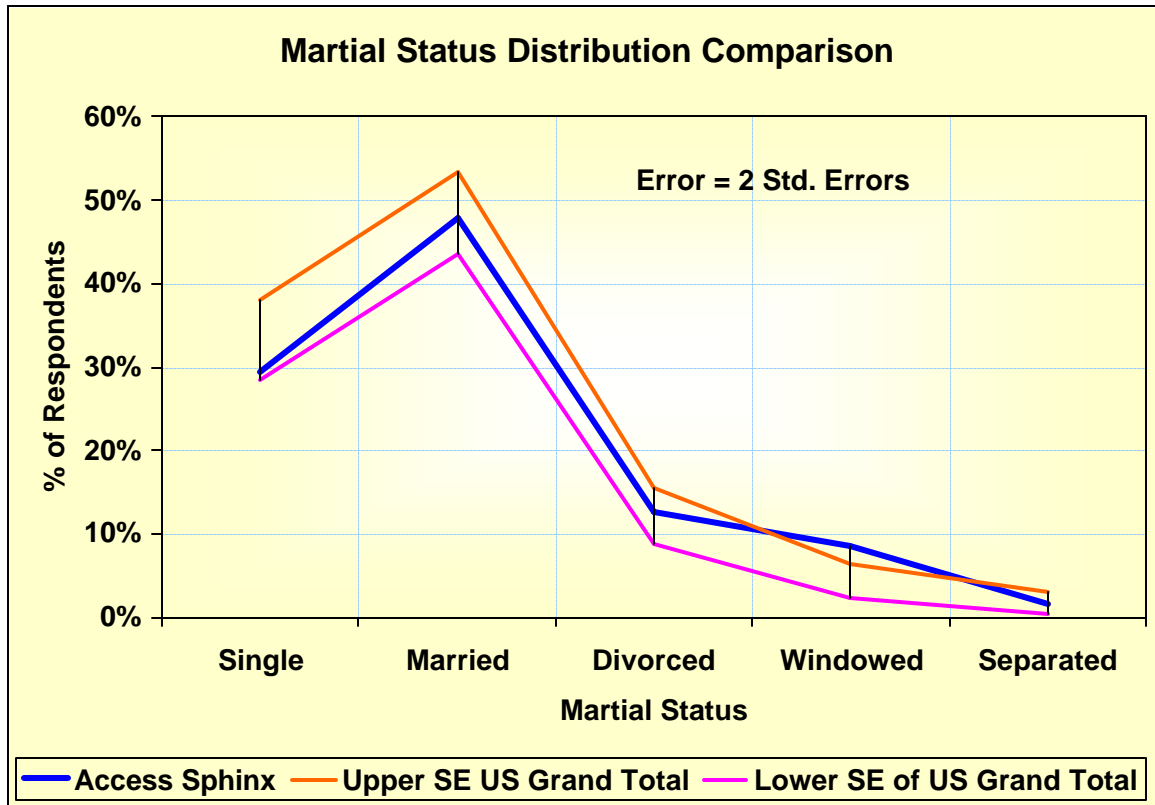
2.3. EDUCATION DISTRIBUTION

Education achievement levels were not quoted. As such they are likely to vary between the source data and the reference. Below is the Education Achievement Distribution for Access Sphinx compared to the US Grand Total. Even though this variable was not quoted, it would be expected that the education distribution to remain within the specified limit and is shown to be so. However, this distribution is closer to the limit than either the age or income distributions. Based on the χ^2 test there is greater than a 36% change that the panel education distribution for Access Sphinx is the same as for the US Grand Total.



2.4. MARITAL STATUS DISTRIBUTION

Similarly, the frequency of marital status was not controlled by quota. It was allowed to vary with the source. Below are the results of the comparative distributions for Access Sphinx compared to the US Grand Total. Note that we have used standard definitions of marital status. There may be some complications in these definitions but we believe that it should not affect the comparison. The marital status distribution shows points beyond the acceptable range. This may be a random fluctuation, but it is statistically significant. Based on the χ^2 test there is greater than a 0.08% change that the panel marital status distribution for Access Sphinx is the same as for the US Grand Total.



3. *STRUCTURAL SEGMENT DISTRIBUTIONS*

In the previous section we examined singular measures to identify the performance of data obtained from various sources. These reflect the operations of the panel or list source, but not the nature of the data itself. Typically, panels and lists are filtered to balance demographics against some external standard such as the known general population. However, this still does not assure that the source reflects the targeted group of respondents or even the larger population. Valued segmentation is designed to capture the distribution of alternative groups of customers that are expected to impact the studies executed using these sources of data. They reflect the potential bias or hopefully the lack of it in the panel. Ultimately the goal would be to balance panels to better reflect the target population distribution of these segments.

Pragmatically Valued Metrics

The key issue is to identify driving factors that need to be addressed to assure an effective sample. These are “pragmatic” issues in that they represent concerns regarding the makeup of panels and sources as they would affect studies. Here we use the term “pragmatic” in that it is not a theoretical construct of what should happen, but what is expected to happen. Studies involving purchases will depend on the distribution of “buyer behavior” segments and media studies would depend on the distribution of “media use” segments, for example.

Quality Clusters

In order to construct a functional segmentation scheme, the underlying groups should be based on the identification of high quality clusters. Whether it is based on cluster analysis or any other method of identification the underlying structure should be:

- **Unique** - The cluster structure should be singular, and unambiguous.
- **Distinct**- Distinction is measured by the degree that segments’ variable values are different from the grand mean and that of other segments.
- **Reliable**¹ – Reliability reflects the clarity of the assignment into clusters. That is there is little ambiguity in the assignments.
- **Functional** - The resulting clusters must represent an important and varying structure that is meaningful to the process of collecting intelligence. As such, the structure should indicate variation within countries for which it is expected to change. That is, the definition of the clusters should not be a trivial universal principle which does not reflect variation.

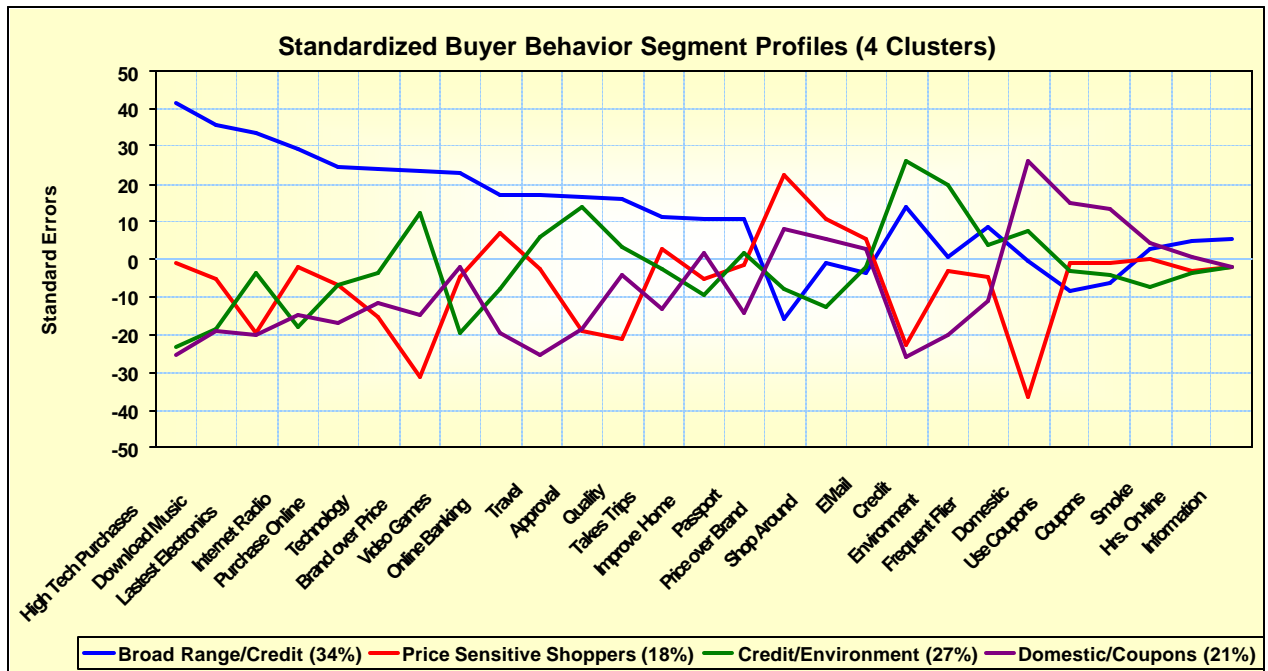
¹ Reliability is tested as part of the formation of the segmentation model using Multinomial Logit Regression. These logit models are then used to assignment segments in with data sets not used in the original segmentation exercise. The existence of a stable highly reliable regression model is critical for the use of segmentation as an analytical tool for progressive sets of data.

Segmentation Definitions

Three segmentation schemes are being used in this evaluation focusing on: (1) Buyer Behavior, (2) Sociographic Factors, and (3) Media Use Factors. These are not the only segmentation schemes that can be developed for this process. However, these were well supported by the test questionnaire and fulfilled the requirements as structural segments.

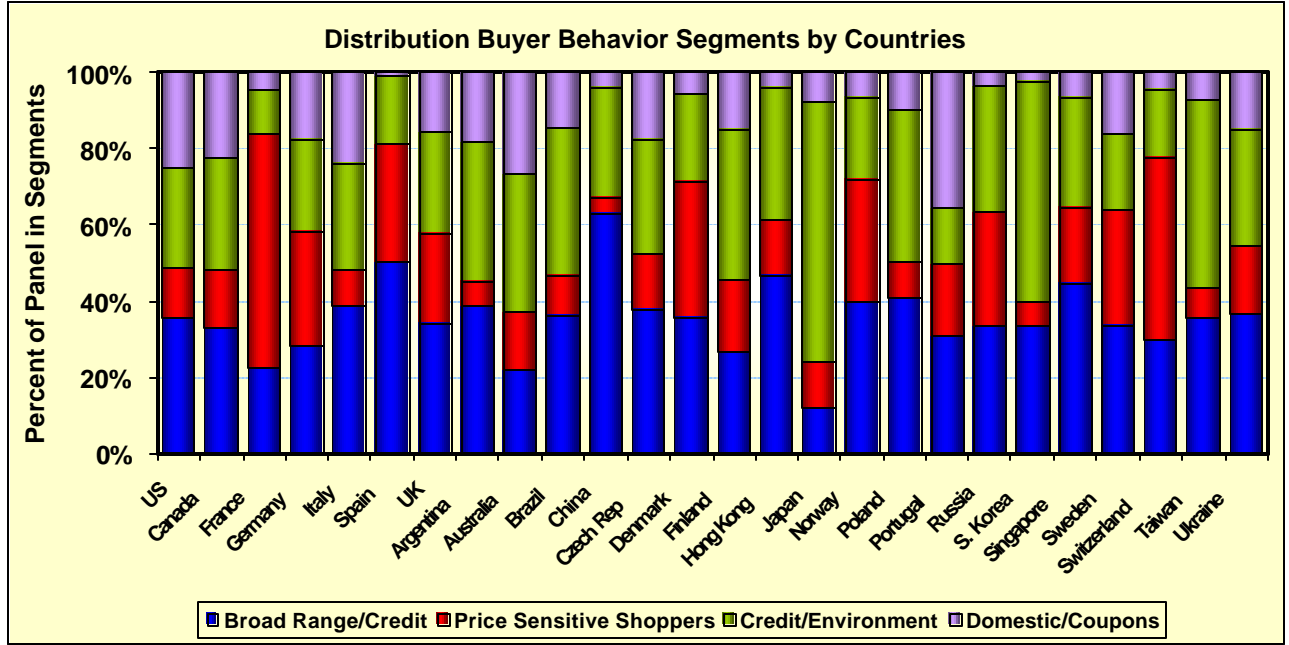
3.1. BUYER BEHAVIOR SEGMENTS

The buyer behavior segments are intended to capture the variability in the attitudes and actions regarding the purchase of a broad range of products. The standardized profiles are shown below and reflect the response to 37 input variables.

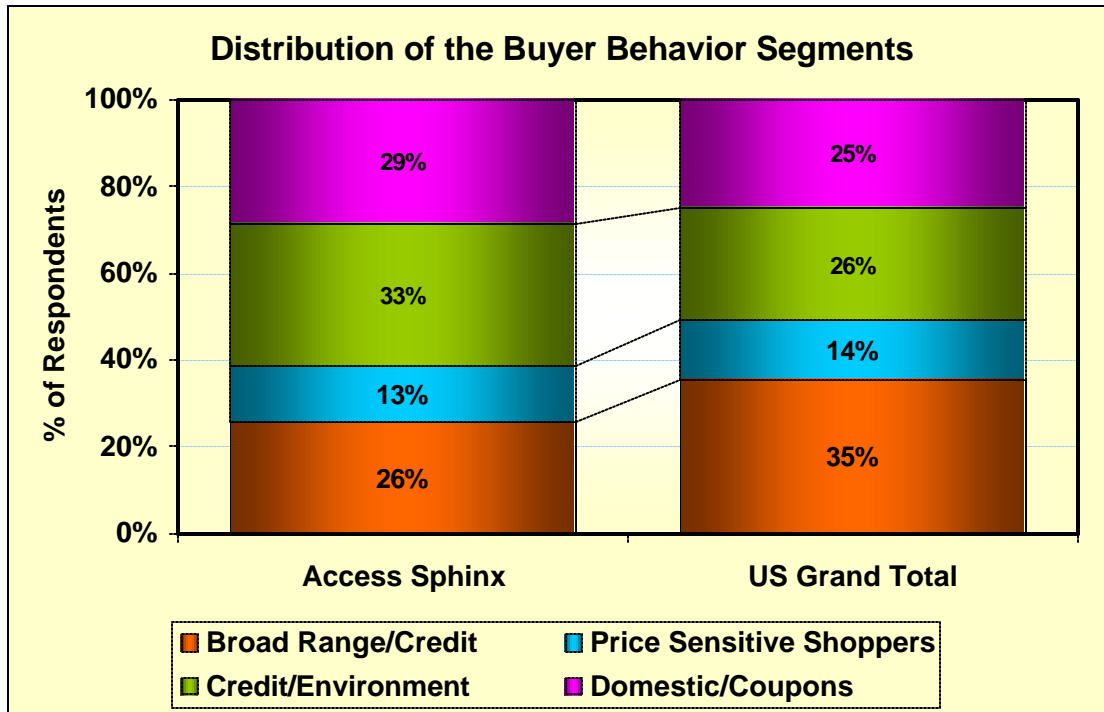


The titles of the segments reflect the strongest loading variables making up the segment. The purpose of this scheme is to reflect differences between sources of data and the general grand mean for that representing region. It is important to note that the distribution of these segments vary widely between different countries and regions as shown in the chart below. These are cultural variations that are expected to vary. However, we expect the distribution of these segments among panel and sources of data within regions to be consistent.

Evaluation of Access Sphinx Data Source

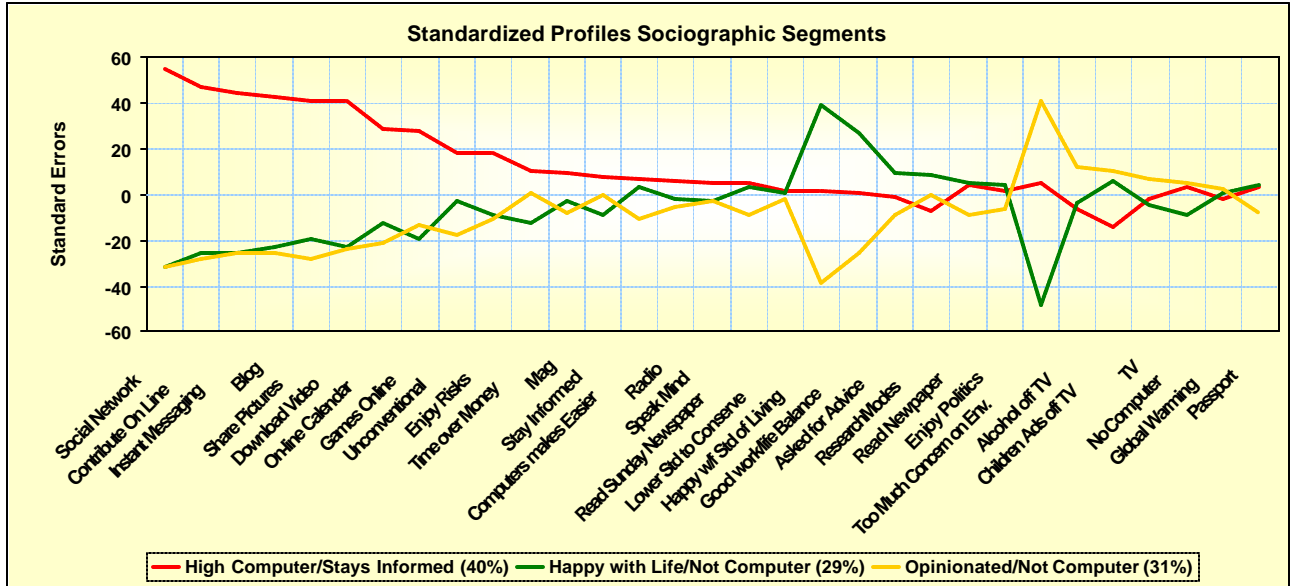


Below is the comparison between the buyer segment distributions for Access Sphinx and the US Grand Total. These distributions are significantly different. Based on a χ^2 test, there is less than a 10% chance that the distribution of buyer behavior segments for Access Sphinx is the same as for the US Grand Total. The largest difference is 28% relative change with the Broad Range/Credit segment. Note that the average expected variation is 4.2%.

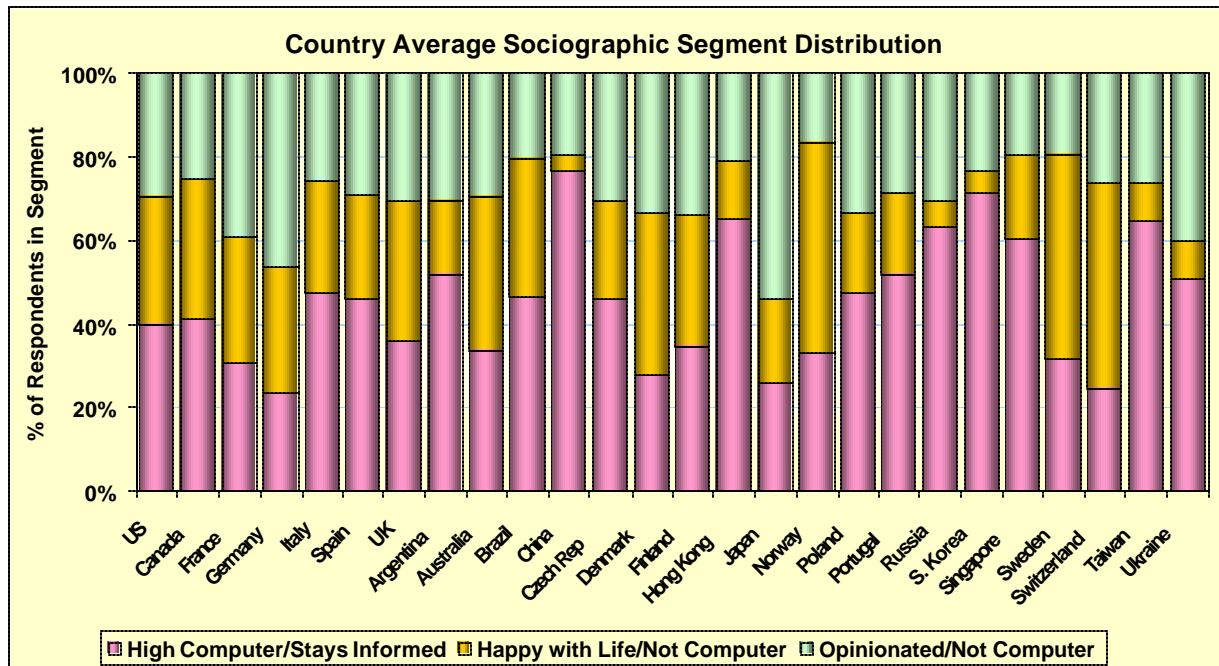


3.2. SOCIOGRAPHIC SEGMENTS

The sociographic segments are intended to capture the variability in behavior and attitudes regarding a broad range of life style decisions. The standardized profiles are shown below and reflect the response to 31 input variables.

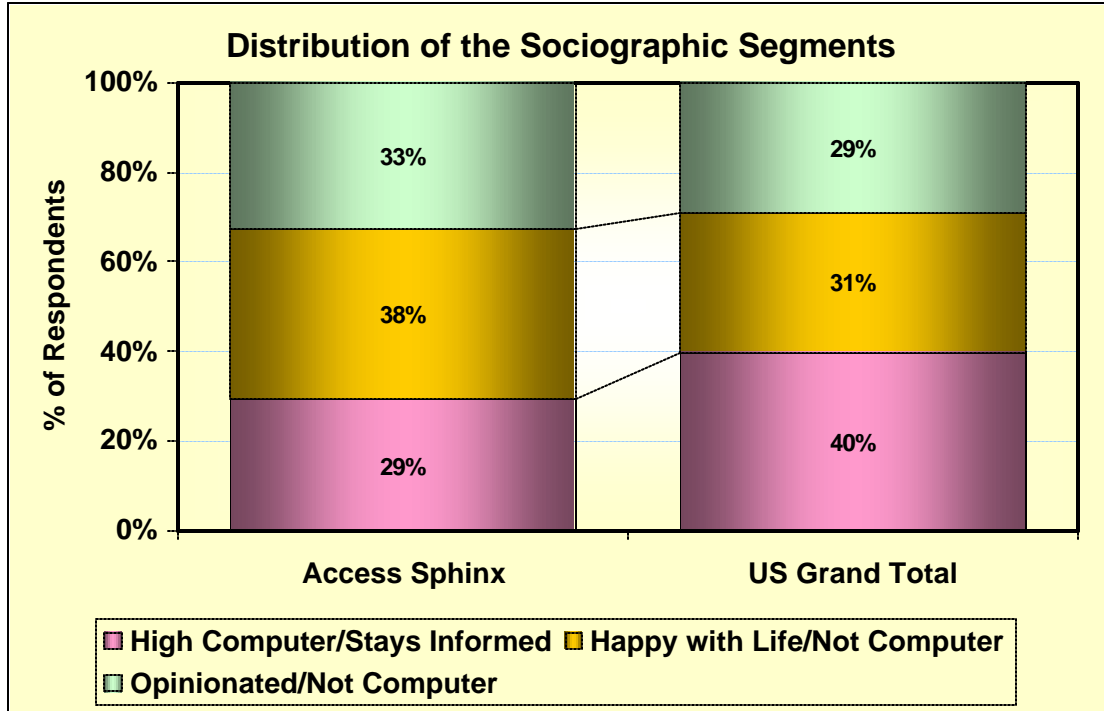


As in the case of the buyer behavior segments, the titles of the segments reflect the strongest loading variables making up the segment. It is important to note that the distributions of these segments are expected to vary widely between different countries and regions as shown in the chart below. However, we expect the distribution of these segments among panel and sources of data within regions to be consistent as before.



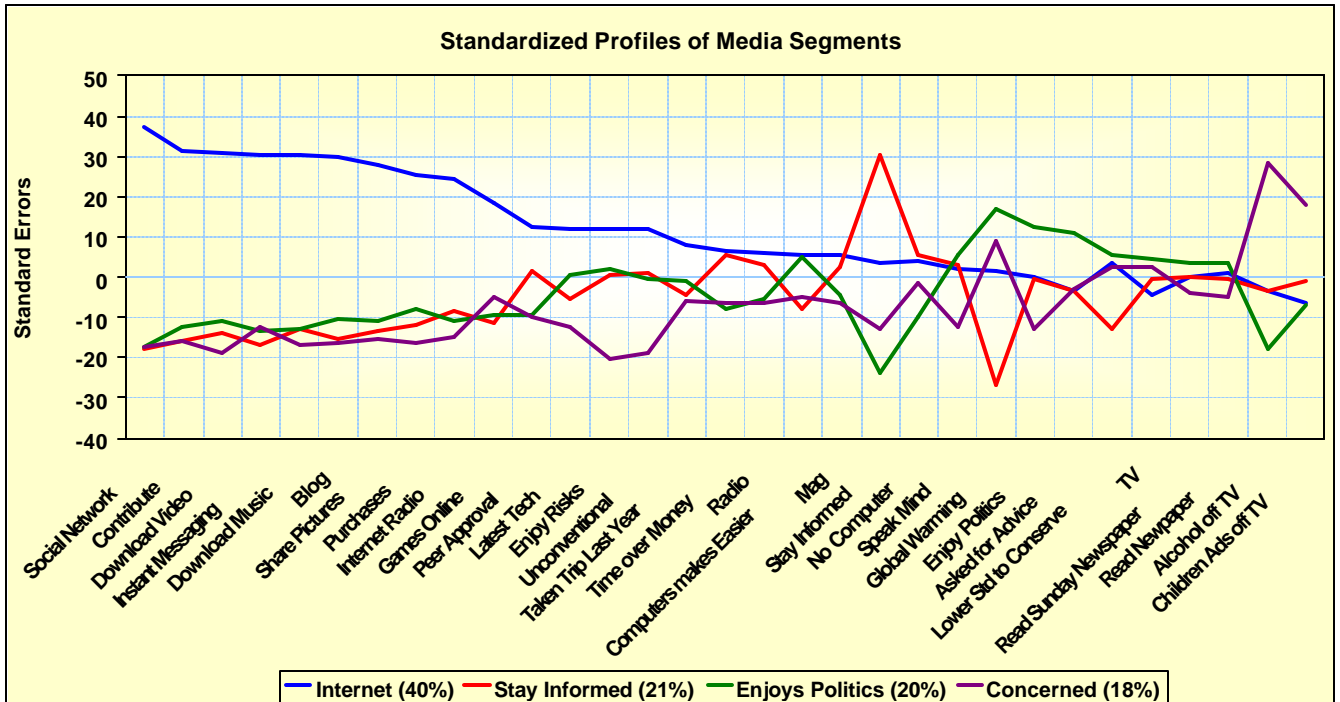
Evaluation of Access Sphinx Data Source

Below is the comparison between the sociographic segment distributions for Access Sphinx and the US Grand Total. These distributions are very significantly different based on χ^2 test, there is less than a 3% chance that the distribution of sociographic segments for Access Sphinx is the same as for the US Grand Total. The largest difference is 27% relative change with the High Computer/Stays Informed segment. Note that the average expected variation is 4.7%.



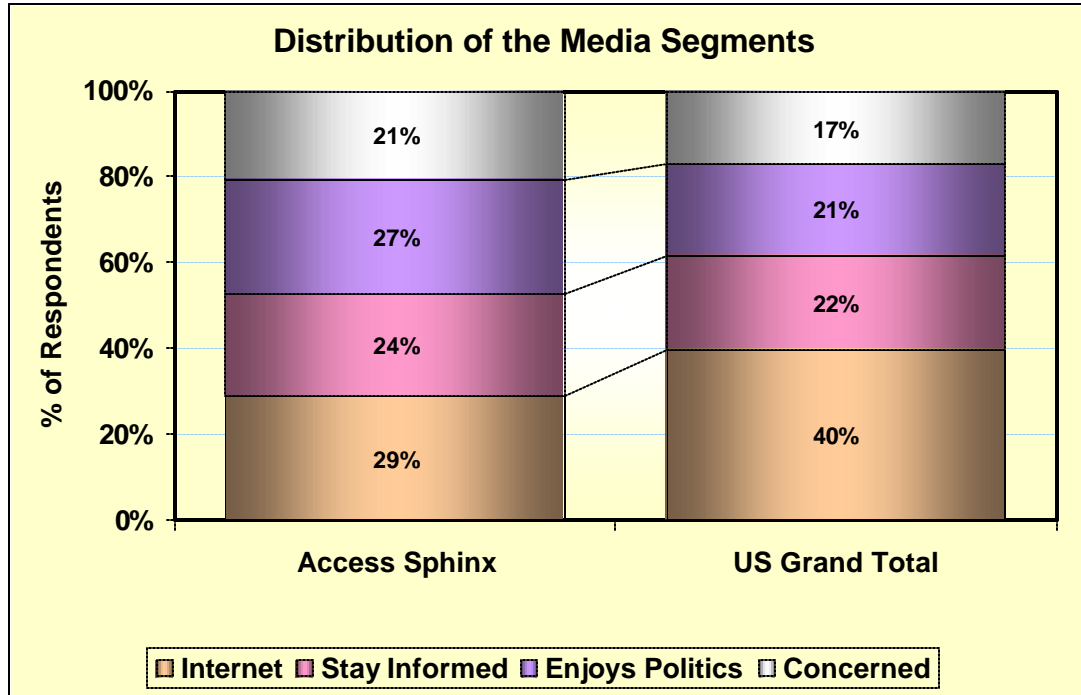
3.3. MEDIA SEGMENTS

The media segments are intended to capture the variability in the use of various sources of communications and activities. The standardized profiles are shown below and reflect the response to 31 input variables. The variables used were combinations of those also used for the buyer behavior and sociographic segmentation but focused on media issues only. As in the case of the other segmentation schemes, the titles of the segments reflect the strongest loading variables making up the segment.



Evaluation of Access Sphinx Data Source

Below is the comparison between the media segment distributions for Access Sphinx and the US Grand Total. These distributions are significantly different based on a χ^2 test; there is less than a 8% chance that the distribution of media segments for Access Sphinx is the same as for the US Grand Total. The largest difference is 27% relative change with the Internet segment. Note that the average expected variation is 4.2%.



4. SOURCE PERFORMANCE

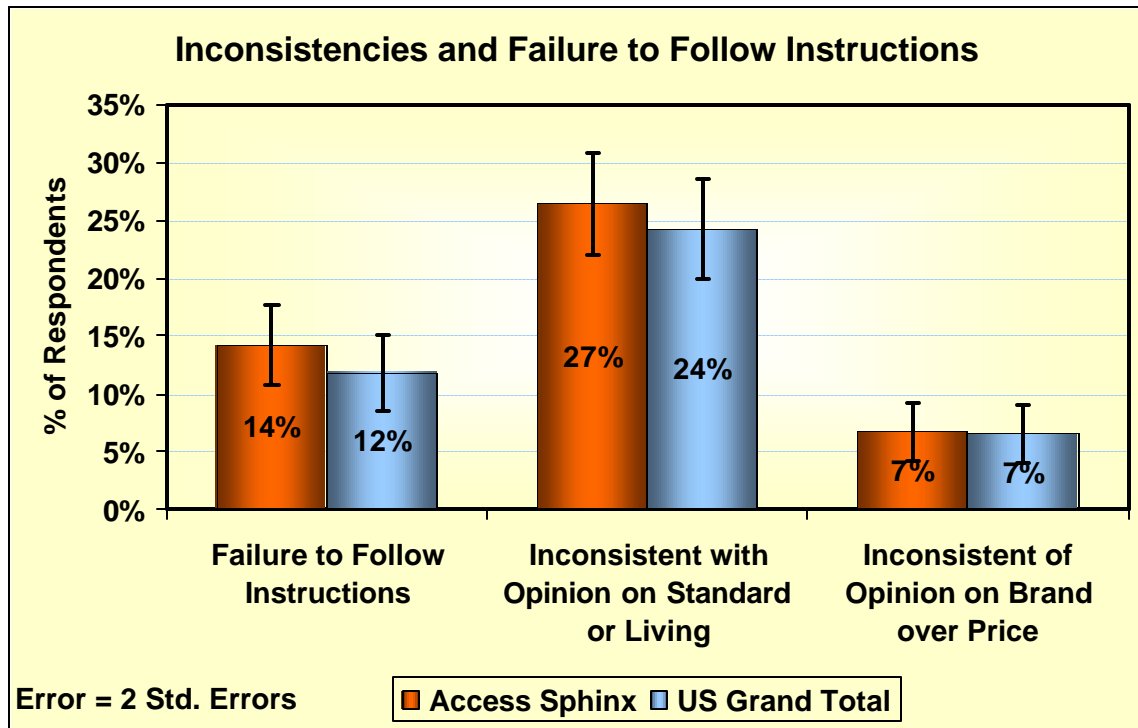
There are several types of panel performance information that might be useful. These reflect the nature of the data source and its functional performance. These are often structural issues reflecting the mechanism of generating and encouraging participants.

- ***Performance***- The incidence of errors in the execution of questionnaires reflects the quality of the panel. These are “checks” designed into the testing instrument. They include but are not limited by: (1) inconsistency in responding to multiple questions and (2) the failure to follow instructions
- ***Characteristics*** - The characteristics or structure reflects the nature of the participants in the panels. In general, these focus on issues and concerns with the long term maintenance and in particular the tendency of containing “professional” participants. These metrics may include participants, who belong to multiple panels, have been on panels for an extended period of time or who take multiple surveys frequently.
- ***Deviates*** – Respondents occasionally show extra ordinary characteristics. These are not errors, just extreme behavior which provides a potential warning of problems. These include: (1) “speeders” who finish their questionnaire in extra ordinarily short time and (2) “straight-liners” who tend to give the same answer to a large number of questions.

4.1. PERFORMANCE

The quality of results for Access Sphinx compared to US Grand Total is measured by the frequency of improper responses and is shown below. The quality of survey results is always difficult to ascertain. In this testing procedure, specific “trick” questions were used that allow the assessment of proper action. The test instrument has two types of items: (1) questions or more specifically required action that the failure to do is viewed as a “fault”, and (2) inconsistencies where two balanced but opposite questions are asked, answering both strongly positive or both strongly negative would then be viewed as another type of “fault”.

The questionnaire had one instructional question where the respondent was required to put in a specific set of values. If an improper response was made it is viewed as a “Failure to Follow Instructions”. The comparative frequency of this fault is shown below. The expected error (2 standard errors) is shown as the error bar. One would expect approximately 95% of the random errors to exist in this range. The frequency of this fault is, therefore, not significantly different for Access Sphinx compared to the US Grand Total.

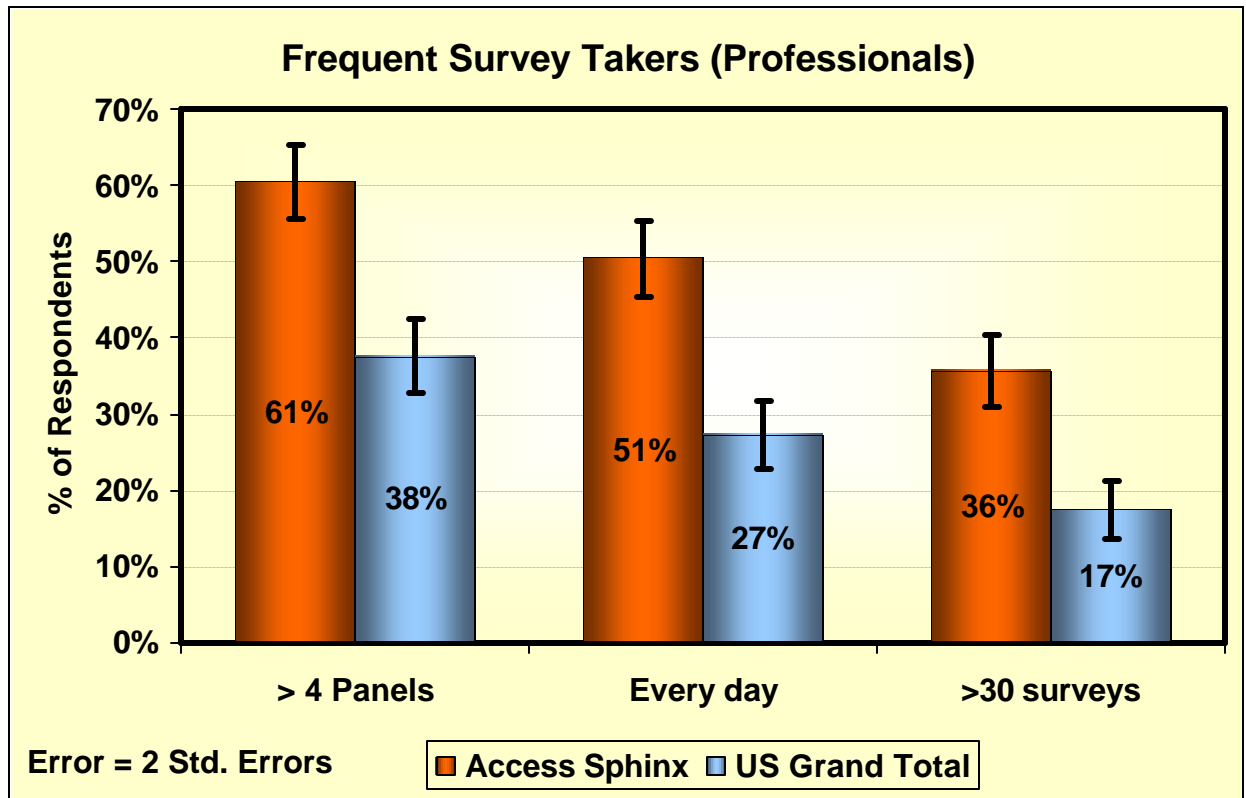


The other two measures capture inconsistent responses: [happy/unhappy with standard of living](#) and [brand over price/price over brand](#). As noted above, these measures are based on having given either strong positive or strong negative values to both directions. As seen above the frequency of the inconsistency of contentment with Standard of Living was not significantly different for the panel against the reference.

4.2. CHARACTERISTICS (FREQUENT SURVEY TAKERS)

A concern regarding online panels is the development of “professional” survey takers. These are members of the panel or data sources that are frequent survey-takers. The frequency of professionals is estimated in terms of four measures: (1) is belonging to 5 or more panels, (2) taking surveys almost everyday, (3) having taken at least 30 surveys in the past month, and (4) panel tenure. All of these are self assessments and as such may be in error, but they represent consistent metrics. Below are the results of the comparison of Access Sphinx and the reference of the US Grand Total.

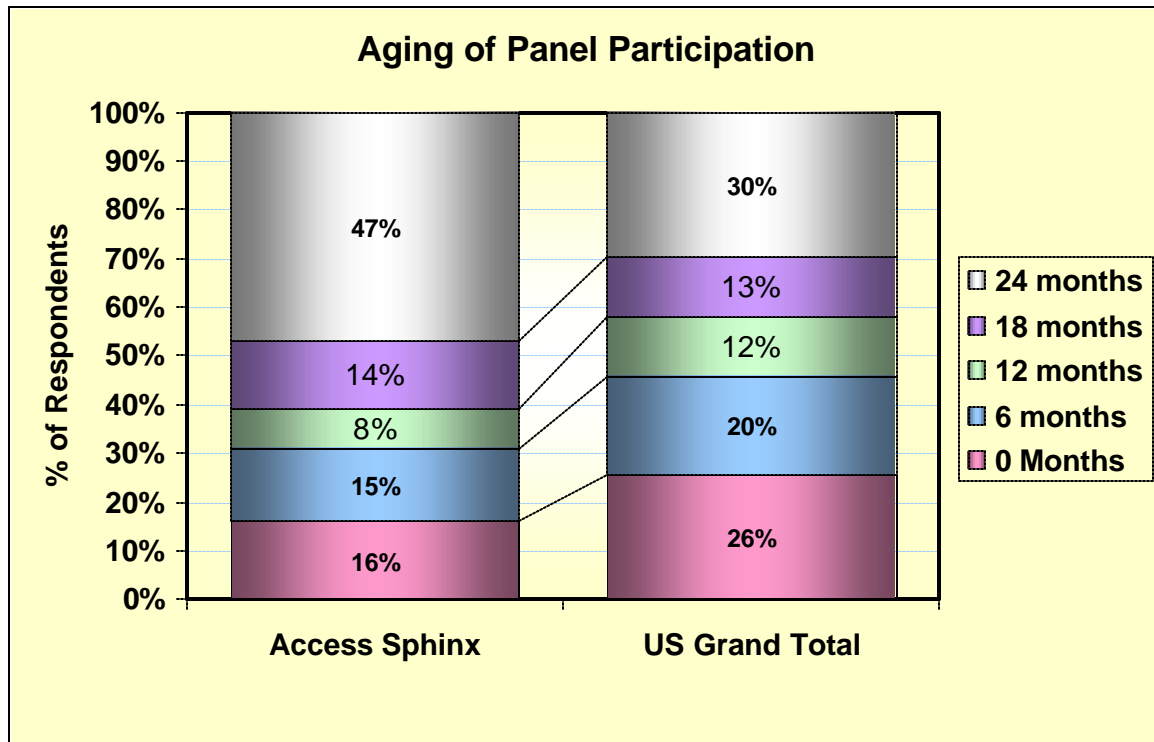
Note that all of the professional measure frequencies for the test data source are higher than the reference. In regards to being in multiple panels, the frequency of Access Sphinx members appear to be significantly different and higher than the US Grand Total. As for taking surveys, the frequency of Access Sphinx members appears to be higher and significantly different from the reference. Similarly, the frequency of taking 30 surveys or more a month was significantly different and higher for the source than the reference.



Evaluation of Access Sphinx Data Source

The complementary issue of belonging to other online panels is the concern over how long have members remained on panels. This is viewed as an aging or acculturation problem. At this point, there is little evidence that aged members of panels cause problems or represent unacceptable participants. But the concern remains.

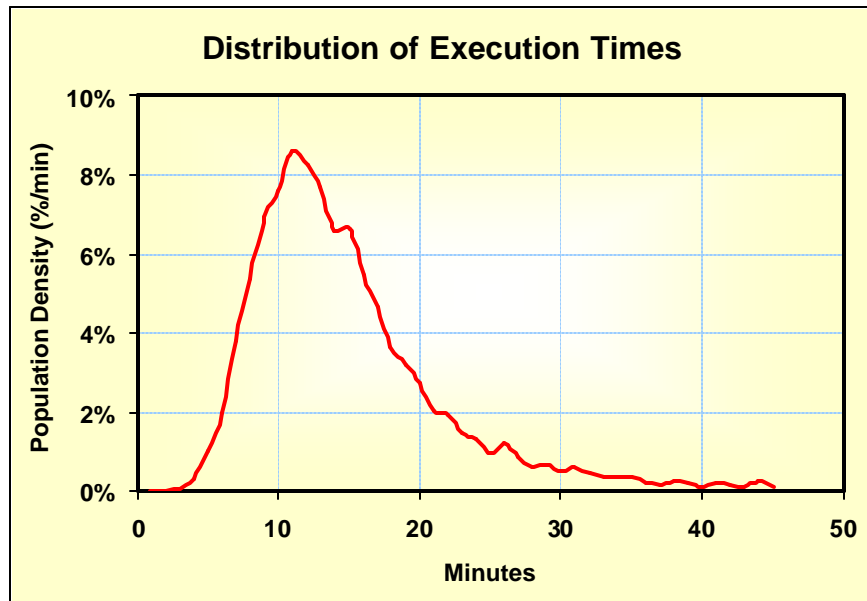
Below is the comparison between the aging of panel participation distributions for Access Sphinx and the US Grand Total. These distributions are very significantly different based on a χ^2 test there is less than a 0.00% chance that the distribution of media segments for Access Sphinx is the same as for the US Grand Total. The largest difference is 58% relative change in the 24 months group. Note that the average expected variation is 3.7%.



4.3. DEVIATE BEHAVIOR (SPEEDERS AND STRAIGHT-LINERS)

The previously performance characteristics that were covered focused on the errors made by respondents, and their participation in surveys and panels. There is a third category of activities that are thought to possibly affect the quality of results. These are the participants who either speed through the survey (speeders) and those who give similar or identical values to blocks of questions in the surveys (straight-liners). These respondents are viewed as exhibiting deviate behavior. However, once again, there is no direct evidence that such behavior results in poor quality survey results.

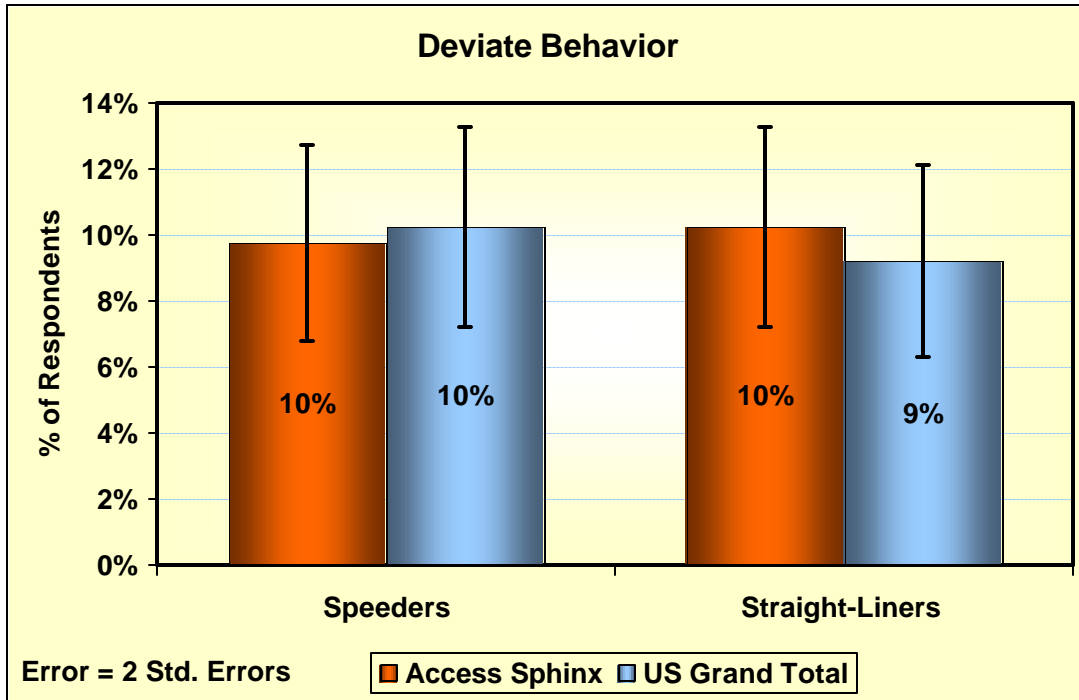
Below is a typical distribution of completion times for the test survey. Note that it has been truncated at 45 minutes. Because of the nature of online surveys, participants may delay execution and thereby run up huge apparent elapsed times. Typically these long times are removed for analysis. Speeders are those that finish the questionnaire very fast. Generally for our test instrument that is less than an overall lower 10 percentile.



Similarly, straight-liners are defined based on the lack of variation in their responses. For the purposes of analysis, the standard deviation over a range of similar questions is used to estimate variation. Straight-liners are defined for this analysis as respondents with a selected standard deviation of 1 unit (out of 7) or less for 30 questions.

Evaluation of Access Sphinx Data Source

The frequency of speeders of Access Sphinx members appear not to be significantly different from the US Grand Total. Similarly, the frequency of straight-liners also appear not to be significantly different for the source than for the reference.



Elaine Trimarchi
Executive Vice President – PRC Certified
Phone (631) 277-7000
Cell (631) 664-1308
etrimarchi@mktginc.com
www.mktginc.com
200 Carleton Avenue, East Islip, NY 11730