

# Are Shoppers Representative of the Population? Using Geofenced Grocery and Convenience Stores to Represent the Population

Davia Moyses<sup>1</sup>, Matt Jans<sup>1</sup>, Ronaldo Iachan<sup>1</sup>  
Lee Harding<sup>1</sup>, James Dayton<sup>2</sup>, Yangyang Deng<sup>1</sup>  
Scott Worthge<sup>3</sup>, Tracy Visconti<sup>3</sup>

<sup>1</sup>ICF, 530 Gaither Rd., Suite 500, Rockville, MD 20850

<sup>2</sup>ICF, 126 College St., Suite 2, Burlington, VT 05401

<sup>3</sup>MFour Mobile Market Research, 19800 MacArthur Blvd., Suite 700, Irvine, CA 92612

## Abstract

General population surveys traditionally sample people from households because people are easy to locate where they live. However, mobile phone technology and geolocation advances have made it simple to locate people in other places. For example, most people visit a grocery or convenience store to purchase food and household items. Can sampling at these locations provide estimates comparable to traditional sampling? This paper describes a piloted method using geofenced grocery and convenience stores. The method samples mobile panel members when they entered geofenced areas around these stores, asking them to answer a few questions, and take a picture of an alcohol, tobacco, or sugar-sweetened beverage display. To evaluate the method, survey responses must be benchmarked against population control totals and a probability survey on the same topic and population. The benefits of this innovative method over traditional phone or mail surveys are quick and inexpensive administration, and the ability to capture images as data.

**Key Words:** nonprobability surveys, mobile phones, innovative data collection methods, nonprobability benchmarking

## 1. Introduction

Nonprobability panels continue to increase in popularity and sophistication, but remain largely untested as replacements for or complements to probability samples. One of the most promising nonprobability panels on the market is MFour's geofenced *Surveys on the Go*<sup>®</sup> panel, which uses the geolocation technology on panel members' smart phones to sample them from specific locations (i.e., areas surrounded by a "geofence").

Geofences are virtual geographic boundaries set around real-world locations, which enable cell phone applications to trigger a response when the device enters or leaves the area. While geofenced surveys are usually used for store intercept market research (i.e., to capture people shopping at a specific store and ask about their shopping experience), this innovative technology can be used to capture a sample of the general population and invite mobile opt-in panelists to complete a survey on any topic when they either enter or leave a given geofence.

This sampling approach has several potential benefits. Logistically, it provides the opportunity to access potential respondents outside of their home and without the use of field data collectors. It is also more cost- and time-efficient than probability samples or on-the-ground intercept surveys. This can benefit general population surveys, surveys targeting rare or hard-to-reach populations, and recreational or environmental surveys. In terms of measurement, there is the possibility to capture real-time data to mitigate recall error and decrease differences that can be found between intent and behaviors in traditional surveys. There is also the option to collect “bonus” data elements, such as the ability to capture images or videos via the cell phone camera during survey administration.

Given these potential benefits, we are investigating whether sampling panelists at geofenced locations is a feasible alternative and/or complement to traditional sampling. Specifically, we are assessing whether a geofenced sample of grocery, convenience, and liquor stores can produce useful population estimates on public health topics, as well as procedurally testing image capture in mobile panel surveys.

## **2. Proof of Concept Design**

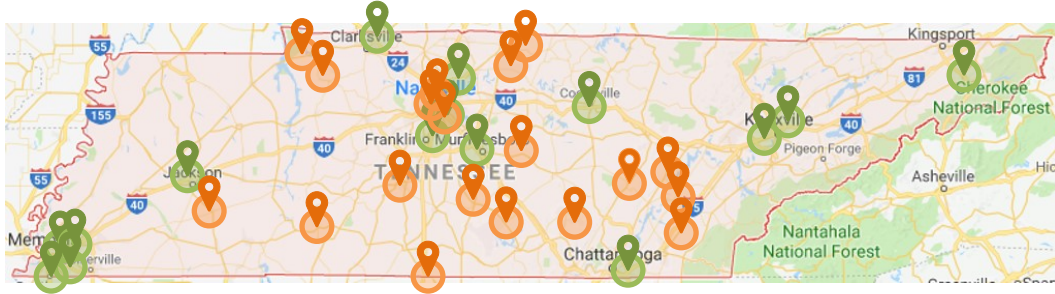
We have developed a rigorous proof of concept design, followed by a concept assessment plan. Though geofenced sampling may be broadly applied, our initial design focuses on geofenced samples as an alternative specifically to household sampling by targeting panelists at a location outside of the home where the general population is likely to routinely visit.

### **2.1 Sampling, Mode, and Fielding**

The population of interest of the proposed design is non-institutionalized adults age 18 and older. The geographic scope of this population may be national or from a narrower geographic scope, such as a single state. A national scope has the benefit of targeting adults with varied backgrounds to assess bias associated with geofenced nonprobability samples in various demographic groups. However, a larger survey completion rate is required to draw reasonable conclusions from the collected data. A state-level scope would require less survey completions to draw conclusions, but potentially offers less diversity in bias analysis, as well as less applicability to broader general population surveys.

Geofences and survey data collection are provided by MFour’s *Surveys on the Go*<sup>®</sup>, which is a mobile opt-in panel with approximately two million active users. Thirty- to forty-foot radial geofences around 12 highly visited grocery, liquor, and convenience store brands are included for sampling purposes. For example, if the grocery store chain Food Mart is classified as “highly visited” on average, geofences around all Food Mart locations within geographic scope are included. Figure 1 provides a visual representation of geofences in Tennessee for two store chains.

All *Surveys on the Go*<sup>®</sup> mobile panelists receive push notifications to complete a brief survey immediately upon entering a selected geofence. Notifications may be visual, audible, or both. The survey remains available to the panelist for 48 hours from the push notification and may be completed after they leave the geofenced area. Visual and/or audible reminders are sent at one, twenty-four, and thirty-six hours. The survey, in general, remains in the field until a set quota of survey completions is obtained.



**Figure 1:** Representation of Two Geofenced Stores in Tennessee  
 Geofences are not presented to scale.

**2.2 Questionnaire Topics**

The brief questionnaire includes two components: basic demographics and health topics. Demographic information is used for eligibility determination and to compare the composition of respondents to the composition of established population surveys. These demographics include state and zip code of residence, place of residence (e.g., private, college housing), age, gender, sexual orientation, ethnicity, race, marital status, education, employment status, and the number of adults in their household by gender.

Health topic data are used to benchmark geofenced survey respondents to known, well-accepted estimates. The health topic questions are presented in Table 1; they are drawn from the Behavioral Risk Factor Surveillance System (BRFSS) and collect data on tobacco, alcohol, and sugar-sweetened beverage behaviors.

**Table 1:** BRFSS Health Topic Questions

<i>Topic</i>	<i>Question</i>	<i>Response Options</i>
Tobacco Behavior	Have you smoked at least 100 cigarettes in your entire life? Do not include electronic cigarettes (e-cigarettes, NJOY, Bluetip), herbal cigarettes, cigars, cigarillos, little cigars, pipes, bidis, kreteks, water pipes (hookahs), or marijuana. Please note that 100 cigarettes is equal to 5 packs of cigarettes.	1. Yes 2. No
Tobacco Behavior	Do you now smoke cigarettes every day, some days, or not at all?	1. Every day 2. Some days 3. Not at all
Tobacco Behavior	Do you currently use chewing tobacco, snuff, or snus every day, some days, or not at all?	1. Every day 2. Some days 3. Not at all
Alcohol Behavior	During the past 30 days, how many days per week or per month did you have at least one drink of any alcoholic beverage such as beer, wine, a malt beverage or liquor?	_____ days per: 1. Week 2. Month  Don't know/Not sure

Alcohol Behavior	During the past 30 days, on the days when you drank, about how many drinks did you drink on the average? Please note: One drink is equivalent to a 12-ounce beer, a 5-ounce glass of wine, or a drink with one shot of liquor. A 40-ounce beer would count as 3 drinks, or a cocktail drink with 2 shots would count as 2 drinks.	_____ Number of drinks Don't know/Not sure
Alcohol Behavior	Considering all types of alcoholic beverages, how many times during the past 30 days did you have [IF MALE, INSERT "5 or more", ELSE IF FEMALE, INSERT "4 or more"] drinks on an occasion?	_____ Number of times None Don't know/Not sure
Alcohol Behavior	During the past 30 days, what is the largest number of drinks you had on any occasion?	_____ Number of drinks None Don't know/Not sure
Sugar-sweetened Beverage Behavior	Not including fruit-flavored drinks or fruit juices with added sugar, how often in the past 30 days did you drink 100% fruit juice such as apple or orange juice? Enter '0' if you did not drink 100% fruit juice in the last 30 days.	_____ times per: 1. Day 2. Week 3. Month  Don't know/Not sure
Sugar-sweetened Beverage Behavior	Now, thinking about sugar-sweetened beverages including regular soda, sports drinks, energy drinks, coffee, tea, and juices that have added sugar, how often in the past 30 days did you drink sugar-sweetened beverages? Enter '0' if you did not drink any sugar-sweetened beverages in the last 30 days.	_____ times per: 1. Day 2. Week 3. Month  Don't know/Not sure

### 2.3 Additional Data Elements

Data collection via mobile phone has the unique ability to ask respondents to capture images or videos via the cell phone camera during survey administration. A single image or video has the potential to provide more detail than a set of survey questions. For example, in lieu of multiple sets of questions to obtain food diary type data, survey respondents can simply take a picture of their meal. To test the usability of the image capture data, respondents are asked to take a picture of any alcohol, tobacco, or sugar-sweetened beverage products either at the store, if they are still inside the geofence, or that they have access to, if they are no longer inside the geofence.

### 3. Concept Assessment Plans

Our concept assessment plan focuses on vetting whether geofenced nonprobability samples produce useful population estimates on public health topics, as well as the utility of image

capture in mobile panel surveys. The plan includes unweighted data benchmarking, weighted data benchmarking, and review of the captured images.

### **3.1 Unweighted Benchmarks**

To evaluate coverage and sample bias, the geofenced sample and respondent composition is compared to the composition of known probability sample surveys, including state-level BRFSS surveys and the National Health Interview Survey (NHIS). Samples and respondents are compared on age, gender, race/ethnicity, marital status, education, and employment status. Location of residence and location of survey invitation from the geofenced respondents are compared to location of residence for the probability respondents. If reasonably equitable demographic and location distributions between nonprobability and probability samples are found, then it is likely that the geofenced sampling method is accurately capturing the same population as the general health survey.

### **3.2 Weighted Benchmarks**

To evaluate geofenced panel accuracy and feasibility, the geofenced sample is weighted and key health behavior estimates are compared to known probability based estimates from BRFSS and NHIS. Geofenced respondents start with a base weight of 1, which is then iteratively raked to population control totals along demographic dimensions. Population control totals are derived from the American Community Survey (ACS). Weighted panel estimates of tobacco, alcohol, and sugar-sweetened beverage behaviors are compared to the probability estimate's confidence interval. If the nonprobability estimate falls within the probability estimate's confidence interval, then it is suggested that the geofenced sampling method is accurately capturing the health behavior constructs.

### **3.4 Image Capture Review**

To evaluate utility of the image capture feature of mobile survey administration, pictures are reviewed against rigorous criteria. Specifically, images are assessed for image quality (such as blurriness, lighting, and alignment/framing) and accuracy to the question prompt (i.e., is the image of tobacco, alcohol, or sugar-sweetened beverages). Willingness to take a picture during survey administration is also assessed. Higher willingness, image quality, and topical accuracy would all suggest that image capture could be a feasible data collection approach.

## **4. Conclusions**

There are several logistical and measurement benefits to using a nonprobability sample derived from geofences around grocery, liquor, and convenience stores. They are time- and cost-efficient and may reduce measurement error. As a complement to or alternative for general population surveys, specifically, the geolocation technology on panel members' smart phones can aid in sampling potential respondents outside of their home, which has historically been the easiest place to locate individuals. As a new and innovative approach to nonprobability sampling, geofenced sampling require rigorous testing. Our proof of concept design was developed with the need for this rigorous testing in mind. The methods proposed allow us to assess various biases, as well as procedurally test image capture in mobile panel surveys. Forthcoming results will inform the potential use of geofenced panels for various, unique survey objectives and populations, including those beyond general population health surveys.