Variance between purchasing behavior profiles in a wide spectrum of online sample sources.

Authors: Steve Gittelman and Elaine Trimarchi

We begin with an essential assumption: *“**Our data is used to advise clients on how to best make their business decisions**”*. Critical to those decisions is not the demography of certain categories of “problem” respondents but rather how those respondents influence conclusions drawn about buying and purchasing behaviors.

It benefits the research industry to have a better understanding of its sampling resources. Online access panels are a critical element of that industry resource. There is a need for transparency in understanding the potential errors in panel research. One category of respondents receiving intense attention is the frequent survey takers or the professional respondents. The concept of a professional respondent focuses on three characteristics: (1) they readily remain on panels for extended time periods, (2) they complete surveys often, and (3) they belong to multiple panels.

There is evidence that respondents who remain in panels over time differ from those who are new to the research process. Weighting has been found to be ineffective. Financial motivations of respondents cause mode effects that are not mitigated by weighting demographics, attitudinal variables and newspaper readership.

Comscore Networks reported in October of 2006 that less than 1% of respondents in the ten largest online survey panels in the United States accounted for 34% of the completed questionnaires. Some report that three-quarters of panel members belong to three or more panels. These multi-panel participants hungrily seek to complete a few interviews a week.

Our study was undertaken to examine these issues and relies on a relatively large number of data sources and panels to determine the variability among them. Major attitudinal differences exist between those who belong to multiple panels, although they can be demographically similar to those in only one panel.

There is a silver lining in that those who belong to multiple panels are by some measures better respondents. They can be more positive towards the process; they are more willing to complete additional surveys, and they view others in the panel process in a positive light. The reported characteristics of multiple panel respondents was refuted by De Wulf (2007) who concluded that “The often used image of lower social class, poorly educated housewife filling out surveys as a complementary source of funding does not seem to be true.” Among many variables tested (age, gender, education, income and professional status) only one appeared to be significant: they had a higher chance of being unemployed or looking for a job.

Research data has migrated to a non-probabilistic frame from a probabilistic frame that we can only long for as a fond memory of the old days. A year ago we witnessed a practitioner publicly confessed that the days of probability sampling were, well, “forget it.” Others have joined the pragmatic chorus refusing to abandon a valued resource under fire, while clinging to a concept that may have dwindled in practicality.

Comparisons between online panels that were created through probabilistic and non-probabilistic methods reveal a higher prevalence of speeders in the non-probabilistic frame. Straight lining, suggestive of inattentive respondents, is far more present in non-probabilistic panel data. More confounding is that substantial “noise” in non-probabilistic data clouds the interpretation of research results. Bivariate analyses were “masked” making it difficult to discern important relationships. Recommendations to clients are confounded by clouded interpretations.

Attempts to purge the data by removing suspect respondents from non-probabilistic data put the whole data set into question. It does not appear to be enough to remove speeders and other forms of inattentive; the remaining data still lacks precision.

METHODS

As our intent was to understand the differences in purchasing behavior between data sources we were privileged to obtain cooperation from sixteen different providers ranging from river to panel to phone in the United States, and to liven up the comparisons, one panel in the United Kingdom. Two companies participated twice and a third five times.

Data was collected from December 2007 to December 2008 at our offices in Long Island, New York. The survey instrument was approximately fifteen minutes in length. A total of 7,600 interviews were collected, about 400 completes per cell. Quotas by ethnicity, income, gender and age were set by completed interviews to replicate the US census. In the United Kingdom income was replaced by social grade. Telephone data collection was performed at our facility on Long Island, New York.

Respondent Typologies

Here we investigate four different respondent categories:

1. ***Failure to follow instructions***: Respondents failed to answer particular questions with designated answers.
2. ***Inconsistencies:*** We asked a few questions that provided statements that appeared mutually exclusive such as: “Brand over Price vs. Price over Brand”.
3. ***Speeders:*** The average length of the survey was fifteen minutes. We chose to designate the quickest ten percent as speeders. Examination of the questionnaire duration curve gave us no rational point to discriminate speeders so we used this rather arbitrary number. Other practitioners have used half of the average as the division point, a no less arbitrary determination.

4. ***Professional respondents***:

- Respondents who reported that they take online surveys "practically every day" representing 25% of the total respondent pool.
- Respondents who reported in an open ended format that they took over thirty online surveys in the past month, representing 15% of the total.
- Respondents who identified on an aided basis $\geq$5 panels that they currently serve as members representing 36% of the total respondent pool.

RESULTS

Panel Sourcing and Respondent Types

Perhaps the most widely discussed measure of panel quality has been the presence of professional respondents (figure 1, panels presented by sourcing subcategory). All three measures were highly correlated (average $\rho = 0.85$) and are thus somewhat redundant. It is here that the variation between panels is overwhelming. Such vast shifts in the frequency of professional respondents imply immense differences between panels. If professional respondents represent different purchasing behaviors it is likely that the panels will also be different by that measure. The root of these differences appears to be related to the sourcing models used by the sample providers.

River" sampling represents our first three sample sources, all waves of the same online sample product.

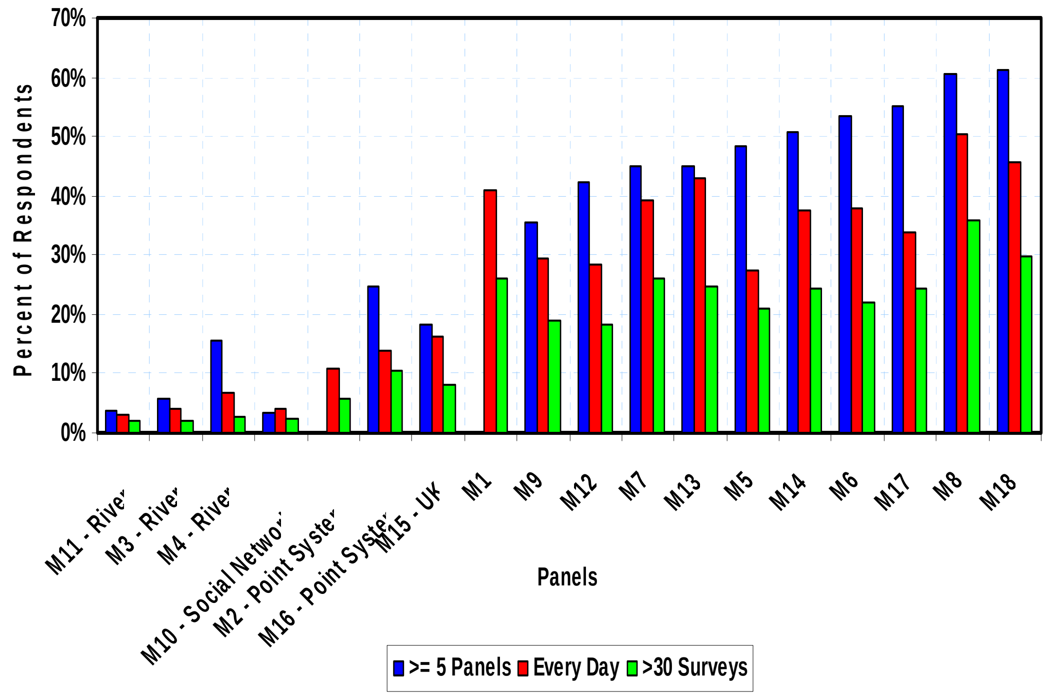
One of the artifacts of River methodology is that repeat responders cannot return with any frequency. Thus we would expect "Rivers" to have a low frequency of professional respondents. M11, M3, and M4 have low professional respondent rates and represent a river source that collects a small amount of data on incoming candidates and exposes them to a small number of (less than three) screeners and no more than two questionnaires. They also block respondents from re-entering the system and do not reuse them as part of their River sourcing model.

Source M10 is a social networking site of the Web 2.0 genre. These sites are non-commercial in that its members do not complete financial transactions as a part of being at the site. Instead the motivation is social exchange. Demographically, the site is younger in age (the provider was unable to complete the older age quotas), better educated, have higher frequency of minors in their households, and are more frequently employed or a student and far less often retired than found in all three professional respondent categories.

Professional respondents are likely excluded from the social networking universe as they do not find it appealing being that they are demographically and probably psychographically adverse. Conversely, those who frequent social networks may find doing an abundance of market research interviews unappealing.

Sources M2 and M16 consist of members who were drawn from point system cultures where certain purchases provide them with a point reward interchangeable with cash or other gratuities; e.g. frequent fliers. We are not speaking here of the incentives that respondents who complete questionnaires receive as a reward.

Figure 1. Professional respondents, sample sources grouped by type[1].



M2 and M16 consist of the most highly educated panels. Their households have fewer minors than other panels. This group consists of the full time employed and the less frequently retired, quite the opposite of the professional respondents.

Inherent in these sources is the ability to make purchases but as professional respondents tend to be less often employed and make few online purchases they are unlikely to participate in such buying communities.

Where the rubber hits the road: Buyer behavior

Data was collected on a total of 27 purchasing related variables (figure 2). Some appear more germane than others, but in total they represent a broad spectrum of questions that give us a first look across panels on this critical measure.

We have summarized the purchasing variables (figure 3), where the percentage is the number of data sources that were significantly different from the mean of all sources, or the "Grand Mean." Included in that mean are the United Kingdom and one cell of CATI interviews. The panels

[1] Panel membership was not asked in M1 and M2.

range dramatically around the purchasing measures. The range in variability peaks at high tech purchases where 37% of the panels significantly differ from the grand mean while there are none who do so around "price over brand".

Figure 2. Buyer Behavior Variability across panels by behavioral measures.

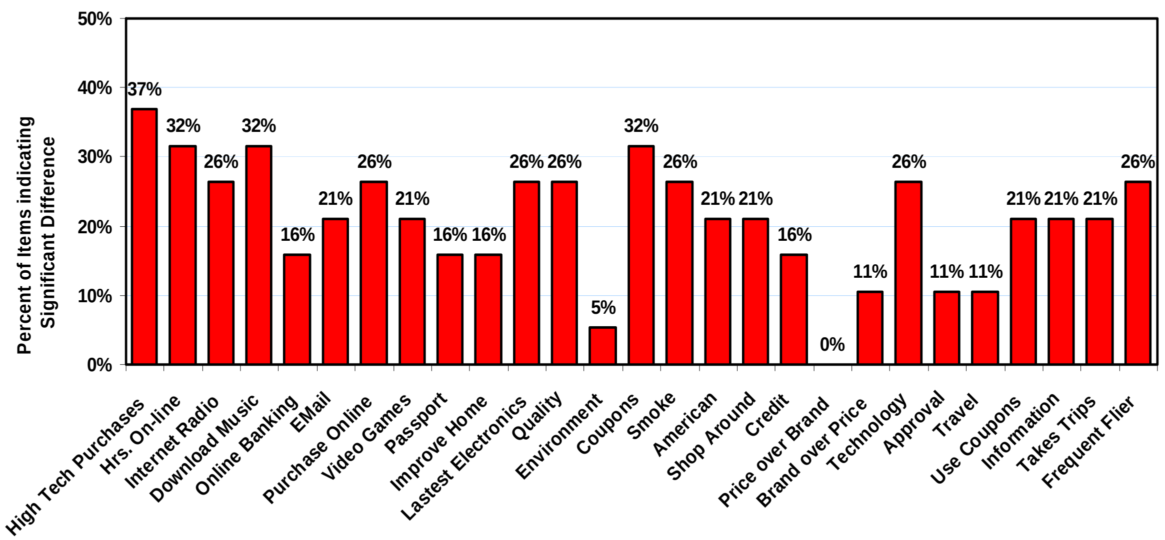
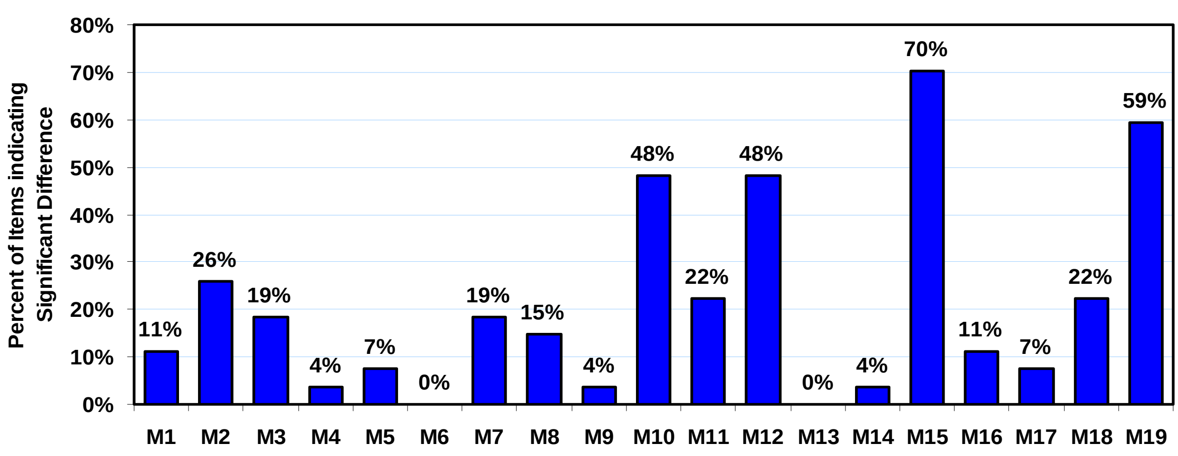


Figure 3. Frequency of significant differences from the grand mean of purchasing data[2].



[2] Significance at three standard errors, where the total n=27 possible buyer behavior measures.

We clustered the data using 27 purchasing measures into five purchasing segments. Table 1 shows the components of the segments in terms of major differences.

The solution provided us with five exclusive segments that reflect combinations of the underlying variables. The five segments appear to vary broadly between panels. Once again our sourcing models seem to impact the behaviors we witness (figure 4). M10, the social network, shows the highest frequency of traditional purchasers, who use credit cards, off line. We have learned to expect the reverse of these internet savvy respondents. They will shop until they drop but are on the web for social exchange not purchasing. It would appear that the social network sites employed by this vender are very different purchasers.

M2 and M16 are apt to be found doing their banking and purchasing on line by credit card as one might expect of those who belong to an on-line purchasing community driven by a point reward system.

The remaining access panels seem rather homogeneous.

When we cast the five segment buying behavior solution against various categories of respondent behavior; patterns of similarity and difference standout (figure 5). Professional respondents might be best understood as price conscious, non purchasers of many of the items we listed, who are frequently online but prefer not to use credit cards. Instead they like to shop around looking for bargains.

Those who failed the validity test, that is, did not follow a simple instruction to enter a particular answer appear strikingly similar to the professional respondents, except they have a higher predilection to bank online and are not as price conscious. Those who expressed an inconsistency on brand vs. price and price vs. brand profess to be off-line traditional purchasers who use their credit cards. Given the choice they would choose brand over price, spend considerable time online, do their banking online and will use their credit cards.

Those who are caught up in an inconsistency over happiness are strikingly similar to the speeders. They are traditional on and off-line purchasers who like the convenience of their credit cards.

Table 1, Buyer Behavior Segment Description

	Description	Internet	Credit	Brand over Price	On-line Banking	Shopper	Scale	
Segment 1	*Purchasers/Credit Cards/ Not On-Line*	blue	red	orange	blue	red	Very High	red
Segment 2	*Shoppers/No Credit Cards*	white	blue	light blue	white	orange	High	orange
Segment 3	*Non-Purchasers/On-line/ No Credit Cards/Price*	orange	light blue	white	orange	blue	Neutral	white
Segment 4	*On-Line/Credit Cards*	white	red	white	light blue	white	Low	light blue
Segment 5	*On-line/Not Price/OLBanking*	red	white	red	red	light blue	Very Low	blue

Figure 4. Buyer behavior segments by panel.

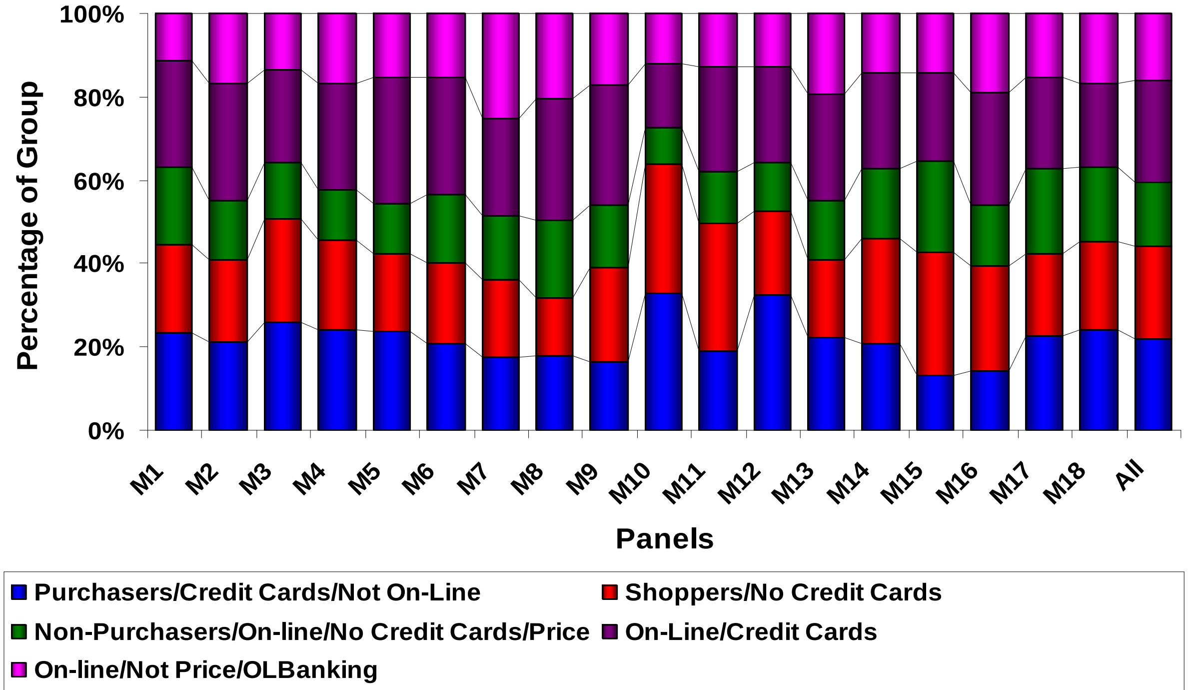
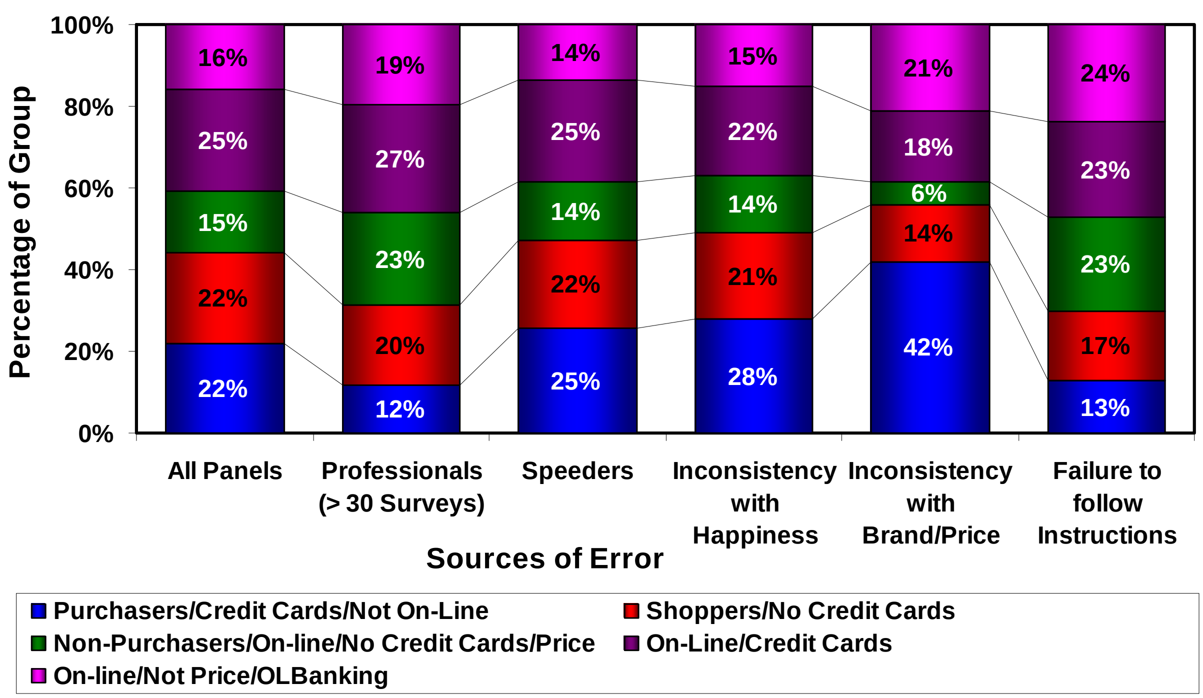


Figure 5. Buyer behavior segments by respondent type.



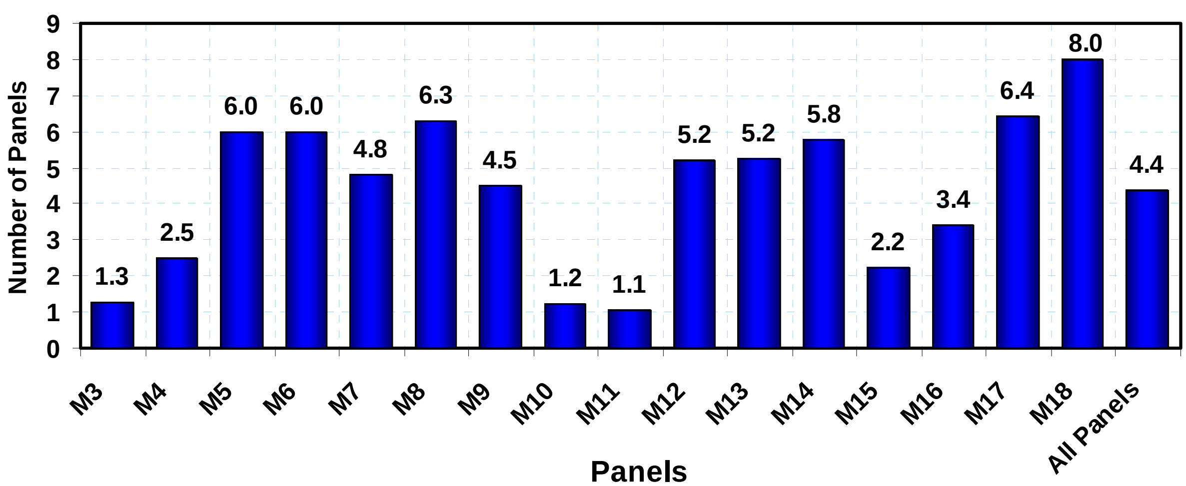
DISCUSSION

Cross panel comparisons are rare. In what must be considered a landmark study, Vonk, Ossenbruggen and Willems (2006) compared 19 of 30 panels then existing in the Netherlands. Their methods were different, e.g. they did not quota control to replicate census, but instead relied on release of a 1000 respondents per participating company.

This data shows considerable variance in the buyer behavior represented between sample providers. Vonk et. al. found substantial differences between panels with response rates ranging from 18% to 77%. Response rates, are a likely consequence of both management and recruiting. Those obtained through traditional research methods such as questionnaires seemed far more responsive than those who are self selected. A prior relationship, such as the completion of a market research interview, potentially drives an emotional reason to respond to an invitation, whereas those who approached the process on their own volition would have little loyalty generated by human connection.

These data indicate that respondents differ dramatically and that sourcing mode is a critical driver. Our measures of professional respondents are particularly telling, one example, the average number of panel memberships (figure 6) ranges from 1.1 to 8.0. Once again, certain sourcing models seem to generate different results: River respondents, M3, M4 and M11 all taken from the same source were 1.3, 2.5 and 1.1 respectively. Our singular social networking site, M10, was once again an outlier at 1.2 panels per respondent. M16 sourced from a point system environment that appears to exclude the demography associated with professional respondents was low at 3.4 and the United Kingdom, a completely different sourcing environment was at 2.2. With these somewhat unique sourcing models eliminated the remaining access panels range from 4.5 to 8, a relatively homogeneous grouping.

Figure 6. Average number of panel memberships per panel[3].



[3] M1, M2 and CATI were not asked.

Vonk et. al. found that panels recruited by telephone, snowballing and traditional research were meaningfully different from more abundant and typical processes of self registration, via website surveys, from existing panels, via banners and by purchase of e-mail lists and subsequent offerings of participation. River sampling, social networks and point system cultures were not identified in their data.

Professional respondents and the cloud that surrounds them seem to drive this situation as the respondent pool within the panels is different from the respondents who actually complete the questionnaires. Here the sourcing differences seem to impact on the very presence of the professional respondents. The essence of the outliers is that they have a low frequency of such respondents.

Casdas et.al. (2006), found that multiple panel members were younger, less educated; more often female; not working full-time; more often worked part-time and were more likely to rent as opposed to own their principal residence. In addition they showed less reliability on attitudinal in contrast to behavioral measures. In many cases our data here agreed.

The impact of these respondents on behavioral data resonated in Casdas et. al. (2006). The predilection of multiple panel respondents to purchase a motor vehicle in the next twelve months correlated with the number of panel memberships. As the number of smokers, home owners, and broadband users increased in their sample so did motor vehicle purchasing intent. Weighting did not reliably cure the situation. They conclude, “Every panel will have its own unique profile of panel membership…”

In the Netherlands, 8% of respondents were deemed to be professional respondents, while in our measure the number appears to be considerably more than that ranging from 37% who belong to greater than four panels; 26% who complete a survey just about every day to and 17% who report completing 30 or more surveys in the past month.

We find that professional respondents are statistically less well educated,. Their incomes are lower, they are less likely employed, their households have fewer young members and their outlook seems less confident.

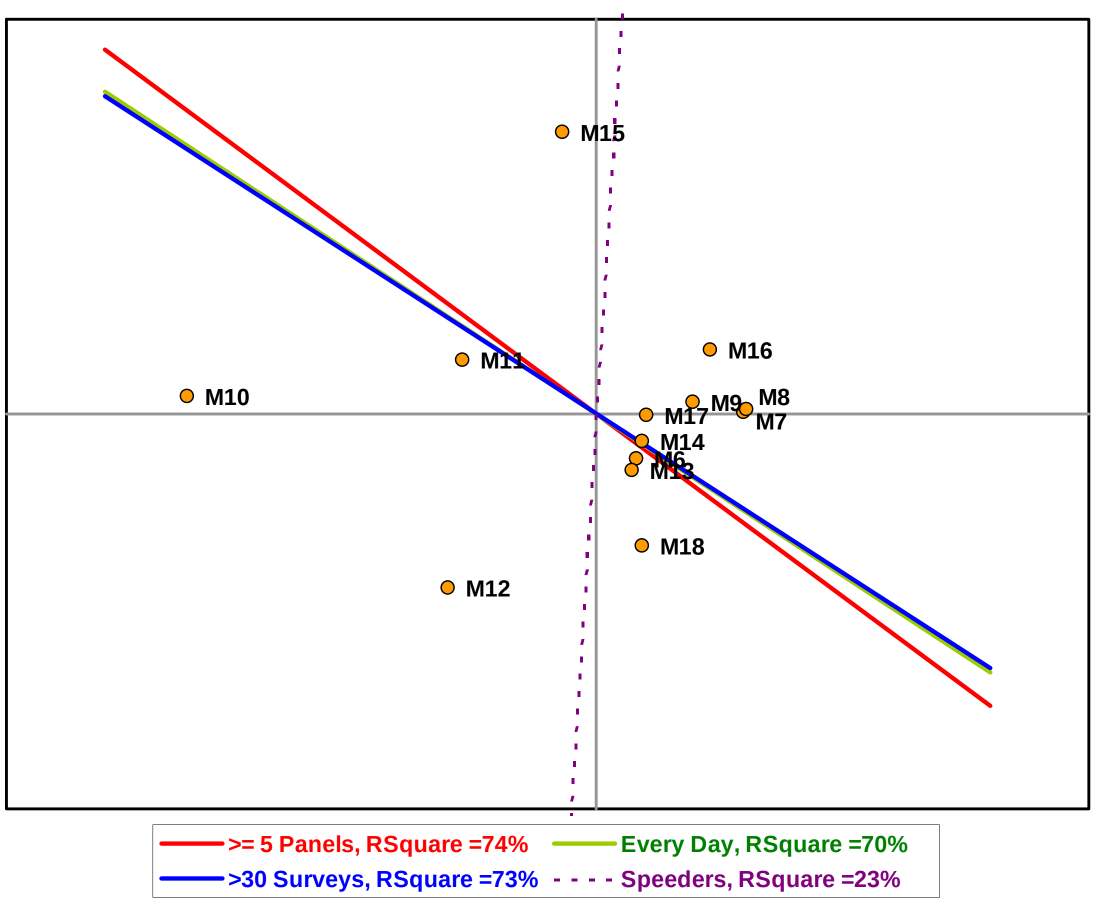
Vonk et. al. also discovered that different respondent characteristics appeared in different frequencies among panels. They attributed the differences to panel recruitment models. “Both professional and inattentive respondents are both much more likely to be approached from purchased databases or by self-selection…” While inattentiveness and professionalism was highly correlated in their results (ρ =0.76, n=19) it does not appear to be in ours (ρ=0.06, n=18).

We calculated the deviations in a matrix, in this case between 13 panels and 27 purchasing questions. This was used as an input for an MDS analysis (Multiple Dimensional Scaling). The resulting coordinates are plotted in (figure 7). Each point on the MDS plot represents a panel. The positioning of the panels to each other translates to a degree of similarity. Thus those panels that are found clustered together would be considered to be most similar. Conversely, those farther apart would be considered to be most different. Once again we find that the sourcing of panels is somehow connected to the buyer behavior/purchasing difference variability. M11, river; M10, social network and M15 are clear outliers. The three professional respondent vectors are highly correlated (R^2 of 0.70 to 0.74, n=13) with panel distribution and provide a connection between purchasing/buying behavior in panels and the presence or absence of these respondents.

In these data, the speeders were significantly correlated ($R^2 = 0.23$) but not to the degree of the professional respondents.

We subjected the data to a principal component analysis to obtain a different analytical perspective, particularly on the speeders. In each measure professional respondents explained the majority of variability in buyer behavior represented by the panels ($R^2 > 36\%$). In contrast, speeders appeared to explain little of the distribution with R^2 approaching 0.4%. There is a considerable body of research that links speeders with professional respondents as a quality issue. However, here they are de-linked.

Figure 7. MDS Plot. Variation between panels appears attributable in large part to the presence of professional respondents[4].



[4] Data in this case is based upon 27 purchasing attributes.

CONCLUSION

Panels are a diverse community. Sourcing models have the ability to drive the presence or absence of various respondent constituencies; particularly professional respondents. These avid survey takers can cause shifts in the purchasing behaviors that typify various sample sources. Sourcing models that aggregate demographic and psychographic concentrations of respondents likely to take frequent surveys will exhibit one set of buying/purchasing profiles, while those that lack broad representation may differ for other reasons. The old adage, "do not put all your eggs in one basket" holds true for market research sampling today. Researchers must learn to use multiple sources just as they diversify an investment portfolio.

REFERENCES

Casdas, Dimitrio; Brian Fine, Con Menictas. 2006. Attitudinal Differences: Comparing people who belong to multiple versus single panels. Panel Research 2006. ESOMAR World Research Conference.

De Wulf, Kristof; Sam Berteloot. (2007) Duplication and multi-source online panel recruitment: real quality differences or idle rumours? ESOMAR World Research Conference, Panel Research 2007: good quality, good business, Orlando, 28-30 October 2007, p. 49-62.

Vonk, Ted; Robert van Ossenbruggen, Pieter Willems., (2006). The effects of panel recruitment and management on research results. A study across 19 online panels, ESOMAR Panel Research Conference, Barcelona.

ACKNOWLEDGMENTS

We are particularly grateful to Gene Lieb at Custom Decision Support, Inc., for analytical advice.

About the Authors:

Steven Gittelman is President of Mktg, Inc.
Elaine Trimarchi is Executive Vice President of Mktg, Inc.