

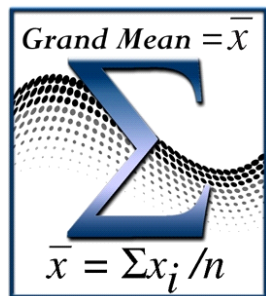
Consumer Data Source Evaluation

The Performance Evaluation of:

Tiburon - Russia

December 7, 2010

By



SAMPLE SOURCE AUDITORS™
A DIVISION OF **Mktg.**
Incorporated

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. Marital Status Distribution.....	8
3. Structural Segment Distributions.....	9
3.1. Buyer Behavior Segments	10
3.2. Sociographic Segments.....	13
3.3. Media Segments.....	15
4. Source Performance	17
4.1. Performance	18
4.2. Characteristics (Frequent Survey Takers).....	19
4.3. Satisficing Behavior (Speeders and Straight-Liners).....	21
4.4. Q-Metrics™	23

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 cannot 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 cannot 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 cannot 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, 501 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.

Typically for “random” and non-structured sampling, some form of quota corrections are used to help balance the resulting sample. These quotas may reflect the specific needs of

Evaluation of Tiburon - Russia Data Source

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.

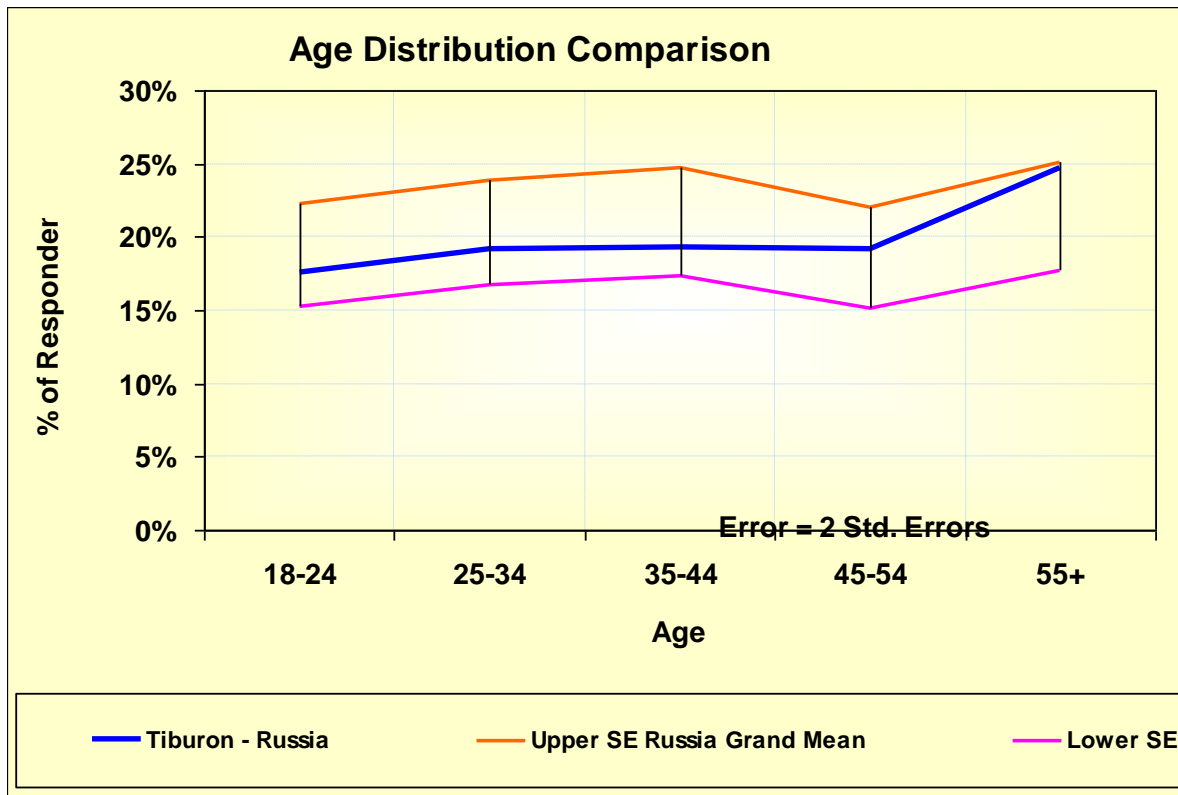
The reference or basis of comparison for this test of Tiburon - Russia data source is the Russia Grand Mean. This is viewed as a “Grand Mean” value. However, note that we have used exclusively standard online data sources and panels. Alternative sources of data such as random phone dialing were not included 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 quota controlled variables should be in-line with those of the reference. Both age of the respondent and income group were quota controlled. However, there were several other demographic characteristics measured that should be in-line with the reference though they were not specifically in the quota.

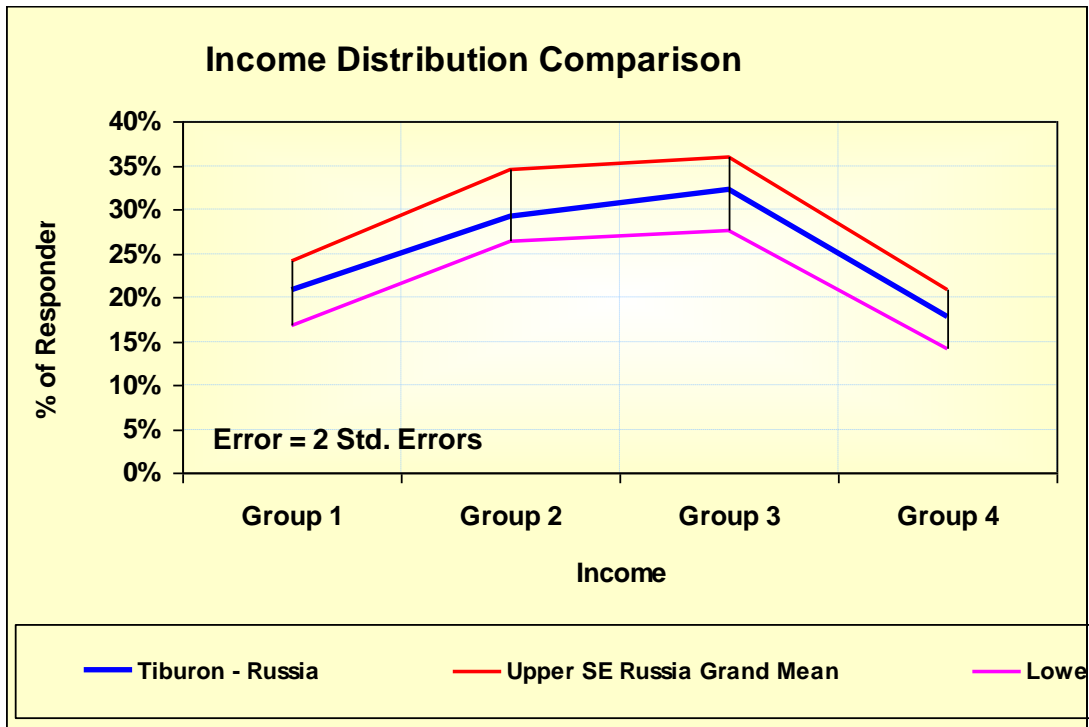
2.1. AGE DISTRIBUTION

A comparison of the age distribution for Tiburon - Russia against the Russia Grand Mean 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. 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 40% chance that the panel age distribution for Tiburon - Russia is the same as for the Russia Grand Mean.



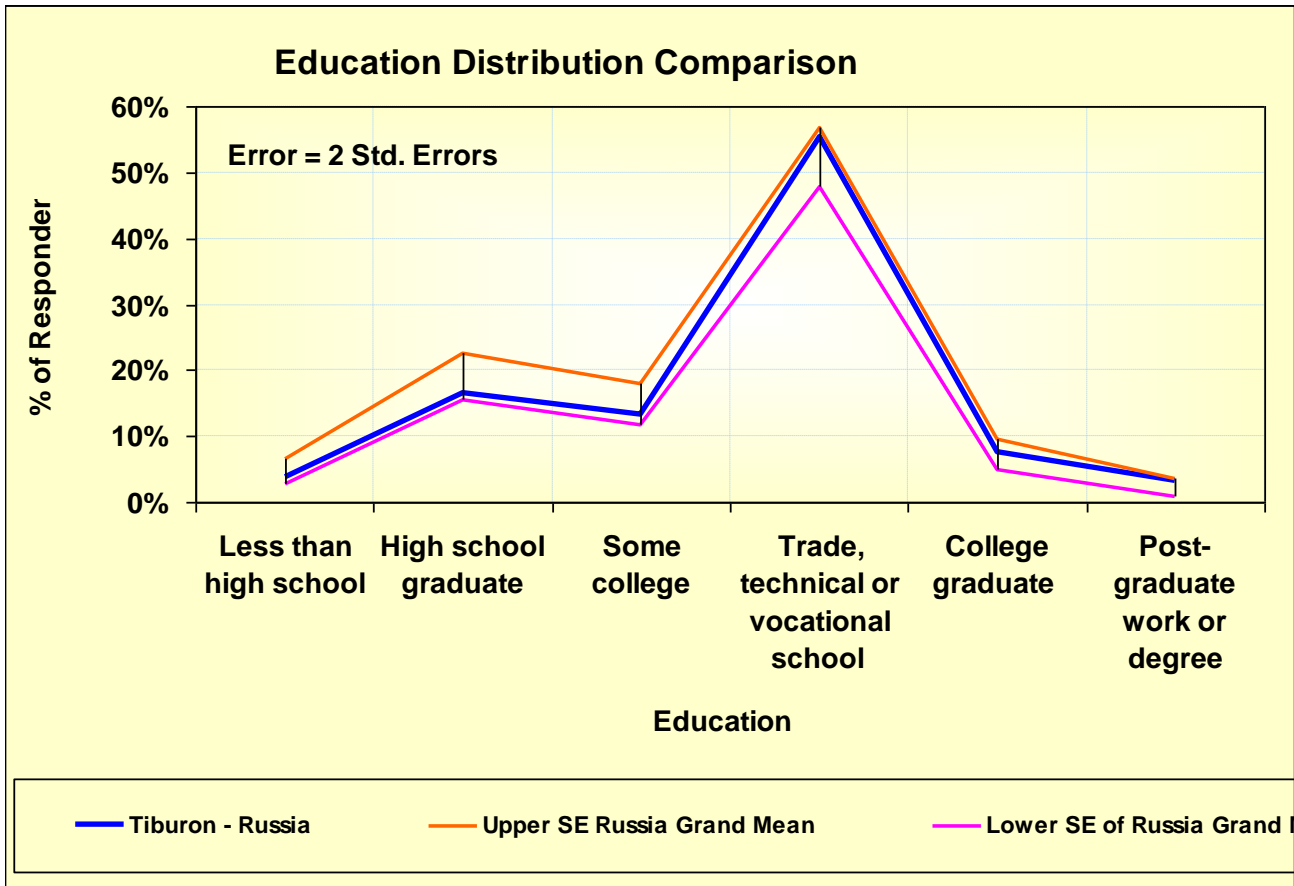
2.2. INCOME DISTRIBUTION

Similarly the income distribution for Tiburon - Russia was quota controlled and as such it would be expected to fall well within the confidence interval of the reference, the Russia Grand Mean. 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. The income distribution remains within the specified limits. Based on the χ^2 test there is greater than a 30% chance that the panel income distribution for Tiburon - Russia is the same as for the Russia Grand Mean.



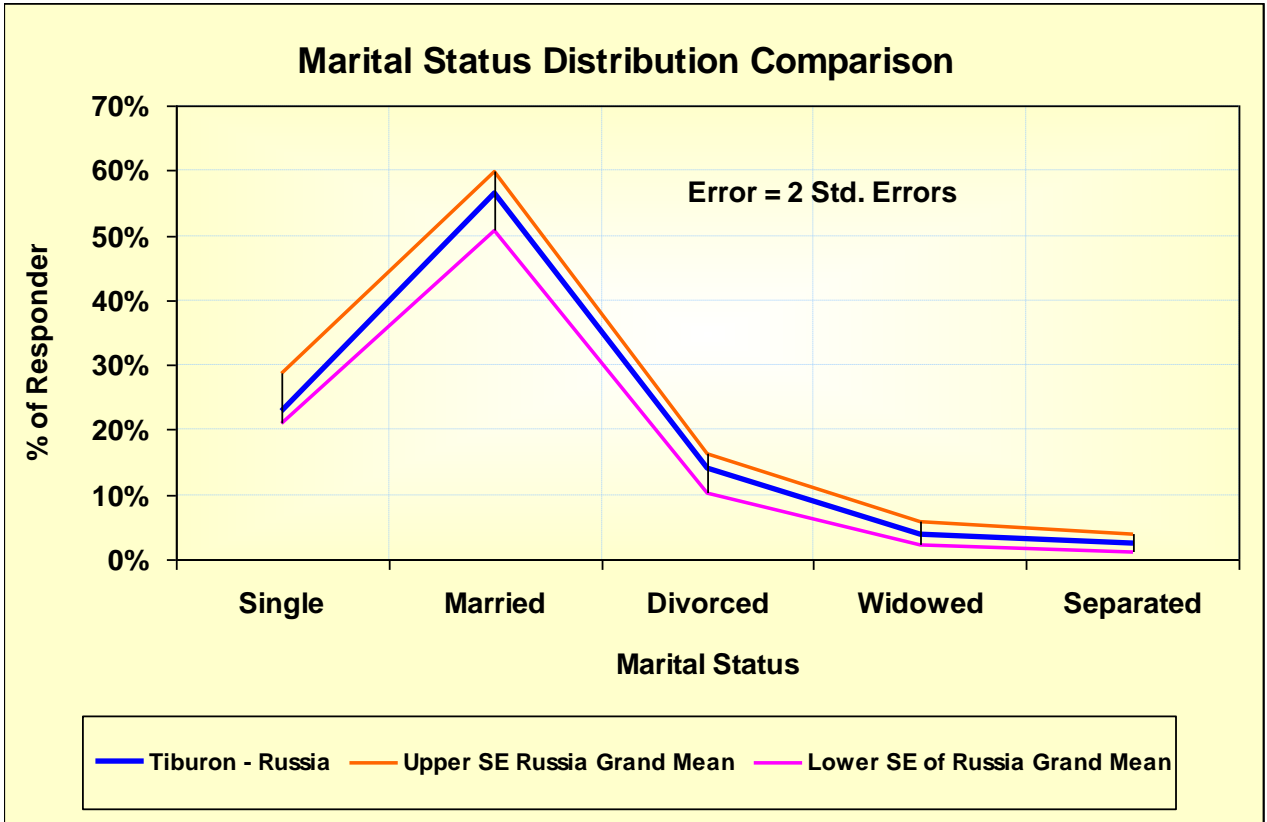
2.3. EDUCATION DISTRIBUTION

Education achievement levels were not quota controlled. As such they are likely to vary between the source data and the reference. Below is the Education Achievement Distribution for Tiburon - Russia compared to the Russia Grand Mean. The education distribution remains within the error bounds. Based on the χ^2 test there is greater than a 0% chance that the panel education distribution for Tiburon - Russia is the same as for the Russia Grand Mean.



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 Tiburon - Russia compared to the Russia Grand Mean. 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 remains within the error bounds. Based on the χ^2 test there is greater than a 85% chance that the marital status distribution for Tiburon - Russia is the same as for the Russia Grand Mean.



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 such that 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. 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.

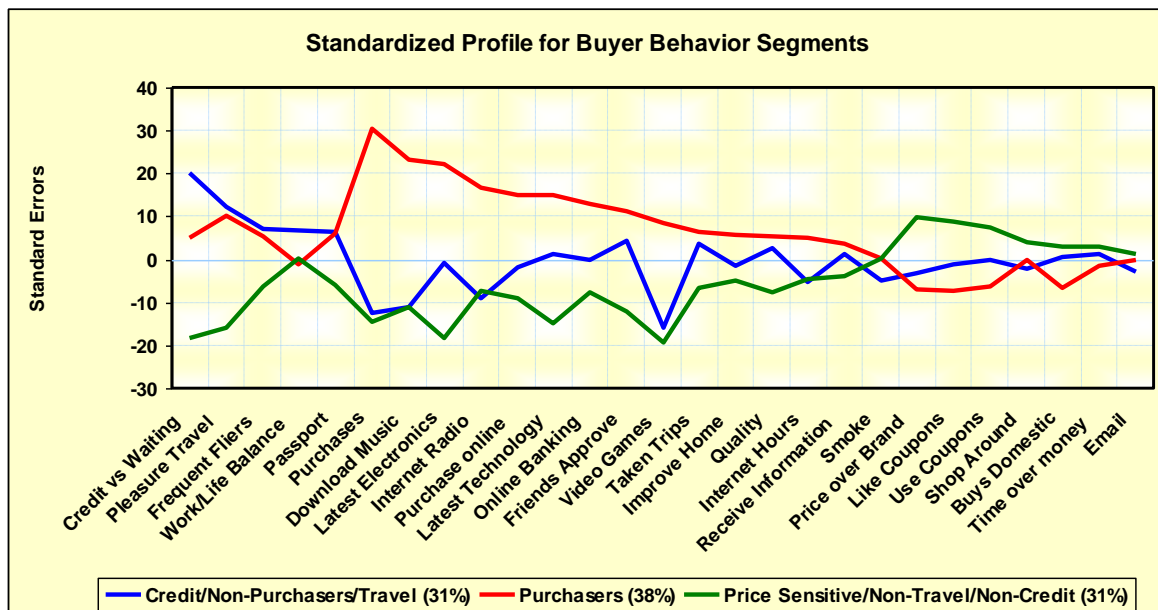
¹ Reliability is tested as part of the formation of the segmentation model using Multinomial Logit Regression. These logit models are then used to assign segments in data sets that were not used in the original segmentation exercise. The existence of a stable highly reliable regression model is critical 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.

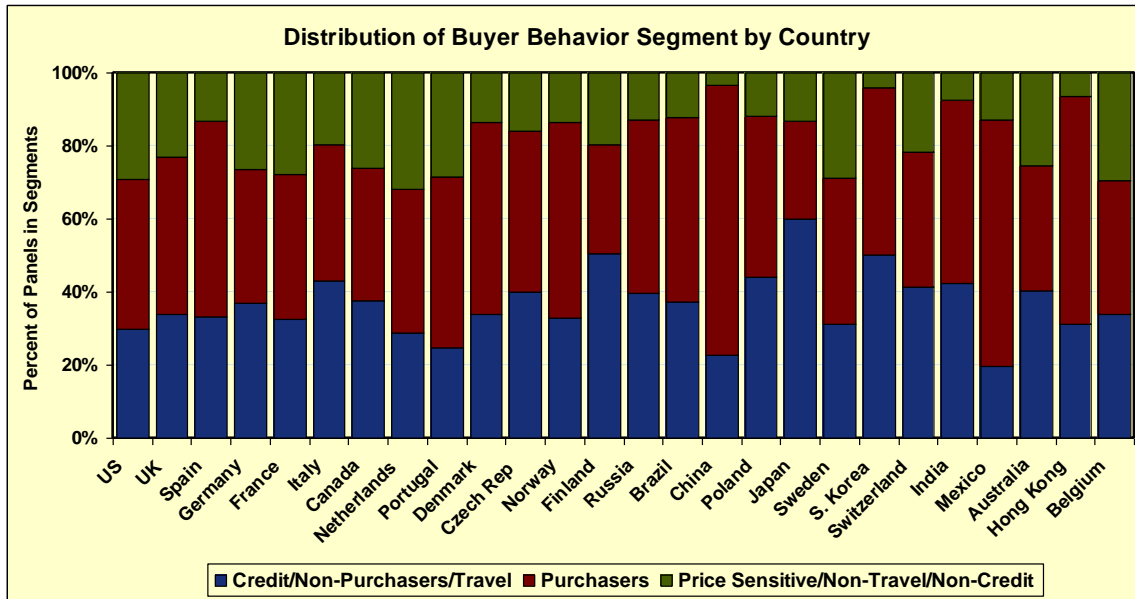
3.1. BUYER BEHAVIOR SEGMENTS

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 36 input variables.



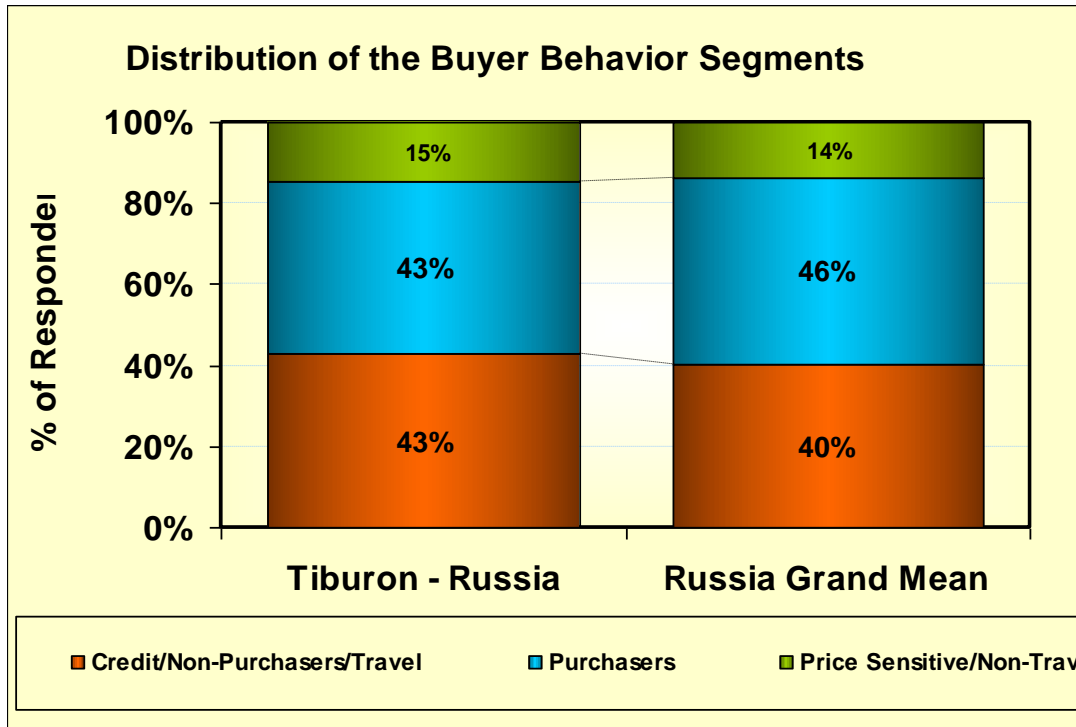
Evaluation of Tiburon - Russia Data Source

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 expected cultural variations. However, we expect the distribution of these segments among panel and sources of data within regions to be less variable.



Evaluation of Tiburon - Russia Data Source

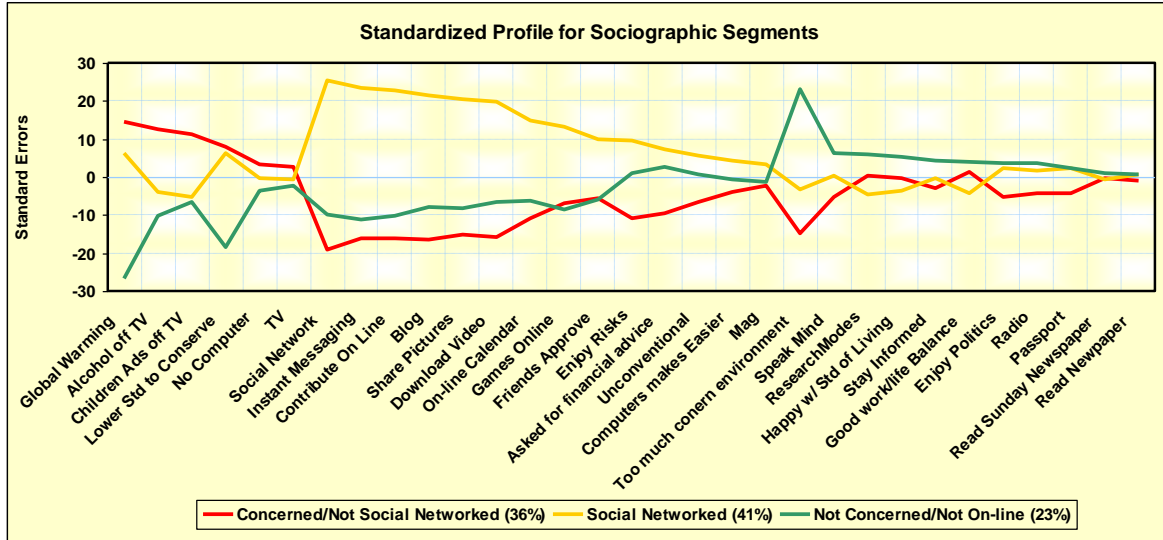
Below is the comparison between the buyer segment distributions for Tiburon - Russia and the Russia Grand Mean. These distributions are border-line different². Based on a χ^2 test, there is less than a 36% chance that the distribution of buyer behavior segments for Tiburon - Russia is the same as for the Russia Grand Mean. The largest difference is a 7% relative change with the Purchasers segment. Note that the average relative difference is 6.0%.



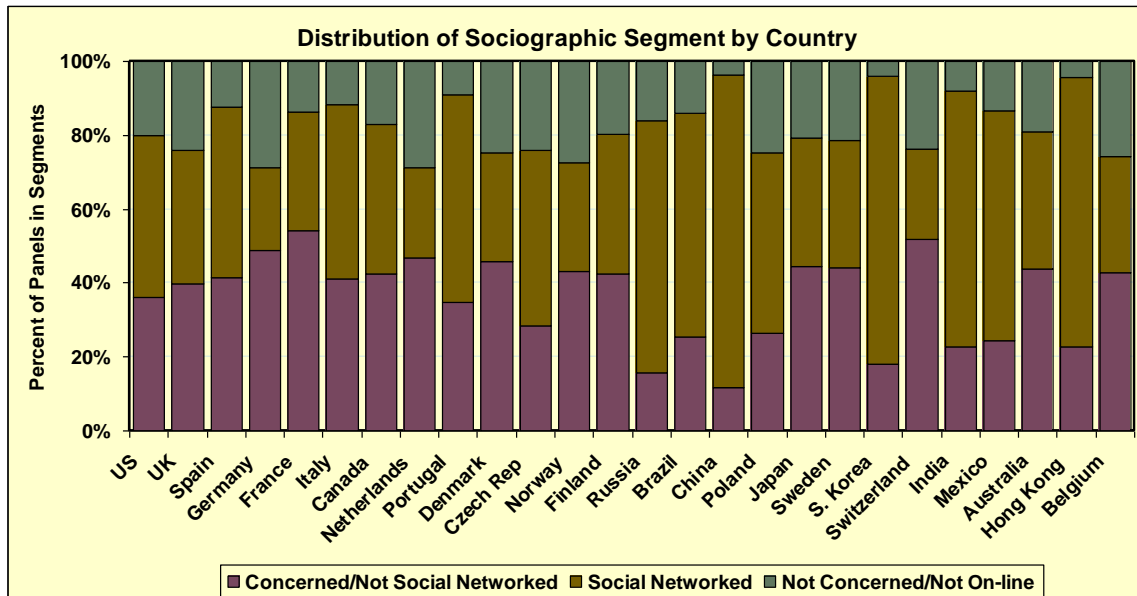
² The measure of differences between the distributions is based on the maximum of the differences of each of the distribution elements divided by its standard error. If the ratio is less than or standard error the distribution is considered not significantly different, if it is between and it is viewed as “border-line”, and greater than standard errors is viewed as significantly different.

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.

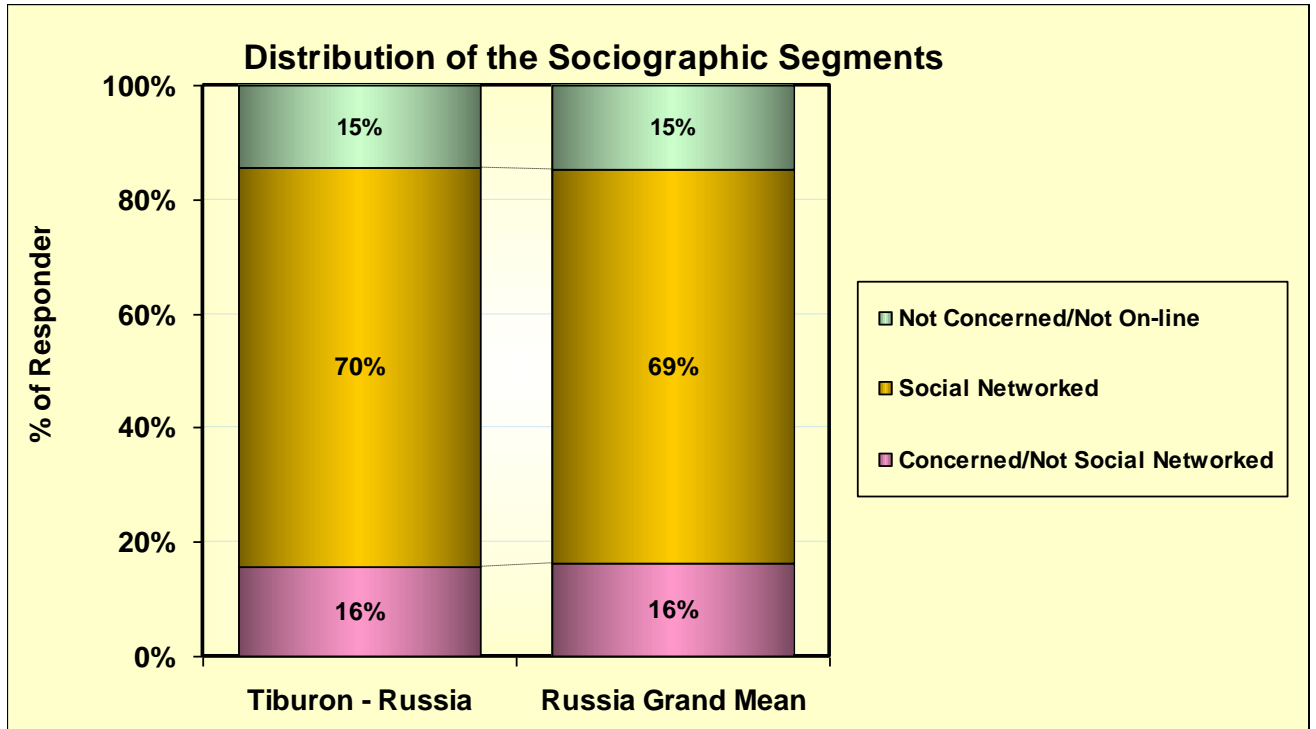


As in the case of the buyer behavior segments, the titles of the segments reflect the strongest loading variables making up the segment. It is important to note that the distributions of these segments are expected to vary widely between different countries and regions as shown in the chart below. However,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



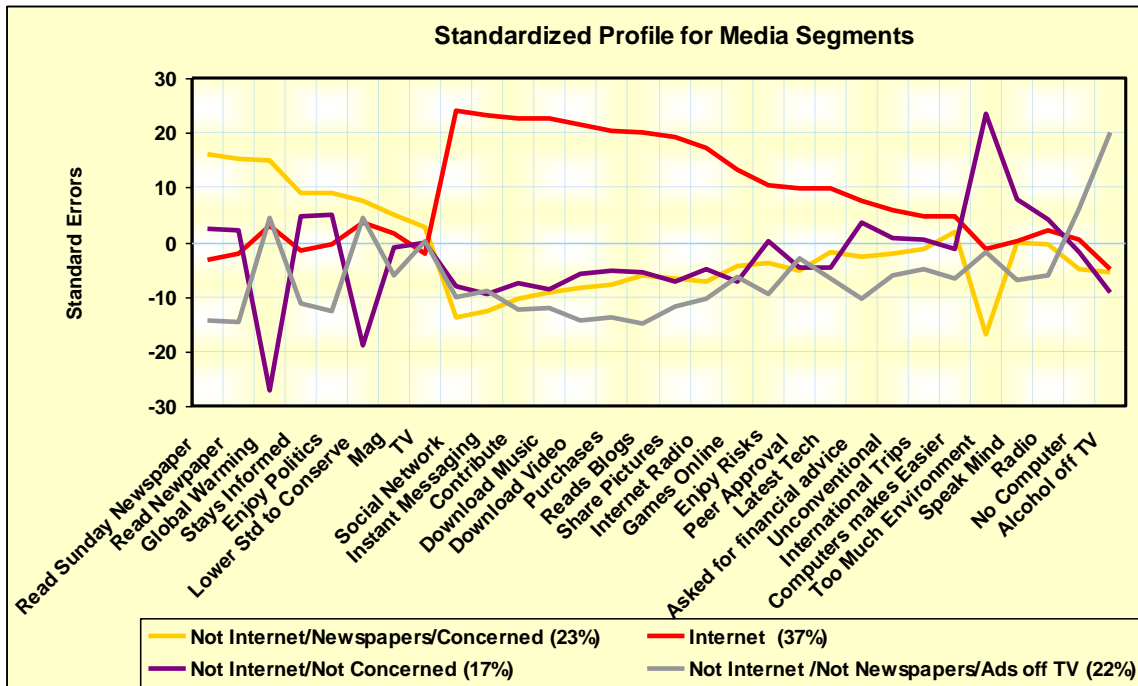
Evaluation of Tiburon - Russia Data Source

Below is the comparison between the sociographic segment distributions for Tiburon - Russia and the Russia Grand Mean. These distributions are not significantly different. Based on χ^2 test, there is less than a 95% chance that the distribution of sociographic segments for Tiburon - Russia is the same as for the Russia Grand Mean. The largest difference is 3% relative change with the Concerned/Not Social Networked segment. Note that the average relative difference is 1.6%.



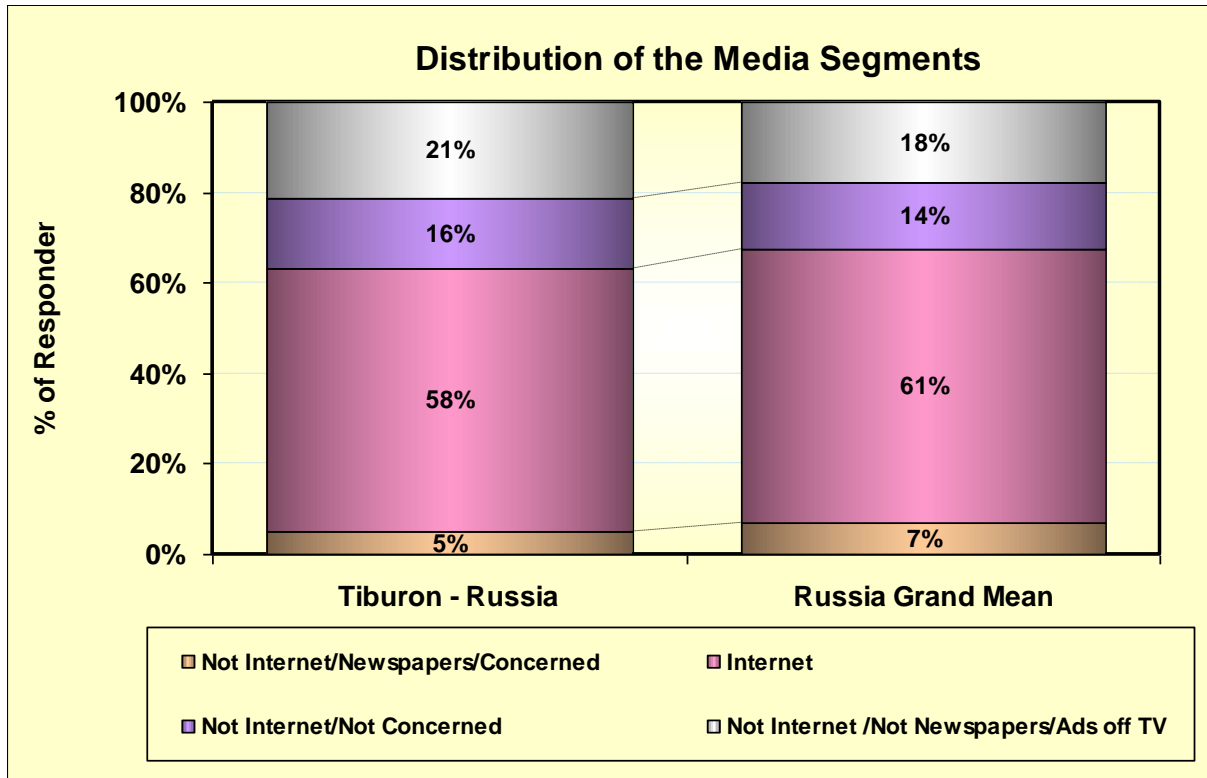
3.3. MEDIA SEGMENTS

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 Tiburon - Russia Data Source

Below is the comparison between the media segment distributions for Tiburon - Russia and the Russia Grand Mean. These distributions are border-line different. Based on a χ^2 test; there is 11% chance that the distribution of media segments for Tiburon - Russia is the same as for the Russia Grand Mean. The largest difference is 25% relative difference with the Not Internet/Newspapers/Concerned segment. Note that the average relative difference is 13.8%.



4. SOURCE PERFORMANCE

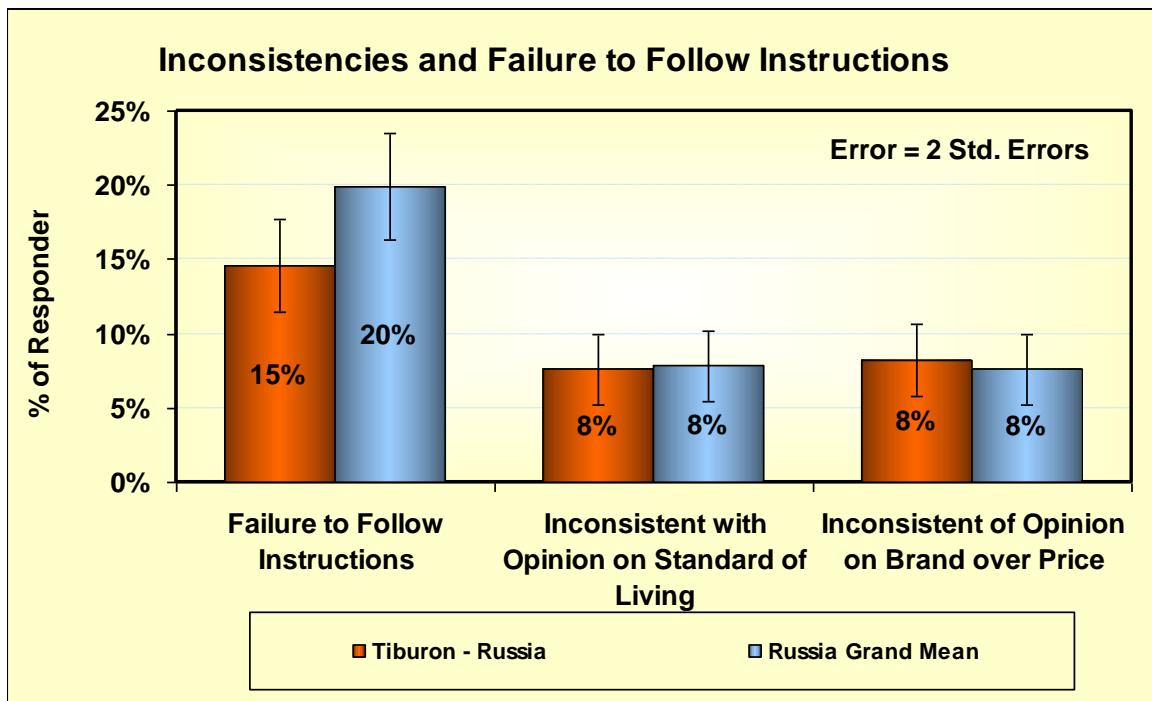
There are several types of panel performance information that might be useful. These reflect the nature of the data source and its functional performance. These are often structural issues reflecting the mechanism of generating and encouraging participants. While each of these measures reflects the potential of respondents generating erroneous responses, each one in isolation may only be a random error. It is in the incidence of multiple errors that provides insight into the potential of erroneous responses. This is explored with Quality Segments based on the number of errors that each respondent does.

- **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.
- **Satisficing** – Respondents occasionally show extraordinary characteristics. These are not errors, just extreme behavior which provides a potential warning of problems. These include: (1) “speeders” who finish their questionnaire in extraordinarily short time and (2) “straight-liners” who tend to give the same answer to a large number of questions.
- **Quality Segments** – The Quality Segments are based on the number of noted issues including errors, indicating “professional” behavior, or showing deviate behavior. As previously noted, the quality segments are used to identify the potential for erroneous responses.

4.1. PERFORMANCE

The quality of results for Tiburon - Russia compared to Russia Grand Mean 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 a more specifically required action that the failure to do so 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, significantly different for Tiburon - Russia compared to the Russia Grand Mean.

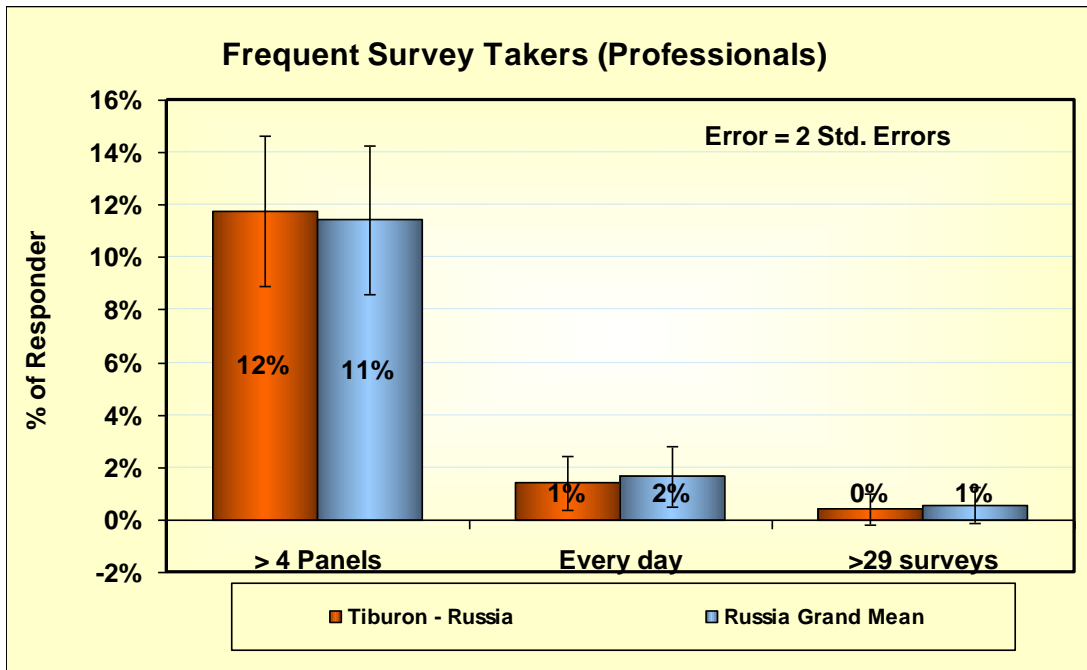


The other two measures capture inconsistent responses: happy/unhappy with standard of living and brand over price/price over brand. As noted above, these measures are based on having given either strong positive or strong negative values to both directions. As seen above the frequency of the inconsistency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

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) belonging to 5 or more panels, (2) taking surveys almost every day, (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 Tiburon - Russia and the reference of the Russia Grand Mean.

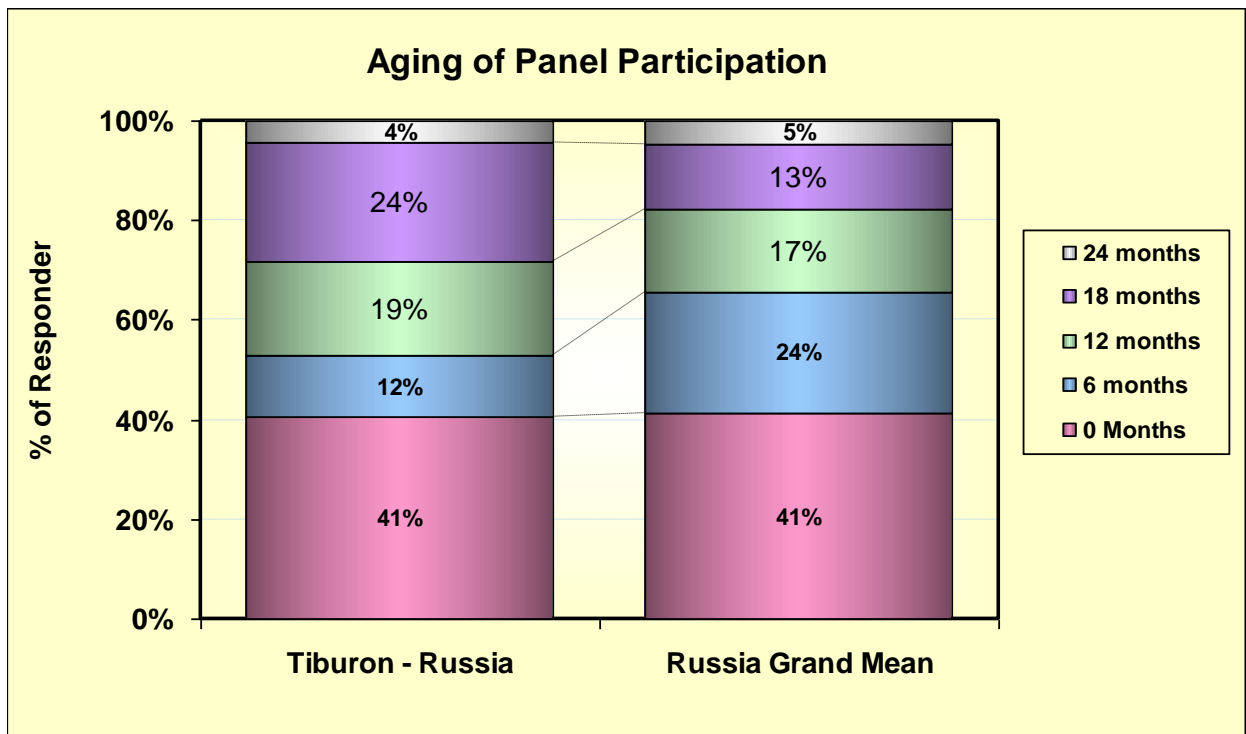
In regards to being in multiple panels, the frequency of Tiburon - Russia members appear to be not significantly different and higher than the Russia Grand Mean. As for taking surveys, the frequency of Tiburon - Russia members appears to be lower and not significantly different from the reference. Similarly, the frequency of taking 30 surveys or more a month was not significantly different and lower for the source than the reference.



Evaluation of Tiburon - Russia Data Source

The complementary issue of belonging to other online panels is the concern over how long members remained on panels. This is viewed as an aging or acculturation problem. There is evidence that changes in panel members' tenure can cause shifts in data.

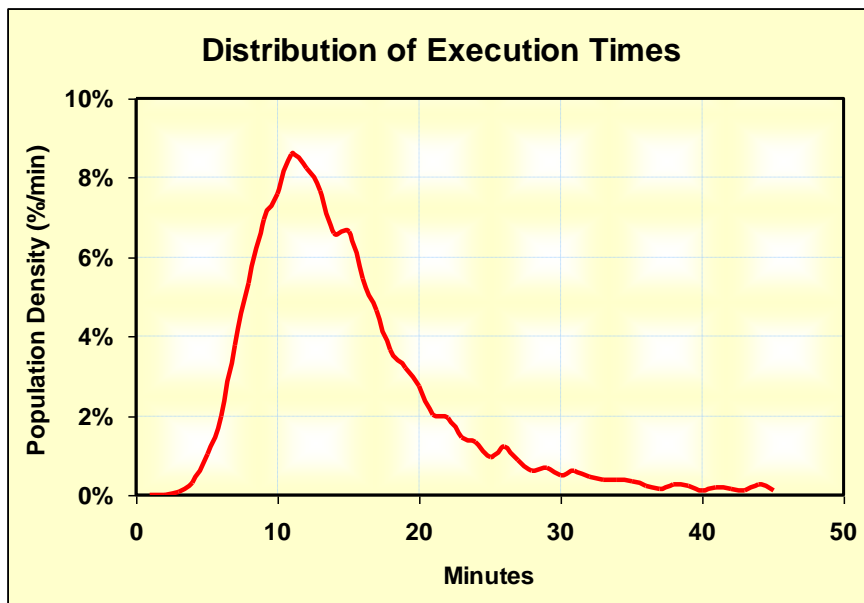
Below is the comparison between the aging of panel participation distributions for Tiburon - Russia and the Russia Grand Mean. These distributions are significantly different. Based on a χ^2 test there is 0% chance that the distribution panel tenure for Tiburon - Russia is the same as for the Russia Grand Mean. The largest difference is 87% relative change in the 18 months group. Note that the average expected difference is 3.3%.



4.3. SATISFICING 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 can be viewed as potential satisfiers. However, once again, there is no direct evidence that such behavior results in poor quality survey results.

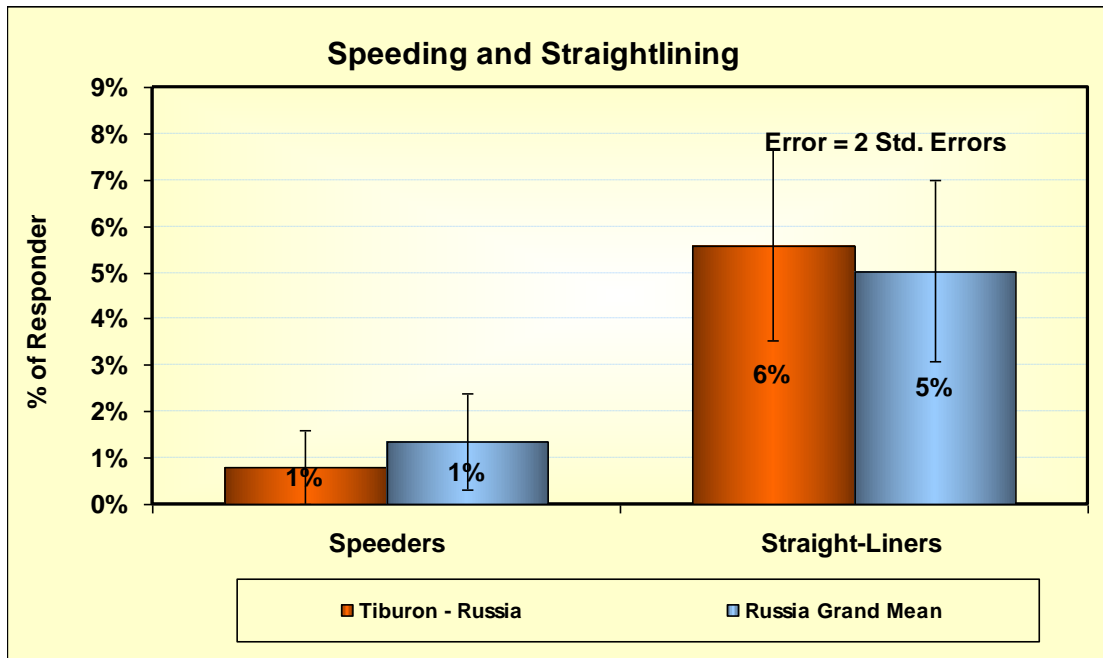
Below is a typical distribution of completion times for the test survey. Note that it has been truncated at 45 minutes. Because of the nature of online surveys, participants may delay execution and thereby run up huge apparent elapsed times. Typically these long times are removed for analysis. Speeders 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 Tiburon - Russia Data Source

The frequency of speeders of Tiburon - Russia members appears to be borderline different from the Russia Grand Mean. The frequency of straight-liners appears not to be significantly different for the source than for the reference.



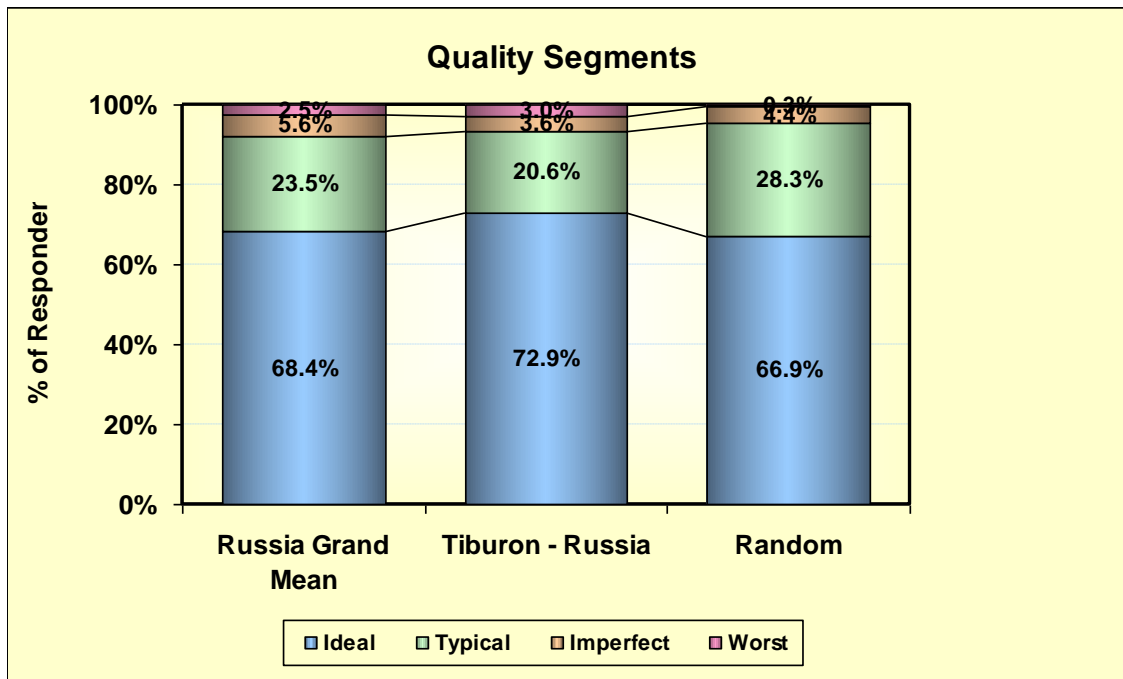
4.4. Q-METRICS™

As previously noted the Quality Segments are based on the number of noted issues including errors, indicating “professional” behavior, or showing satisficing behavior. There are six indicators in this estimation: three performance measures, one measure of professionalism, and the two measures of satisficing behavior. Four segments are used corresponding to: (1) no error (Ideal), (2) one error (Typical), (3) two errors (Imperfect), and (4) three or more errors (Worst). In this context, it is the Worst segment which is of the greatest concern since it represents those who are most likely to give erroneous responses.

Below is the comparison between the Quality Segment distributions for Tiburon - Russia, the Russia Grand Mean and the expected random distribution. As previously noted, each indication in isolation may be a random error and therefore the distribution of errors would follow some type of independent, random, distribution. While this would imply that the occurrence of the errors are not linked and are independent, it still does not imply that those individuals who make up that group are not prone to produce erroneous results. Only that, it is not an unexpected large group.

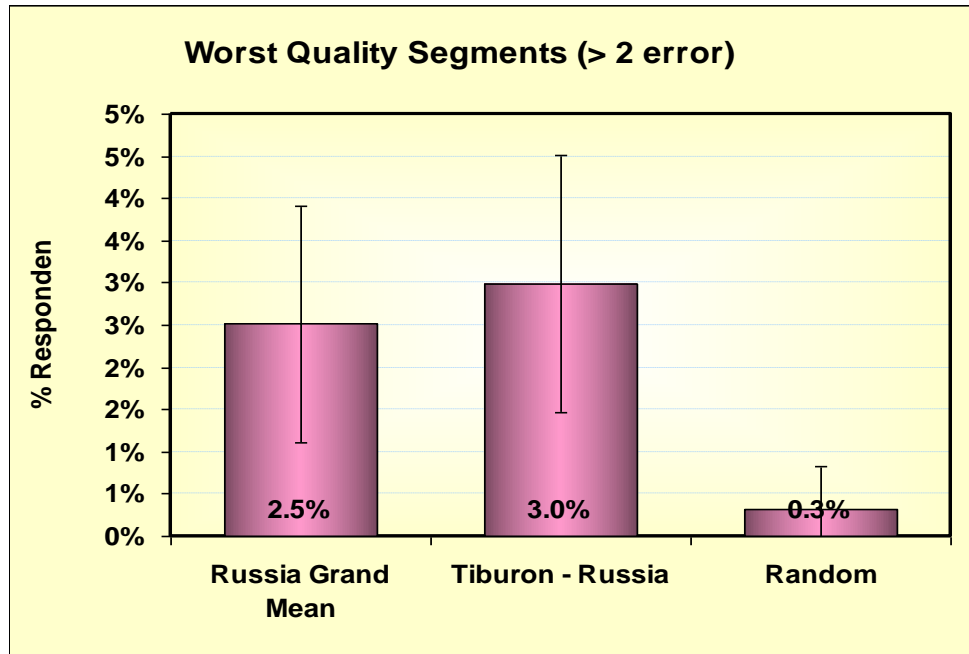
The Tiburon - Russia and Russia Grand Mean the distributions are significantly different. Based on a χ^2 test, there is a 12% chance that the distribution of the Quality Segments for Tiburon - Russia is the same as for the Russia Grand Mean. The largest difference is 35% relative change with the Imperfect segment. Note that the average relative difference is 67%.

The Tiburon - Russia and the expected random distribution are significantly different. The χ^2 test yields a 0% chance that the distribution is the same as the random process.



Evaluation of Tiburon - Russia Data Source

Below are the frequencies of the Worst quality segment. Notice the relatively large error bound. This is due many to the relatively small incidence of multiple errors.



SAMPLE SOURCE AUDITORS™
A DIVISION OF **Mktg.**
Incorporated

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