Assessing the Research Methodology, Validity, and Representativeness of CivicScience Survey Data

Dr. Lamar Pierce
Associate Professor of Strategy
Olin Business School
Washington University in St. Louis

Dr. Aaron Chatterji
Associate Professor of Strategy
Fuqua School of Business
Duke University

Dr. Jason Snyder
Assistant Professor of Strategy
Anderson School of Management
UCLA

Dr. Alessandro Acquisti
Associate Professor of Information Technology and Public Policy
Heinz College
Carnegie Mellon University

June 2012

SYNOPSIS

The following paper provides an overview of the data collection methodology employed by CIVICSCIENCE (CS) to gather, organize, and analyze public opinion data. The main objectives include:

1- To explain the CS data collection, quality assurance, and reporting techniques
2- To examine the relative representativeness of CS respondent demographics compared to known demographics of land line households and the US Census
3- To highlight specific examples where CS data has been used to mimic, predict, or improve the capabilities of conventional consumer research methods

CONCLUSION: It is the finding of the authors, upon rigorous assessment of the company's approach, that CivicScience has developed an effective, defensible, and accurate model for gathering and analyzing public opinion in the U.S.

DATA COLLECTION TECHNIQUE

CS manages a network of web-based polling applications distributed across third-party websites, social media assets, mobile applications, and a proprietary web portal to engage consumers ("Respondents") in attitudinal research. A strong consideration in the recruitment of participating web assets is the demographic and geographic composition of its visitors.

CS polls deliver three questions to respondents during each session, including two attitudinal questions, the first designated as an "Engagement" question, designed to compel respondents to participate, a second "Value" question, designed for commercial purposes, and a third demographic or "Profile" question. Respondents receive no monetary incentive and can only view the results after completing the poll.

On a daily basis, CS distributes a library of poll questions across that day's respondent base to achieve an approximately equal number of answers to each question every day. This library includes questions about current events, pop culture, local and national brands, products, media properties and personalities, technology usage, social and policy issues, political topics, lifestyle, and personality traits. Custom questions on behalf of CS customers, delivered to the general population or a targeted subset, may also be included in the daily library.

As a default, the two Value questions for a given respondent are randomized to prevent fraudulent voting and to ensure equal distribution of responses over time, unless a respondent is designated for a specific targeted question. The third, Profile question is delivered in priority order, beginning first with gender or age, until both are completed, then six additional demographic questions until each is completed, and finally a larger list of 140 profile questions delivered at random until each is completed.

Respondents are identified in the CS database by a unique, anonymous digital alias including a browser cookie, IP address, and other available credentials including a Facebook Connect ID or a unique User ID from a logged-in website. This digital alias enables CS to: 1) Identify unique respondents over multiple visits and across multiple sites; 2) Append responses from multiple polls to build a longitudinal profile of each respondent; 3) Ensure that no respondent answers the same question more than once (unless deliberately re-asked); 4) Target questions to a specific respondent based on profile-type or prior answers; 5) Identify the geographic location of each respondent.

DATA QUALITY: FREQUENTLY ASKED QUESTIONS

How does CS prevent respondents from "stuffing the ballot box" or, in other words, organizing biased respondents to produce an artificial measure? At any given time, there are a minimum of 2,000 active poll questions being deployed across websites that host CS polls. Given that "Value" questions are delivered at random to each respondent, the likelihood of a specific question appearing in a specific poll is roughly 1 in 2,000. A respondent must answer the initial "Engagement" question, check for their desired question in the Value slot, refresh the page if their desired Value questions does not appear, then complete this cycle up to 1,400 times until they find the question they were looking for. Any unusual page-loading behavior of this nature would be quickly flagged by the CS web servers and the offending respondent would be pruned from the results.

How does CS prevent and/or identify lying or misrepresentation by respondents? First, CS attempts to discourage lying by creating intrinsic incentives for voting honestly. The Detailed Results interface enables respondents to see how they compare to peers and to learn interesting insights about themselves. These features are devalued if the respondent answers inaccurately. Secondly, CS uses a series of algorithms to identify conflicting combinations of answers in a respondent's profile based on high-entropy relationships of answers. For example, consider an individual who in one poll identifies herself as under 18 years of age and then, in a subsequent poll, claims to have a 401k retirement account. The CS system would flag this individual in the database as a potentially-fraudulent respondent. Future questions could then target that individual with a validating question or prune them from the results. This same logic is also used to identify multiple users on the same computer or cookie.

How does CS account for respondents who regularly clear their cookies or otherwise do so deliberately to vote multiple times on the same question? The CS reporting interface only includes respondents in weighted or "formal" results if they answer a minimum of six poll questions, including both age and gender. As it pertains to fraudulent voting, this is first mitigated by the measures outlined in Question One above. Combined with the Age/Gender requirement, this means that a fraudulent respondent would need to complete the refresh/find/answer cycle two times for every one vote they would contribute to the formal results. Again, the CS web servers would detect this behavior automatically.

DATA REPORTING

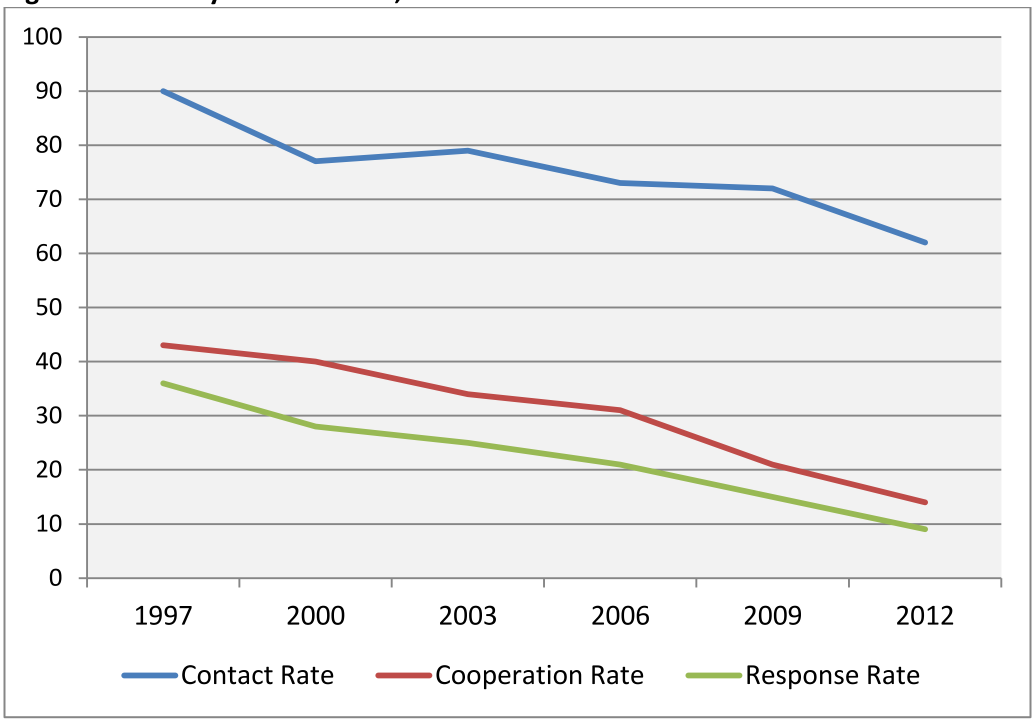
The CS data reporting interface provides access to real-time results as they are collected, with tools for Census weighting by age and gender, cross-tabulation, margin of error calculations, time-trend analysis, and data exportation. Results can be sorted by geography and/or time.

Because each respondent in the CS database has answered varying combinations of questions over time, data mining algorithms explore the data to find groups of respondents who have answered specific combinations of queried questions. When questions are selected for cross-tabulation, the reporting interface runs a standard Chi-Squared Test to determine whether or not there is an association between the questions, and then Tschuprow's T is calculated to measure the strength of correlation. To ensure large enough sample sizes to achieve reliable cross-tabulation figures, CS recommends a minimum single-question sample of 1,000 respondents for analysis

COMPARING CS SAMPLE DEMOGRAPHICS TO LANDLINE PHONE SURVEYS

In a 2012 Report[1], the Pew Foundation revealed that only 9% of US households respond to telephone surveys, compared to 36% in 1997. Similarly, survey contact rates – the percentage of households in which an adult was reached – and cooperation rates – the percentage of households contacted that yielded an interview – have been declining steadily as well, as Figure 1.0 shows.

Figure 1.0: Survey Success Rates, 1997-2012



Further, a December 2011 report from the National Center for Health Statistics[2] found that only 68.4% of US households own landline phones. With 9% survey response rates among 68.4% of households, this means that typical landline-only phone samples are reaching only 6.15% of the adult population.

Not only is CS able to reach a large proportion of the population relative to traditional, phone-based surveys, but the raw demographic composition of its respondents is superior to that of typical landline sample frames as well. The following charts compare the demographics of United States landline households to the national CS respondent population and, finally, the baseline demographics obtained from the US Census[3].

[1] "Assessing the Representativeness of Public Opinion Surveys"
[2] "Wireless Substitution: Early Release Estimates From the National Health Interview Survey, January-June 2011"
[3] National landline numbers come from the November, 2010, Pew Research Center report titled "The Growing Gap between Landline and Dual Frame Election Polls."

Figure 2.0: CivicScience vs. Landline Phone Respondents by Gender (National)

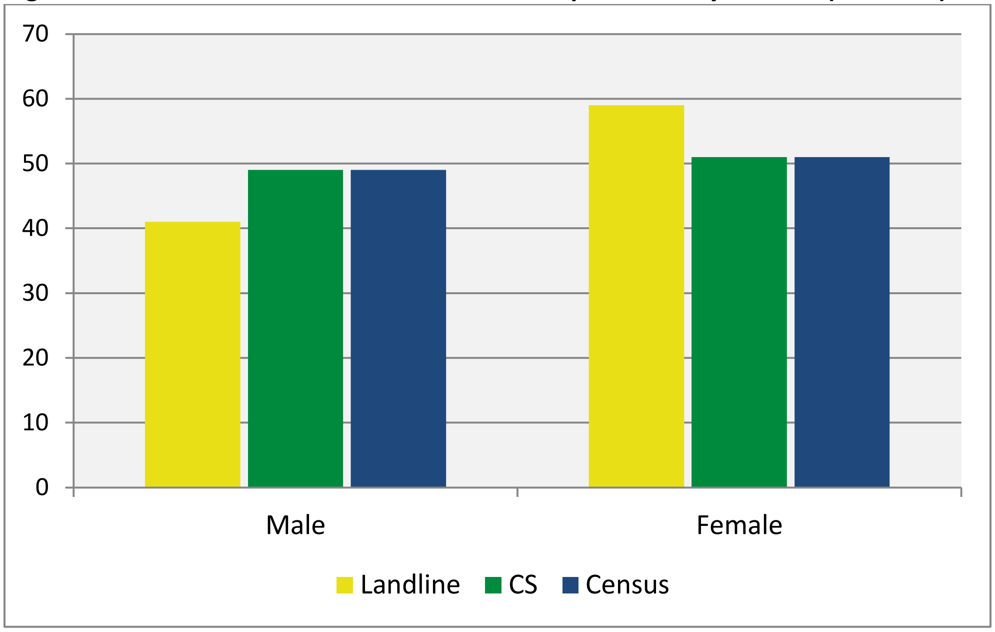
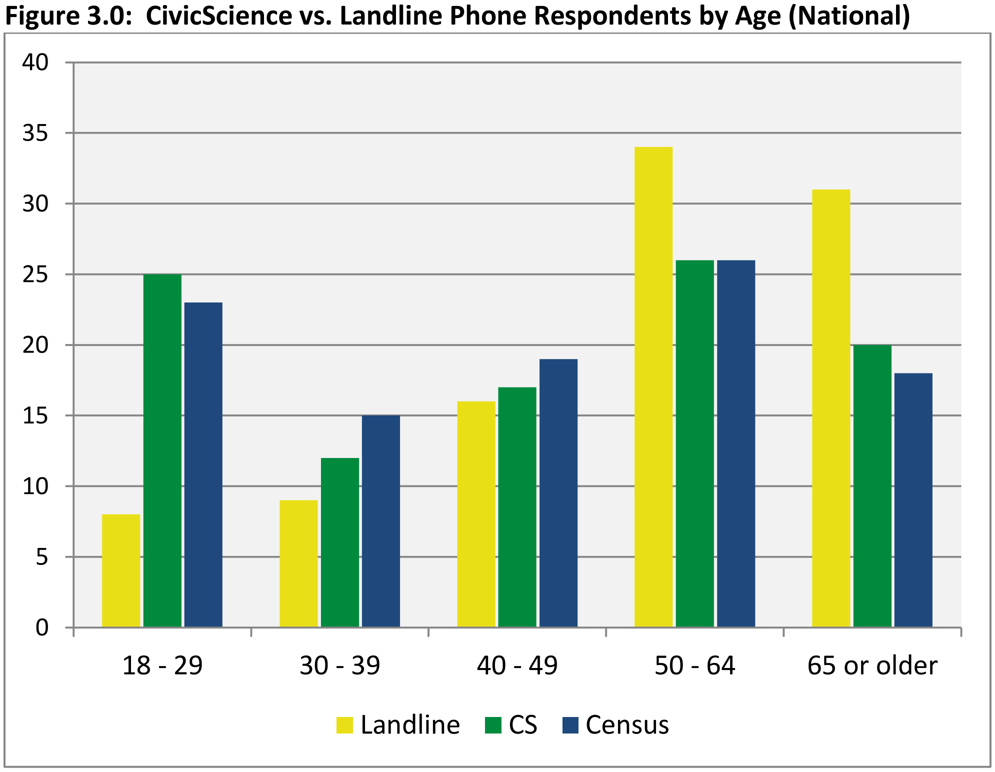


Figure 2.0 shows that landline households (or at least respondents to polls conducted via landline phone) are significantly biased towards women. The gender distribution of CS respondents nationally is virtually identical to that of the Census.

Figure 3.0: CivicScience vs. Landline Phone Respondents by Age (National)



In Figure 3.0, we see that the CS respondent population is closer to the real demographics of the Census for every single age group. Among 18 to 29 year-old respondents, an age group notoriously unreachable by landline telephones, CS slightly over-samples this hard-to-reach population.

Figure 4.0: CivicScience vs. Landline Phone Respondents by Race (National)

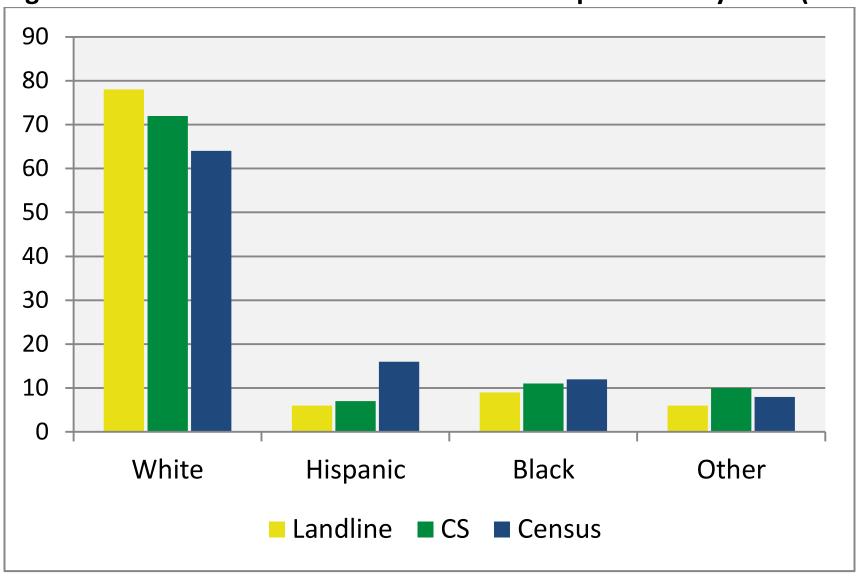
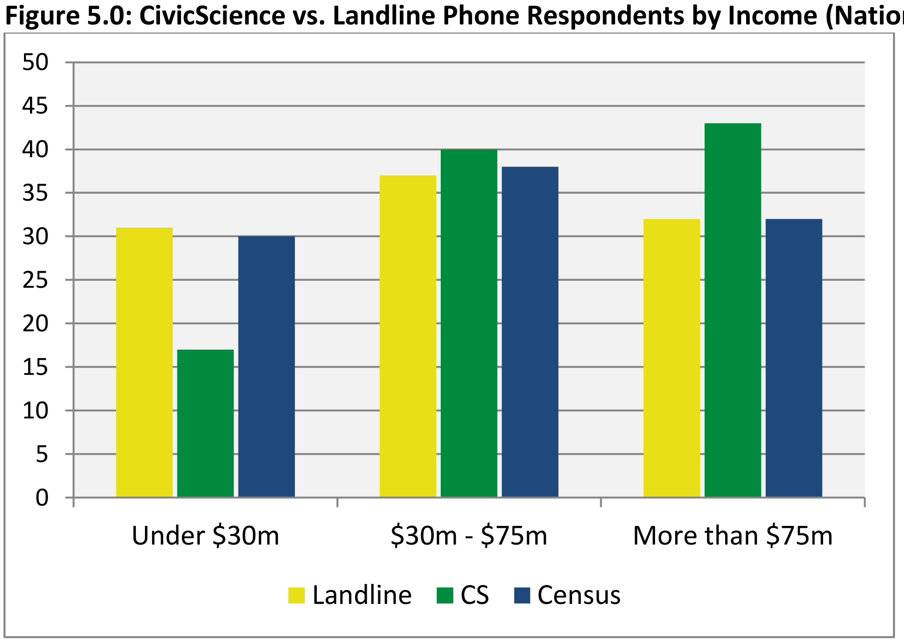


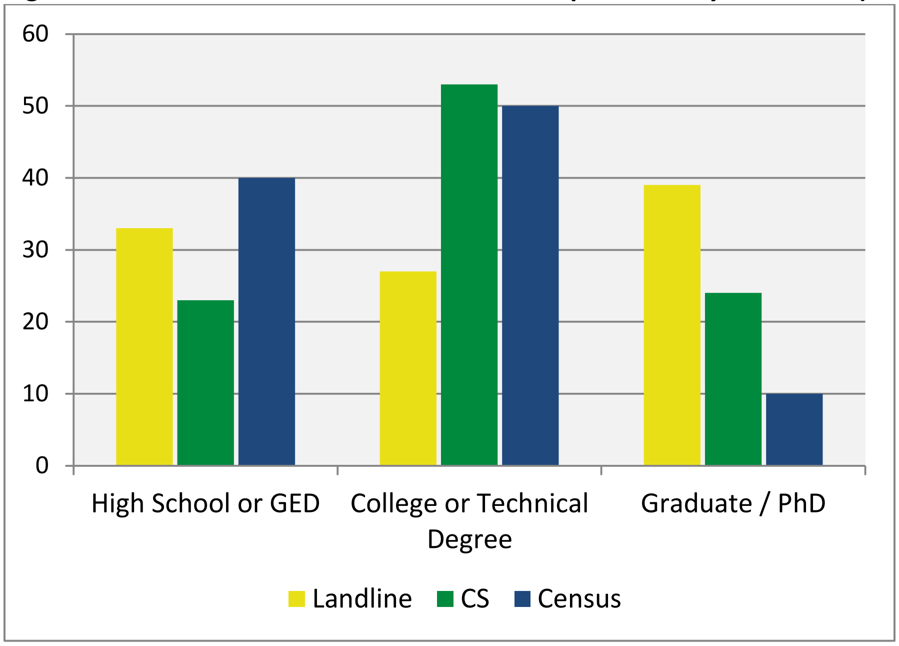
Figure 4.0 shows that the CS respondent population is closer to the real demographics of the Census in every Race category except for "Other," where both CS and landline homes differ by two percent.

Figure 5.0: CivicScience vs. Landline Phone Respondents by Income (National)



In Figure 5.0, we see that the CS respondent base is not as representative of the Census in terms of expected annual income as are landline households. This variance is likely attributable to landline surveys' over-sampling of "65 or Older" respondents, where retirees claim significantly lower income numbers.

Figure 6.0: CivicScience vs. Landline Phone Respondents by Education (National)



In Figure 6.0, we see that neither the CS respondent base nor landline users are highly representative of the Census on the basis education. For higher levels of education, CS more closely matches the Census, but is under-representative of those who have earned a GED or high school degree.

It is important to note that CS has a considerable advantage over phone-based polling in terms of sample size. For example, a phone poll with a sample size of 800 will struggle to reach a substantial number of Hispanic citizens who have a landline phone and will respond to a survey. Yet even though CS may also under-sample Hispanic citizens to a lesser degree than landlines, CS reaches a higher number of respondents (approximately 243,000 Hispanic respondents in CS database), resulting in less overall error after reweighting.

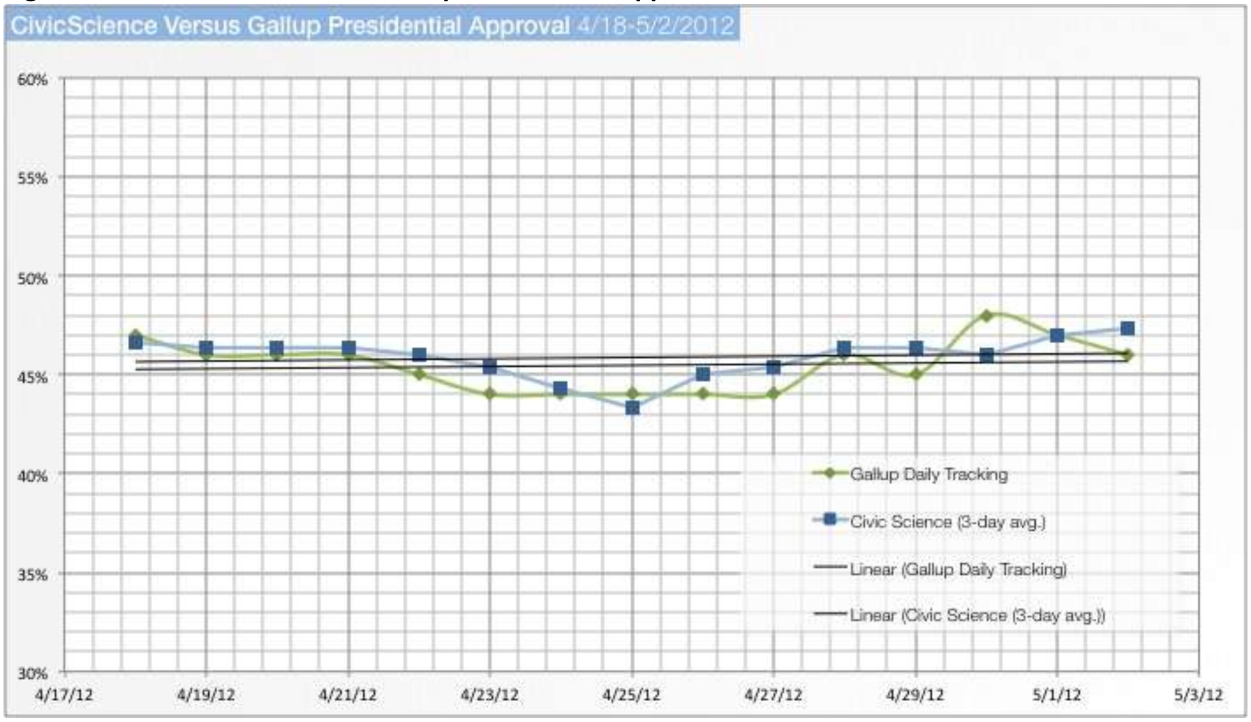
BENCHMARKING AGAINST PREVAILING BEST PRACTICES

CS has worked closely with leading political pollsters, marketing research scientists, economists, and industry executives to extensively test the reliability of our data collection model and its ability to predict events such as local political elections, product sales, movie box office performance, and consumer reaction to price increases. We are happy to share case studies in these areas upon request. Highlighted examples have been provided below:

Example 1- Political Sentiment

One recent example looks at a comparison between CS data and the esteemed Gallup tracking report, which measures the approval rating of President Obama on a daily basis. The graph below shows daily numbers from April 18th through May 2nd. We regularly collect approximately 1,245 responses daily to our Obama approval question (or 18,683 during the two-week timeframe below). Our daily scores, like Gallup's, were based on a three-day rolling average.

Figure 7.0: CivicScience vs. the Gallup Presidential Approval Tracker



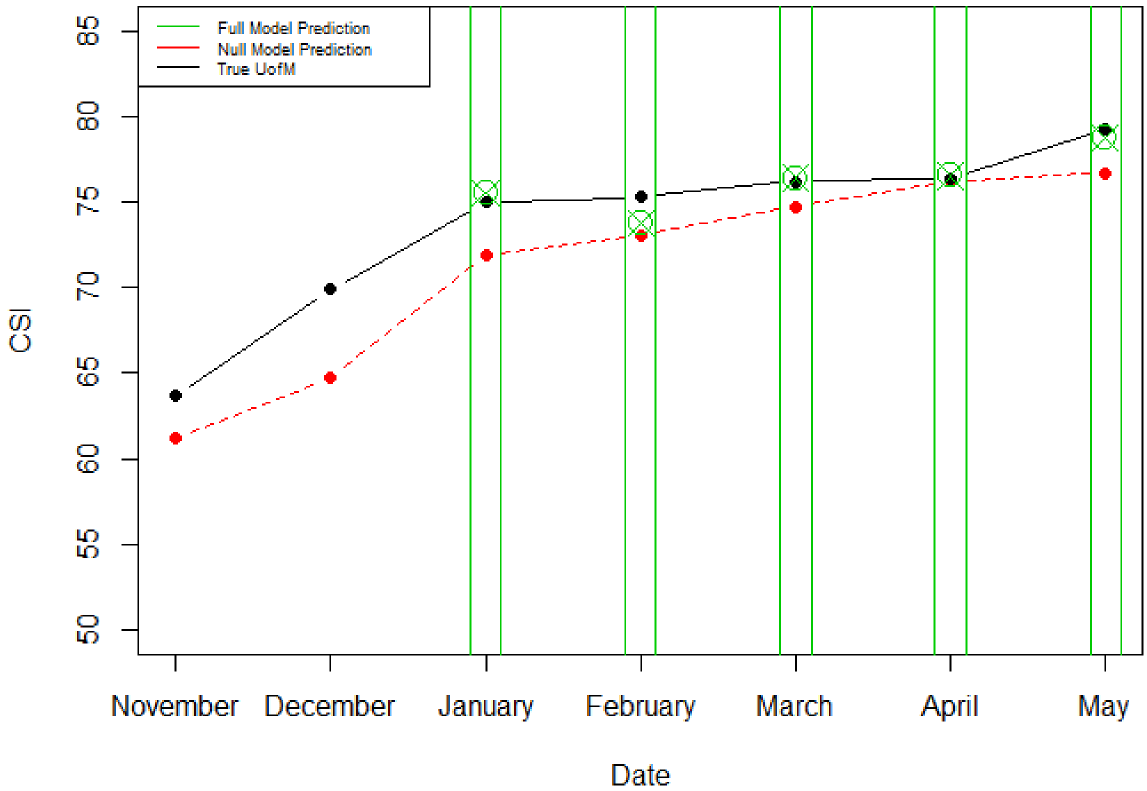
With very little reweighting (gender and age), CS numbers are able to closely mimic the Gallup measures. Despite, some daily fluctuations, it's worth noting that our linear score during the two-week time period was exactly the same as Gallup's --- 46%. And, on no single day did the CS measures vary beyond Gallup’s stated margin of error (+/- 3%).

Example 2- Consumer Sentiment

In June, 2011, CS began asking consumer confidence-related questions, with the goal of tracking the University of Michigan Consumer Sentiment Index (Michigan CSI). The Michigan CSI is based on a phone survey of 500 US households. Based on this monthly survey, the University publishes a numeric estimate of the nation’s consumer confidence level. The Michigan CSI is a highly regarded survey, and its results are closely scrutinized by investors. Prior to the Michigan CSI’s release each month, scores of economists and market analysts attempt to predict the final number. Whenever the Michigan CSI estimate comes in significantly higher or lower than the market consensus, stock prices swing dramatically.

Throughout the early months of 2012, CS collaborated with the Carnegie Mellon University’s Graduate Statistics Department to devise a model to track and predict the Michigan CSI using CS data. Figure 8.0 below shows the model’s ability to track the index through the first five months of 2012 The black line represents the real monthly Michigan CSI number, while the green “targets” represent the CivicScience prediction. The red line represents a predictive model that does not include CS data.

Figure 8.0: Tracking the University of Michigan Consumer Sentiment Index



As can be seen in the chart above, CS results were able to pinpoint the eventual Michigan CSI in every month except February 2012, where our estimate was still within Michigan's margin of error.

Example 3- Brand Preference Measurement

In the Summer of 2011, the NPD Group, a respected consumer marketing research firm in New York commissioned an independent study of CivicScience data as part of due diligence in consideration of making an investment in the company. NPD retained Joel Rubinson, President of Rubinson Partners, a professor at the NYU Stern School of Business, and formerly the Chief Research Officer of the Advertising Research Foundation, to perform an assessment of the validity and reliability of CS data, as it pertained to consumer brand preference. CS regularly tracks sentiment toward hundreds of consumer brands, the resulting data from which, Rubinson measured against empirical market share data for each brand. The full white paper is available upon request but the key findings are enumerated below:

Objective 1- To determine whether the CS data collection procedure produces good balance in survey responses across brand or whether there is a sizeable and unintended bias.

> **Conclusion 1**- In each of five brand categories, the number of interviews for each brand and the balance of interviews across demographics were extremely well balanced.

Objective 2- To determine whether respondent demographics reflect the demographics of the web properties from which the surveys are taken.

Conclusion 2- Yes, the age and gender profile of respondents was compared to Quantcast estimates of the demographics of website visitors. The skews [identified for the contributing sites] are well reflected in the respondent profile.

Objective 3- To determine whether "top box" brand liking scores correlate with market share. As a corollary for restaurants, NPD Group provided customer experience rating data to measure brand liking data against customer experience data.

Conclusion 3- For each of the three product categories analyzed, the correlations to market share of the percent of respondents giving top box brand liking responses mostly met or exceeded 0.8, which is comparable to the correlations claimed by brand equity measurement systems. In addition the percent saying "I love it" regarding a given restaurant was significantly correlated to NPD respondents rating their dining experience as "excellent."

Objective 4- To determine whether the demographics of those consumers "loving" a brand match the demographics of known purchasers.

Conclusion 4- When both NPD point-of-sale and CivicScience data sets are converted to indices, the array of indices across age and gender groups, brand by brand, are highly correlated.

Objective 5- To determine whether CS brand liking scores trend in a way that suggests they can be used for brand trending and tracking purposes.

Conclusion 5- While there was some evidence that scores declined a bit over time, these declines almost completely stabilized after weighting the data by age and gender and controlling for the source of the interview. (Note: These weighting algorithms have since been embedded in the CS platform).

Overall Conclusion- "The bottom line is that Rubinson Partners, Inc. concludes that CivicScience data can be validly and reliably used for brand marketing measurement purposes."

FINAL CONCLUSIONS

Following rigorous assessment of the CivicScience data collection methodology and comparisons to prevailing landline research techniques, it is the conclusion of the authors that CivicScience has developed a highly reliable and defensible model for measuring public opinion both. Where CivicScience's raw demographic proportions vary from population norms, the company's large sample sizes enable minimal reweighting measures to achieve accurate results. Furthermore, the computational opportunities among data in this volume and structure enable superior levels of analysis, discovery, and predictive modeling.

ACKNOWLEDGEMENTS AND AUTHOR CREDENTIALS

This paper was developed and reviewed by the following authors, each in their capacity as a formal advisor to CivicScience.

Dr. Lamar Pierce
Associate Professor of Strategy
Olin Business School
Washington University in St. Louis

Dr. Aaron Chatterji
Associate Professor of Strategy
Fuqua School of Business
Duke University

Dr. Jason Snyder
Assistant Professor of Strategy
Anderson School of Management
UCLA

Dr. Alessandro Acquisti
Associate Professor of Information Technology and Public Policy
Heinz College
Carnegie Mellon University