

<https://blog.m-s-g.com/2019/11/14/the-abcs-of-abs-why-the-address-based-sampling-frame-works-so-well/>

The ABCs of ABS: Why the Address Based Sampling frame works so well

ON [NOVEMBER 14, 2019](#) BY [VMERCHANTS](#) IN [ADDRESS BASED SAMPLE](#), [GENESYS](#), [MSG](#)

The past 10 to 15 years have been very good to Address Based Sampling (ABS). ABS has grown so much that it is now perceived as a substitute to random-digit-dial dual frame sample designs, and arguably, it has become the dominant sample survey design in the USA.

ABS is a special type of sampling frame, distinguishable from telephone surveys in its flexibility. The frame can support many methods and modalities: web, phone, and mail. In this article we will briefly explore the popularity of ABS and the problems it attempts to solve.

First, a quick historical lesson. Let's look at what has happened to traditional telephone surveys. Response rates have tanked and many households have scrapped their land lines, forcing survey designers to sample both land lines and cell phone frames. To be fair, as recently as the late 2000's telephone surveys were still doing rather well. They were still cost efficient and dual-frame survey designs (landline and cell phone) were gaining traction. While it is true that response rates were already in decline, data quality was not suffering.

The picture has changed in ten years. Telephone response rates have continued their precipitous decline (now down into the single digits) and associated risks of systematic bias have risen. Researchers have been forced to adapt by choosing alternative methods without sacrificing sampling integrity. ABS is a countermeasure for the trends we have witnessed with telephone response rates. The costs of telephone surveys have risen as well, compared to ABS. Not only has ABS solved some of those problems, it has opened the door to mixed modes of contact and data collection.

ABS from the ground up

The foundation of [Address Based Sampling](#) (ABS) is the United States Postal Service USPS Delivery Sequence File. Marketing Systems Group was one of the first companies to get approval for providing the Computerized Delivery Sequence (CDS) File, which contains just about every deliverable postal address. That's more than 135 million residential addresses to date.

ABS merges the CDS with other data sources that contain geographic and demographic data. This is like cranking up the volume on your guitar amplifier to "11". Data sources consist of both publicly available sources such as the Government's American Community Survey, the Current Population Survey and decennial Census data. Beyond that, ABS can mine commercial databases for additional data. You can append demographics such as age, gender, income, education, and more. By meshing data sources together, the odds for positive matches are increased and the negative impact of coverage lapses are decreased. By targeting the household instead of the telephone number, ABS avoids the under and over coverage downside risks of telephone samples.

The difference maker: Geocoding.

Geocoding is the key ingredient which effectively launched ABS as a valid alternative. Geocoding is the application of geographical coordinates to a corresponding postal address location. Why does this matter so much? It means researchers can reach the majority of U.S. households more inexpensively and faster than ever before.

The basic geocoding method works like this: addresses are coded using linear interpolation, constructing geographical data points within each street segment based on the numeric addresses as end points. The interpolation is accurate to the street level but not necessarily to the actual rooftop level due to factors such as property size and park spaces. Still, you can get very accurate correspondences with geocoding.

There is no better approach for standard mail surveys as ABS has also solved problems with respect to in-person household surveys. Because CDS does not include census geography, it was a problem to design samples for in-person households. In the old days this was solved through costly methods: multi-stage sampling of primary and secondary sampling units based on census blocks and field-testing every address in a segment. ABS removes those obstacles. Every address is geocoded to a census block, with some exceptions such as P.O. boxes, rural routes, and simplified addresses (rural routes, P.O. boxes with no physical address). While it is true that simplified addresses are a nagging problem –the good news is that the scope of the problem has diminished: the number of simplified addresses, once upwards of 10 million addresses, has been reduced to the hundreds of thousands. Not insignificant, but a vast improvement.

In the past, ABS was hampered by some systematic nonresponse factors. For instance, ABS respondents were more likely to be college grads and less likely to be non-White, as compared to RDD samples. Lately however, mitigation efforts have

made real progress due to Census data appends that can be used to predict areas of high nonresponse. You can oversample areas that tend to respond less frequently to ABS surveys. Consumer data also can be appended based on trackable behaviors and predictive models. This too can be used for oversampling nonresponsive areas. In short, there are fewer reservations attached to the use of ABS, hence its increasing popularity.

ABS isn't just a “one-trick pony”

To appreciate the raw power of ABS, you need to think of it as much more than its source USPS CDS file. It is an enhancement of it: CDS plus demographics plus geocoding. The effect is empowering. Researchers can increase the range of analysis options for testing hypotheses and models. And the ease of use has fueled the use of multimode surveys to combat the telephone survey problems mentioned above. ABS is also useful for probability-based panel recruitment, non-response follow ups, and for reaching more inaccessible populations with stratified samples. ABS gives you that flexibility. Samples can be drawn to custom specifications without sacrificing representation.

Key Advantages of ABS

- Single frame. Does away with dual-frame uncertainty.
- Expansive coverage.
- Straightforward weighting protocols.
- Higher response rates, especially when multimodes are used.
- More precise.

For all the reasons mentioned above, ABS is proving to be the best balance between coverage and cost for many researchers, but we can only outline the many factors involved in a short blog article.