HOW USING

DATA QUALITY METRICS

CAN SIGNIFICANTLY IMPROVE THE PRECISION OF ADDRESSABLE MARKETING

An independent evaluation of the validity and applicability of Truthscores™, a new service to Analyze, Improve, and Measure data-driven marketing

A Review of Truthset's Truthscores™
by Rubinson Partners Inc.
Published: February 2021

truth{set} RUBINSON PARTNERS, INC.

Table of Contents

EXECUTIVE SUMMARY

The rise of addressable marketing:

Helping marketers find better, more specific targets for their advertising dollars is the major promise - and challenge - of digital campaigns. The use of "programmatic" technology to buy digital advertising has nearly tripled over the past four years. Programmatic currently accounts for about 80% of digital display because customized messaging benefits both sellers and consumers – and saves money.[1] The challenge remains to make the best use of accurate segments of targetable consumers typically identified by third party data providers. That's the goal.

The challenge of quality control:

Currently, addressable marketing is only part way to that goal. It still promises more than it delivers. The market needs a sniper rifle to hit more select targets - not a shotgun. For example, empirical studies show that the accuracy of consumer segments – particularly when the segments are targeting a desirable demographic (e.g., high-income males) - is "similar to or even worse than what you'd get if you used random chance."[2] Moreover, segment accuracy varies widely across data providers. As it stands now, marketers still need a reliable, independent assessment of the quality of targeting data. Hence, marketers find themselves in a caveat emptor situation.

1. EMarketer, "US Programmatic Digital Display Ad Spending," November 21, 2019
2. Catherine Tucker and Nico Neumann, Harvard Business Review, "Buying Consumer Data? Tread Carefully," May 1, 2020

The Truthset value proposition:

Truthset describes itself *as an unbiased, unconflicted third-party that does not engage in pay-to-play evaluation of data partners. Truthset has built a system that is agnostic to ID spaces, attributes, and marketing channels, one that is able to support many cases beyond marketing and advertising. Importantly, Truthset values complete transparency about its own practices and methodology*

Truthset is a data intelligence company that offers an independent new standard for measuring the accuracy of record-level consumer data. Truthset has a network of data partners that provides large-scale, targetable segments keyed off of HEMs (hashed emails). Truthset also works with validation set partners that provide smaller data assets of self-declared demographics, which are also keyed off of HEMs.

Raising on-target rates of ad delivery:

By shedding light on the quality of the data used for targeting, Truthset allows marketers to increase their campaign on-target rate. They do this by calculating and applying Truthscores™ to targeting data at a record-level. A Truthscore is a number between 0.0 and 1.0 that estimates the probability that a given HEM actually possesses a purported attribute (i.e., demographics, potentially purchase intent, TV watching preferences, etc.). For example, if HEM123 has a "male" Truthscore of 0.2, this means that there is a 20% chance that this HEM is male. By filtering those records in a segment that meet or exceed a given Truthscore threshold, a marketer can improve on-target ad delivery.

Rooted in valid mathematical principles:

To quantify Truthscores, Truthset uses a "Wisdom of Crowds" method.[3] This approach looks for consensus among assertions for the same HEM across different providers. The approach also assesses the accuracy of each individual provider, as measured against the validation sets. For example, five data providers may be equally accurate when measured against the validation set. Assume these five providers all agree that HEM123 is Asian. However, for HEM456, three of them agree and two disagree. In that case, Truthset would assign a higher "Asian" Truthscore to HEM123 than HEM456.

3. The Wisdom of Crowds, James Surowiecki, Anchor Press, 2004

Conclusive validation achieved:

So how well does Truthset's methodology work? Until now, no one has used such data assets and statistical approaches to tackle this use case. To answer this question, a series of statistical hypothesis tests were designed. The result of every hypothesis test – for every combination of attribute and Truthscore threshold that was examined – was significant at the 99% level or higher. The evidence is conclusive that this novel method works and will have an impact on a marketer's ability to more accurately target their desired addressable audience.

How the validation was conducted:

Each statistical hypothesis test began by identifying HEMs from a 20% holdout sample of Truthset's validation sets. For these HEMs, of which there are tens of thousands in total, we compared their self-declared demographic information to their Truthscores, for various attribute values. For example, we examined these HEMs' Truthscores for the 18-24 attribute value – that is, the estimated probability that any given HEM is 18-24 years old.

The overall incidence of the 18-24 demographic in the US adult internet-using population is 14%.[4] We delineated the HEMs from the validation set holdout into two groups: those with Truthscores below this incidence in the US population and those above. For the Truthset methodology to work, there had to be a statistically significant difference in the validated incidence rate of 18-24 year-olds between these two groups of HEMs.

What the validation showed:

This is exactly what we saw. The validated incidence level (of truly 18-24 year-old individuals) was 73.2% among those HEMs above the Truthscore threshold but only 2.6% among those HEMs below the threshold – a remarkable gap. Statistically, this difference is significant at the 99%+ level. In addition, in this example 37% of IDs included in the validation exercise fell above the specified Truthscore threshold. A marketer that worked with Truthset would be able to hone in on only that subset of IDs for advertising and significantly improve its on-target rate.

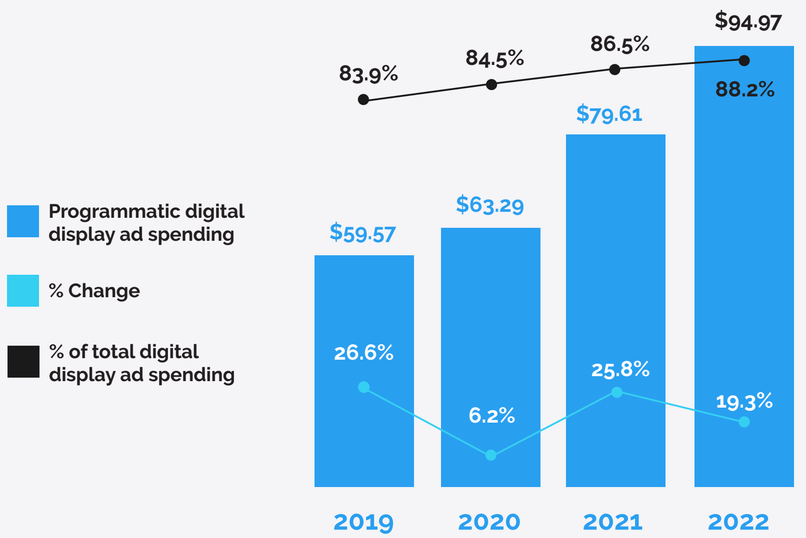
4. Core Trends Survey, Pew Research Center, October 2019.

Rise of programmatic and demand for targetable segments

According to eMarketer, the vast majority (84.5%) of US digital display ad dollars are transacted using programmatic technology today.[5] (See Figure 1.) A main driver of programmatic advertising practices is that a marketer will bid higher for ad serving opportunities for those consumers who are potentially more valuable to their brands, often because that consumer is within the target for the brand which might be defined based on demographics, customer activity records, purported shopper intentions, etc. Many marketers either do not have their own first party data on consumers or wish to expand beyond their footprint, for example to reach non-customers with their ad messages. To accomplish this, marketers will construct an ad tech stack that can act on third party segments which are made available to a marketer on a licensing basis.

Figure 1 : US Programatic Digital Display Ad Spending, 2019-2022

(billions, % change and % of digital display ad spending



Note: digital display ads transacted and fulfilled via automation, including everything from publisher-erected APIs to more standardized RTB technology; including native ads and ads on social networks like Facebook and Twitter; includes advertising that appears on dekstop/laptop computers; mobile phones, tablets and other internet connected devices
(source: eMarketer, July 2020)

5. eMarketer, "US Programmatic Digital Display Ad Spending," November 21, 2019

Third party segments

Marketers can access a third-party segment from a data aggregator either directly or via a data marketplace such as the one that LiveRamp offers. The scale (how many IDs) and the cost (e.g., $1.50 CPM surcharge) are disclosed so a marketer can decide which provider to work with of those that offer similar segments. For example, there are numerous provider sources if a marketer wanted to target consumers with a certain household income range. Currently, scale and price are the two main levers for a marketer when considering to purchase and use third-party data.

"The world is now awash with data and we can see consumers in a lot clearer ways."

~ Mark Levchin ~

Can We?

Missing: a measurement of accuracy

What is missing from the process just described is transparency to the marketer on the accuracy of the segments in the consideration set. For example, there might be five providers of segments of IDs with household income above $75,000 (affluent households), all with roughly similar scale of IDs in their asset. From a data provider's viewpoint, this means that price becomes the main lever, and therefore loses value with every negotiation had between buyer and seller. It is known that not all IDs in a segment are correctly classified and it is suspected that the overall rate of IDs correctly classified varies across segment providers.

To illustrate, consider two providers who each offer 50MM IDs of affluent households. If one provider has 80% accuracy and another has 40% accuracy, the true reach of advertising into the desired demographic segment is different but the total cost will be the same because ads would be served regardless. To a marketer/data buyer, this reflects wasted spend, poor results against KPIs, and difficulty maintaining budget to grow important segments of their audience. Furthermore, if all providers are less than 100% accurate, then it follows that some IDs are accurate within the segment and some are not. Currently, a marketer has no way to unpack a segment to determine which specific IDs are most likely to possess the desired attribute for ad serving. Data providers have no way to differentiate themselves beyond price and scale – only one of which does not have an infinite range.

Unmet need that Truthset addresses

Truthset's methodology is built to address a heretofore unsolved problem. Without Truthset and Truthscores, a marketer has no way of independently evaluating the accuracy of individual demographic assertions from different segment providers, either collectively for the segment or individually for each targetable ID within the segment.

The hypotheses that underlie Truthset's value proposition are:

- The collective accuracy of a segment is often "similar to or even worse than .... random chance" and will vary across providers for the same segment of attribute assertions (e.g., segments of affluent households).[6]

- The likelihood of accuracy for individual IDs within a segment from a particular provider will vary in terms of their underlying probability of being accurate.

- The probability that individual assertions are correct can be accurately estimated by using a "Wisdom of Crowds" method, looking for consistency in assertions for the same ID across different providers.

If Truthset can identify which providers have a higher accuracy for a given targetable demographic segment and if they can sort out which specific IDs are most likely to be the ones that are accurate, they can provide enormous value to a marketer by improving on-target percentages for any addressable marketing use case.

---

6. In addition to Truthset's internal research, there are numerous pieces of academic research, as well as industry articles, that support the idea that consumer data segments available for purchase can be wildly inaccurate, and that this inaccuracy varies on a provider-by-provider basis.

For example, see Catherine Tucker and Nico Neumann, Harvard Business Review, "Buying Consumer Data? Tread Carefully," May 1, 2020.

See also, Digiday, "Why is third party data so often wrong," January 5, 2017.

EVALUATION OF TRUTHSET METHODOLOGY

What is my overall validation approach and brief summary of findings?

The Truthset methodology is predicated on the belief that third-party assertions about characteristics of individual IDs are not always correct. This hypothesis of less than 100% accuracy is logical and supported by the literature as these third-party data are not necessarily self-declared but often the result of inferential signals.

From there, two inferences follow:

Different providers will have different levels of accuracy, that is, the percent of IDs in a segment that truly possesses the purported attribute will vary from provider to provider.

Within a given provider, the accuracy of each HEM-attribute assertion is not equal to the provider's mean but follows some underlying distribution of accuracy probabilities that range from 0 to 100%. For example, if a provider has 75% of HEMs correctly asserted, clearly the individual HEMs do not all have a 75% probability of being correct; some are close to 100%, some are closer to 50%, and some are close to 0%.

Given this logic, Truthset has created a method for distributing HEM-attribute assertions across this range of probabilities that they judge to be reasonable and where evidence demonstrates the statistical accuracy of their method. Let us examine the linchpins of their method.

I have evaluated the Truthset system on the following characteristics and provide a short synopsis of findings below:

- ### Is the Truthset hypothesis of variability in accuracy about individual assertions viable?

  Yes. It is statistically improbable to imagine that there is no variability in the accuracy of individual HEMs when the average accuracy of a group or segment of HEMs is less than 100%.

- ### Is Truthset's method for determining overall provider accuracy valid?

  Yes. Their approach is based on using opt-in behavioral and survey panels and surveys of consumers who have provided self-declared data about their own demographics. Then these consumers are matched to provider records via hashed emails (HEMs) to determine the "hit rate" of correctly assessing if an assertion is correct. I would judge that to be a best-in-class method.

- ### Is Truthset's method for determining the Truthscore (the probability that an individual HEM possesses a given attribute (an assertion)) valid conceptually and mathematically?

  Yes. Their methodology has roots in well-accepted Bayesian analytics, Wisdom of Crowds methods from economics, and is akin to the "poll of polls" approach (e.g. NY Times, Realclearpolitics.com, Nate Silver's 538 site) for integrating election polling results of correctly assessing if an assertion is correct. I would judge that to be a best-in-class method.

Does Truthset's methodology meet statistical validity standards?

Yes. I devised a series of statistical hypothesis tests that align to the primary marketer use case. I directed Truthset's Data Scientists to sort and delineate HEMs into two groups based on their Truthscores for a desired attribute value - those that fell above vs. below a specified Truthscore threshold. For each attribute tested, the Truthscore threshold was set as close as possible to the desired attribute value's incidence in the US adult internet-using population.[7] If Truthscores are accurate, the incidence of HEMs that truly possess the desired attribute value among IDs above vs. below the threshold should be significantly different. In addition, the true incidence of the given attribute value among those IDs that fall above the threshold should be much higher than that of the average internet population. In fact, this is what we found for all assertions analyzed.

For example, African Americans are 11.2% of the US adult internet-using population. If we examine HEMs that fell above vs. below the nearest Truthscore threshold (i.e., a Truthscore threshold of 0.1), the true incidence of African Americans among HEMs that fell above vs. below the threshold is 61.5% vs. 5.7%, respectively.

This supports the use case of a marketer asking a provider to create a subset segment that is above a given Truthscore threshold in order to improve the on-target percentage for consumers who truly possess a given, desired attribute value.

7. Core Trends Survey, Pew Research Center, October 2019

Is the Truthset method of using validation datasets sound for determining overall provider accuracy?

Self-declared data for demographic profiling is generally accepted as the most accurate way of determining demographics for a given individual. This is clearly more deterministic than inferential methods such as assuming someone is a male because they visit male-oriented websites, for example. Such declared data exists via survey and behavioral tracking panels where consumers have opted in for their cooperation and self-declared their demographics. Truthset has integrated a number of well- known and respected panels to maximize scale without reducing accuracy.

Hence, matching records from third-party providers on hashed emails (HEMs) and then comparing the assertion about their demographics to self-declared data is the most accurate way to determine the overall average accuracy of a given provider's segment.

Furthermore, the validation sets provide a way to validate Truthscores about a particular HEM/ assertion combination by matching a hold-out sample of Truthscores for validation purposes. This validation exercise is discussed in further detail in a later section.

Is the analytic method for calculating Truthscores conceptually and mathematically valid?

The method for calculating Truthscores based on agreement or disagreement of an assertion about a particular HEM across multiple data providers is referred to as a "Wisdom of Crowds" method by Truthset. We also note that the methodology has common elements with Bayesian approaches.

Such approaches are ways of bringing all data and signals together into a coordinated, most likely estimate of an underlying parameter.

Let's illustrate the logic of Truthscore estimates with a few examples.

Example 1: HEM789 has the same gender assertion, "male" from both Provider A and Provider B. Let's assume Providers A and B have equally high accuracies when making male assertions, as measured against the validation set. This second, corroborating "male" assertion from Provider B strengthens the "male" Truthscore for HEM789, leading to a higher "male" Truthscore for this HEM than the "male" accuracy against the validation set of either provider alone. The inference is that Truthscores are not bounded by the accuracies of individual providers.

Example 2: HEM789 is asserted to be "Hispanic" by Providers A and B but a third Provider C asserts that the HEM789 is not Hispanic. Let's also assume that all providers have equally high accuracies when making "Hispanic" assertions, as measured against the validation set. This conflicting assertion from Provider C reduces confidence around whether this HEM is truly Hispanic, which in turn lowers its "Hispanic" Truthscore below the accuracy levels on this attribute value for Providers A, B, and C.

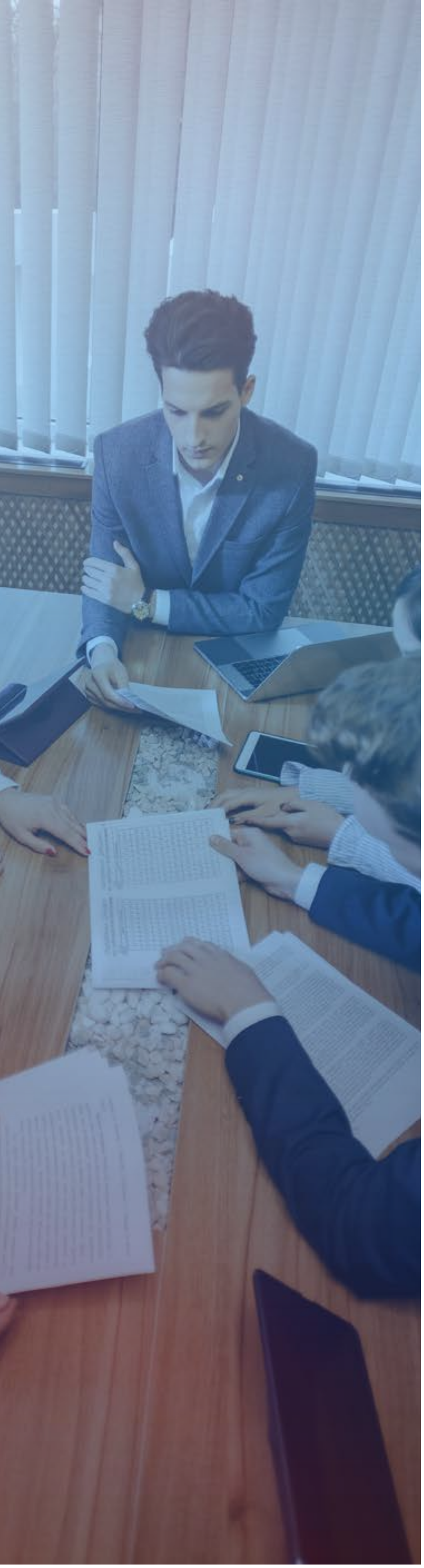
Database of Truthscores:

The result of this process is that Truthset maintains a database of hundreds of millions of HEMs and demographic assertions about each, along with the determined Truthscores.

The set of HEMs (i) that Truthset maintains each have a (potentially distinct) Truthscore for each attribute value (j). We denote these values as Aij . When a given HEM appears in a given provider's data segment with a particular attribute assertion, Aij, the Truthscore for the asserted attribute value is then ascribed to that HEM.

For example, let us say that a particular HEM has a "male" Truthscore of 0.8. If that HEM shows up in two providers' segments of males, that "male" for that HEM will assertion will be assigned 0.8 Truthscore in both provider segments. In general, important properties of Truthscores are:

- The probability that a HEM possesses a given attribute value is a property of the HEM and not of the data provider; hence, once a Truthscore is calculated for a given HEM and attribute assertion, all providers that made the same assertion about the same HEM (e.g., HEM789 is male), will receive the same Truthscore.
- A given HEM can have different Truthscores for each asserted attribute values (e.g., high probability of being a male, but lower probability of being between 21-34 years old).
- Truthscores apply to individual HEMs so a provider segment can be unpacked by HEMs. This means a marketer could select only HEMs with high Truthscores for a desired attribute from a targeting list.
- Truthscores will be updated as Truthset gets information from newly cooperating providers on a particular HEM.



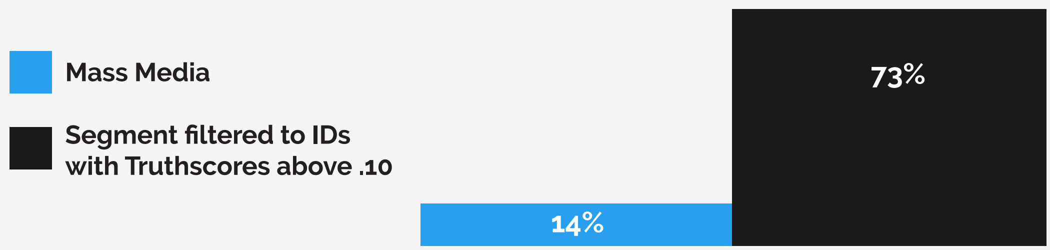
Quantitatively, does Truthset's methodology stand up to independent validation?

Truthscores equip advertisers with the ability to source, buy, and activate upon only the individual HEMs that are above a specified accuracy threshold, thereby improving on-target ad delivery. This concept of higher prediction accuracy of a binary condition as predicted likelihood increases is related to ROC (Receiver Operating Characteristic) curves and confusion matrices; however, Truthset has adapted that statistical approach for determining model validity to be more directly related to the use case of a marketer targeting a subset of HEMs that meet a threshold of likelihood for truly possessing the desired attribute. Think of the use case as a marketer wanting to improve the share of IDs in a particular targetable segment actually possessing the desired demographic, so they want to eliminate those IDs that are less likely to possess the desired attribute.

Consider a real example taken from Truthset's data. Say a marketer wants to target individuals aged 18-24, a group with an incidence in the US adult internet-using population of about 14%.[6] If we look at HEMs with a Truthscore of 0.1 or less for the 18-24 attribute value, which is 63% of all total HEMs included in this validation exercise, the actual incidence of HEMs in that age group among all HEMs with a Truthscore below 0.1 is less than 3%! In other words, ads delivered to 63% of HEMs would be almost completely wasted relative to the intention of targeting 18-24 year-olds! On the other hand, the incidence rate of true 18-24 year-olds among those HEMs with a 18-24 Truthscore greater than 0.1 is close to 73% (vs. a population incidence of 14% which is closer to the on-target delivery rate for mass media ad buys). (See Figure 2). Hence, by weeding out HEMs with a low Truthscore for a desired attribute, a marketer can eliminate a substantial amount of potential ad waste.

7. Core Trends Survey, Pew Research Center, October 2019

Figure 2 : % ads delivered on target to 18-24 year olds



A statistical hypothesis test of the validity of Truthscores is constructed as follows:

- Create a holdout sample of HEMs from the validation sets. These are HEMs found in the validation sets that were intentionally not used to estimate provider accuracies. Therefore, it is possible to compare these HEM's true demographics from the validation set to their Truthscores.

- Divide the holdout sample HEMs into two groups; those below vs above the given Truthscore threshold (determined by the true incidence of a given attribute value in the US adult internet-using population) for each attribute value of interest (e.g., males, 18-24 year olds).

- Within each group, examine the true incidence of HEMs that possess the given attribute value, according to the validation set. If the Truthset methodology works, we will see:
  - The incidence of HEMs truly possessing the given attribute value is statistically significantly different between the two groups of HEMs, and specifically, this incidence rate is higher among the HEMs that fall above the chosen Truthscore threshold.
  - For the group of HEMs that fall above the Truthscore threshold for the given attribute value, the incidence of HEMs truly possessing the given attribute value in that group is statistically significantly higher than the incidence of the given attribute value in the US adult internet-using population.

Results

The results conclusively prove the validity of the Truthscore methodology. For every attribute value that was tested, the difference in the true incidence of the given attribute value between the two groups (i.e., those HEMs that fell above vs. below the given Truthscore threshold) was significant at the 99% level or higher. In addition, the true incidence of the given attribute value among the group of HEMs that fell above the Truthscore threshold was always statistically significantly higher (at the 99% level) than the respective incidence in the US adult internet-using population. Note that these results are robust as they are based on extremely large holdout sample sizes. For most attribute values, the two groups of HEMs (i.e., those above and below the threshold) both contained tens of thousands of HEMs; for other attribute values, the sample sizes in the two groups are as large as over 100,000 HEMs each.

Figure 3 shows that the incidence of 18-24 year-olds in the US adult Internet population is approximately 14%. The true, validated incidence of individuals aged 18-24 among those HEMs above the chosen Truthscore threshold is 73.2% (i.e., approximately five times the incidence of this demographic in nontargeted media).

Figure 3: Comparing Incidence of Desired Age Among HEMs with Truthscore Below vs. Above Threshold

Attribute Value: Age	Incidence in US, Internet-Using Population	Truthscore Threshold Used	Validated Incidence of Attribute Value Below Threshold	Validated Incidence of Attribute Value Above Threshold
18-24	14.0%	0.1	2.6%	73.2%
25-34	20.5%	0.2	4.6%	82.1%
35-44	18.2%	0.2	4.5%	76.2%
45-54	17.1%	0.2	3.8%	70.3%
55-64	15.3%	0.2	4.1%	68.7%
65+	15.1%	0.2	4.6%	63.1%

This result is achieved by weeding out those HEMs that fall below the given Truthscore threshold – among which there is only a true, validated incidence rate of individuals aged 18-24 of 2.6%. The HEMs that fall below the Truthscore threshold are 63% of HEMs used in this validation exercise, so Truthset has improved any provider's segment of this demographic by weeding out 63% of HEMs that would be unproductive to target and would have created (avoidable) ad waste.

These results are really quite remarkable. Think of it this way: a provider offers a segment containing a list of IDs that are all purported to possess a desired attribute, demographic or otherwise. At the start, a marketer has no way of knowing what percent of those IDs truly possess that attribute, and more importantly which exact records accurately represent this attribute, and which do not. Without any direct interrogation (i.e., surveys, which are impossible to do at the scale of tens of millions of IDs), Truthset has created a statistical method for sorting these IDs into sub-segments based on attribute value Truthscores. As demonstrated by my above validation exercise, delineating IDs based on Truthscores results in huge, statistically significant differences in the true, validated incidence rate of the desired attribute.

TRUTHSET METHODOLOGY

An independent evaluation of the validity and applicability of Truthscores, a new service to Analyze, Improve, and Measure data-driven marketing.

The information in this section was provided by Truthset either via their own description or via my interviewing of key personnel.

Overall Description

Truthset is a data intelligence company, focused exclusively on evaluating the accuracy of consumer segment profiling as offered by third party providers of targetable segments. Truthset has created a network of data partners who provide targetable segments along with HEMs, and validation partners who offer self-declared demographic data from individuals, matchable to the data partner records via HEMs.

Truthscores™ - An Introduction

Based on these data assets, Truthset has built a proprietary methodology that assigns Truthscores to each HEM in their database, which are numerical values between 0.0 and 1.0. Truthscores quantify the probability that a record-level assertion is accurate. An assertion is an attribute (e.g., demographics, purchase intent, TV watching preferences, etc.) that is purported by the data provider to be descriptive of a particular HEM. A HEM will have a distinct Truthscore for each attribute (e.g., gender), and each distinct value within an attribute (e.g., male and female). Across every distinct value within an attribute, the Truthscores for the same HEM will sum to 1.

For instance, if a hypothetical Data Provider A asserts that Consumer C is age 45-54, and Truthset assigns that assertion a score of 0.67, it means Truthset believes that Consumer C has a 67% chance of being 45-54. Put differently, this .67 Truthscore also means that Provider A's 45-54 assertion about Consumer C is 67% likely to be accurate.

Truthscores enable data providers to showcase how their data assets compare to those from other data providers, evaluate and prove which attributes they score well on, and conduct further diagnostics on their data sources, data science/modeling practices, and evaluate onboarding partners. For advertisers, Truthscores create the ability to source, buy, and activate upon only the individual HEMs that are above a specified accuracy threshold.

Overview of Analytic Process for Calculating Truthscores

As a general overview, the Truthscore algorithm can be summarized in the following steps: First, for every attribute (e.g., gender) and attribute value (e.g., female) of interest, the average accuracy of each provider's data is established against a validation set.

Secondly, a Bayesian version of a Wisdom of Crowds voting algorithm is then applied to every possible combination of HEM and attribute. The provider accuracies computed in the above step are direct inputs to the Wisdom of the Crowd algorithm.

Finally, Truthscores are evaluated against a random 20% hold-out of the original validation set, so that their performance against an independent, highly accurate set of HEMs can be checked.

Truthscores for each HEM become elements in a proprietary database managed by Truthset. These Truthscores are then reapplied to each of the appropriate provider HEM-attribute pairings to report out on the average Truthscore of the provider for each attribute value, and to create further diagnostics showing the scale of each data provider's asset broken into deciles of Truthscores (e.g., 0.0-0.1, 0.1-0.2, etc.).

For more in-depth information about Truthset's proprietary Truthscore methodology, contact Truthset.

Data Partner Network

To develop Truthscores, Truthset works closely with its Truthset Data Partner Network, a select cohort of data owners and vendors. Truthset has obtained consumer record-level data from these large scale data partners who provide third-party segments for targeting and other use cases.

Taken together, these data partners can be called "Data Partner Network" and their data, at the time of publishing this whitepaper, include upwards of 650 million distinct HEMs. Each partner in the Truthset Data Partner Network provides their full file of consumer data – not a sample, as those can be gamed, and samples are not what marketers are currently buying/activating against in reality.

Truthscores are the result of applying an adapted form of the Wisdom of Crowds voting algorithm to the entire universe of data received from each participant in the Truthset Data Partner Network.[8] For any given HEM-attribute combination, there can be multiple assertions from different data providers in the Truthset Data Partner Network. These assertions are not always in perfect agreement-that is, there is sometimes conflicting information for a given HEM-attribute combination in Truthset's universe of data, even at times within a single provider's data asset. Truthset digests and leverages this information in order to compute HEM-attribute Truthscores.

Validation Sets

Truthset does not rely on any one data source to establish truth about individuals and their associated demographic attributes. Instead, Truthset has partnered with several independent providers of small, highly accurate panels and other data sources. The individual data sources are ingested and combined, to form a single data asset - the validation set. Providers of validation set data are entities that specialize in opt-in recruitment of consumers. Their data assets currently consist entirely of self-reported, declared data (i.e., demographic information directly given by panelists during an intake survey), as opposed to demographic data that is the result of modeling or extrapolation.[9]

8. The Wisdom of Crowds, James Surowiecki, Anchor Press, 2004
9. GreenBook, "The Publisher's Secret Weapon, Using Self-Declared Data to Build Audience Profiles," May 22, 2017

These validation set data sources are currently a mix of behavioral panels and survey data, and other possible opt-in sources of declared data. Taken together, Truthset's combined validation set contains over one million HEMs, with almost all records having complete information for the main demographic attributes of interest.

A small fraction of these HEMs are found in two or more data sources from which the validation set is ultimately comprised. For this overlapping set of HEMs, any possible disagreement between validation sets is examined and reconciled. More specifically, if for a given HEM there is any evidence of conflicting information from multiple providers of validation data for any demographic attribute, then that HEM and all of its associated attributes is removed entirely from the validation sets.

The validation data serves three roles in the Truthscore methodology:

- Allows Truthset to measure the average accuracy of each data provider when making assertions about HEMs.
- Enables Truthset to test and calibrate the output of the Truthscore algorithm.
- Empowers Truthset to evaluate the performance of Truthscores and detect incremental improvements in different approaches/methodologies used to produce Truthscores.

Mitigating Skews in Validation Panels

As is well known, most panels contain various biases and do not perfectly represent a given population. Therefore, it is crucial to correct demographic bias in Truthset's validation sets, in order to prevent this underlying skew from distorting Truthscore estimates. Generally speaking, if a certain demographic (e.g., young people, males, etc.) is over-represented or under-represented in a particular validation set, this will distort analyses of how well any given provider performs when evaluated against that validation set. The extent to which Truthscores would be distorted by underlying demographic skews in the validation sets increases as the skew itself intensifies.

As an extreme example, suppose one validation set contains 90% female HEMs. Truthset's process matches HEMs from the validation set to a segment of HEMs from Data Provider A. Data Provider A, in truth, has an 80% accuracy rate of asserting a HEM is a male. Since the match rate between this data partner and the validation set is 9 times higher for those HEMs that are inaccurately classified by Data Provider A – that is, the HEMs that are actually female according to the validation set but that are asserted to be male – the data partner's accuracy at making male assertions will appear to be well below 50%, when it is really 80%. This is why corrective, population re-factoring weights must be applied to validation sets at the record-level, in order for these data assets to more closely represent US demographics.

The Truthset team has developed an approach to derive corrective weights for underlying skew in panel data. The approach is as follows: for all demographic attributes under scoring/measurement -- as well as for compound attributes (e.g., age and household income, age and gender, etc.), Truthset evaluates the extent to which the demographics in our validation sets deviate from those found in the general population of the United States (for individuals over the age of 16). For demographics which are under- or over-represented in the validation sets, Truthset derives and applies a record-level weight. For example, if men aged 18-24 are twice as common in the US general population as in Truthset's validation sets, then these records are weighted by 2.0 when calculating a provider's accuracy for this segment against the validation sets.

The demographics of the US general population are currently derived from the Public Use Microdata Sample (PUMS) from the American Community Survey as well as the Core Trends Survey from Pew Research. PUMS is a particularly useful dataset, since it contains a statistically representative sample of 1% of the US general population on both a state and federal level and is updated annually. For Truthset, data from Pew and PUMS establish demographics targets for both single (e.g., age) and compound attributes (e.g., age+gender) to which Truthset's validation sets are corrected. Truthset periodically refines the corrective weights applied to validation sets based on industry-best practices for enumeration along with all available reputable data sources from which appropriate correction factors can be derived (e.g., data from PUMS and Pew Research). In addition, with every successive refresh of data from the data providers that collectively comprise the validation set, Truthset aims to both re-evaluate the demographic skews and re-derive the corrective weights created.

APPENDIX

Appendix A : Detailed Validation Data for Demographic Assertions at Different Truthscore Threshold Values

These tables provide full details on how Truthscore threshold values differentiate the true incidence of HEMs possessing the purported attribute.

Comparing true incidence levels of various attributes for holdout samples at various Truthscore™ threshold values

Males Incidence = 49.0%			Females Incidence = 51.0%		
Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold
0.80	21.1%	86.4%	0.80	27.5%	87.2%
0.70	19.0%	84.8%	0.70	22.0%	85.5%
0.60	17.7%	83.1%	0.60	19.7%	84.6%
0.50	16.3%	81.4%	0.50	18.6%	83.7%
0.40	15.4%	80.3%	0.40	16.9%	82.3%
0.30	14.5%	78.0%	0.30	15.2%	81.0%
0.20	12.8%	72.5%	0.20	13.6%	78.9%

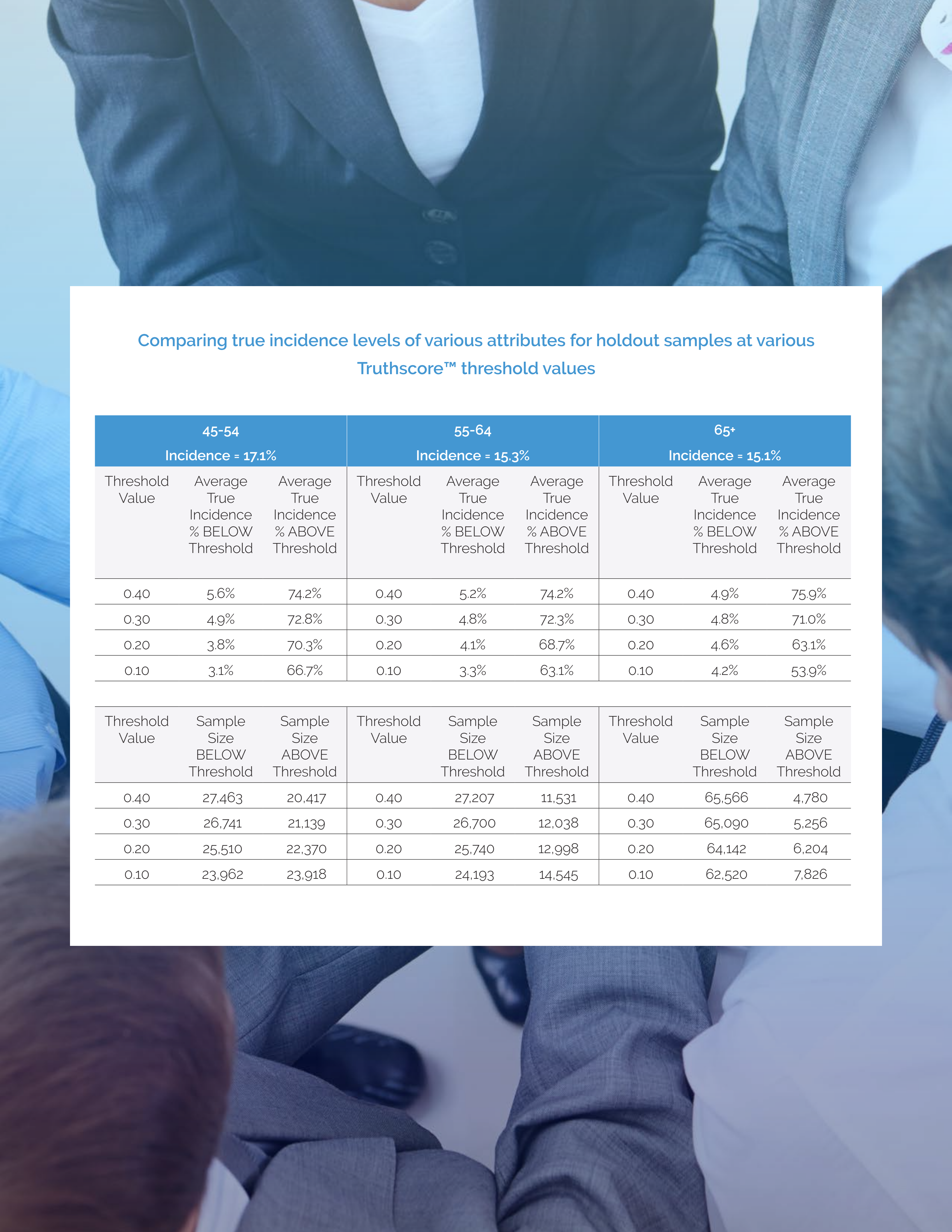
Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold
0.80	216,994	69,832	0.80	116,377	170,449
0.70	208,335	78,491	0.70	101,873	184,953
0.60	202,083	84,743	0.60	95,562	191,264
0.50	195,510	91,316	0.50	91,316	195,510
0.40	191,264	95,562	0.40	84,743	202,083
0.30	184,953	101,873	0.30	78,491	208,335
0.20	170,449	116,377	0.20	69,832	216,994

In general, the tradeoff is that the higher the selected Truthscore threshold, the higher the true incidence of the desired attribute, but the lower the reach in terms of total HEMs above the selected threshold. Regardless, at any threshold, Truthscores demonstrate remarkable difference in true incidence levels of the desired attribute value for HEMs above vs. below the chosen Truthscore threshold.

Comparing true incidence levels of various attributes for holdout samples at various Truthscore™ threshold values

18-24 Incidence = 14.0%			25-34 Incidence = 20.5%			35-44 Incidence = 18.1%		
Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold
0.40	5.2%	76.7%	0.40	7.0%	85.0%	0.40	5.7%	79.4%
0.30	4.0%	76.1%	0.30	6.0%	83.7%	0.30	5.2%	78.3%
0.20	3.1%	74.4%	0.20	4.6%	82.1%	0.20	4.5%	76.2%
0.10	2.6%	73.2%	0.10	3.7%	80.3%	0.10	3.3%	72.3%

Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold
0.40	36,828	17,830	0.40	32,380	34,116	0.40	29,915	20,517
0.30	36,039	18,619	0.30	31,371	35,125	0.30	29,379	21,053
0.20	35,144	19,514	0.20	30,385	36,411	0.20	28,485	21,947
0.10	34,533	20,125	0.10	28,926	37,570	0.10	26,795	23,637



Comparing true incidence levels of various attributes for holdout samples at various Truthscore™ threshold values

45-54 Incidence = 17.1%			55-64 Incidence = 15.3%			65+ Incidence = 15.1%		
Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold
0.40	5.6%	74.2%	0.40	5.2%	74.2%	0.40	4.9%	75.9%
0.30	4.9%	72.8%	0.30	4.8%	72.3%	0.30	4.8%	71.0%
0.20	3.8%	70.3%	0.20	4.1%	68.7%	0.20	4.6%	63.1%
0.10	3.1%	66.7%	0.10	3.3%	63.1%	0.10	4.2%	53.9%

Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold
0.40	27,463	20,417	0.40	27,207	11,531	0.40	65,566	4,780
0.30	26,741	21,139	0.30	26,700	12,038	0.30	65,090	5,256
0.20	25,510	22,370	0.20	25,740	12,998	0.20	64,142	6,204
0.10	23,962	23,918	0.10	24,193	14,545	0.10	62,520	7,826

Comparing true incidence levels of various attributes for holdout samples at various Truthscore™ threshold values

Hispanic Incidence = 14.6%			Asian Incidence = 5.1%			African-American Incidence = 11.2%		
Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold
0.40	4.7%	72.1%	0.40	2.0%	66.1%	0.40	6.9%	66.3%
0.30	4.5%	71.5%	0.30	1.8%	65.7%	0.30	6.3%	64.3%
0.20	4.5%	71.3%	0.20	1.8%	63.9%	0.20	5.8%	62.6%
0.10	4.2%	69.2%	0.10	1.7%	62.0%	0.10	5.7%	61.5%

Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold
0.40	63,796	31,204	0.40	117,502	2,483	0.40	102,566	16,442
0.30	63,398	31,602	0.30	117,113	2,851	0.30	100,963	18,045
0.20	63,309	31,691	0.20	116,955	3,030	0.20	99,587	19,421
0.10	61,971	33,029	0.10	116,760	3,225	0.10	98,912	20,096

Comparing true incidence levels of various attributes for holdout samples at various Truthscore™ threshold values

HHI<$50K Incidence = 47.2%			HHI btwn $50-100k Incidence = 27.7%			HHI>$100K Incidence = 25.1%		
Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold	Threshold Value	Average True Incidence % BELOW Threshold	Average True Incidence % ABOVE Threshold
0.70	31.6%	59.2%	0.50	31.8%	43.7%	0.50	23.9%	56.9%
0.60	30.0%	57.2%	0.40	29.9%	41.5%	0.40	22.4%	53.4%
0.50	27.2%	55.4%	0.30	28.0%	38.5%	0.30	20.6%	49.5%
0.40	24.2%	53.0%	0.20	25.9%	37.3%	0.20	18.7%	44.0%

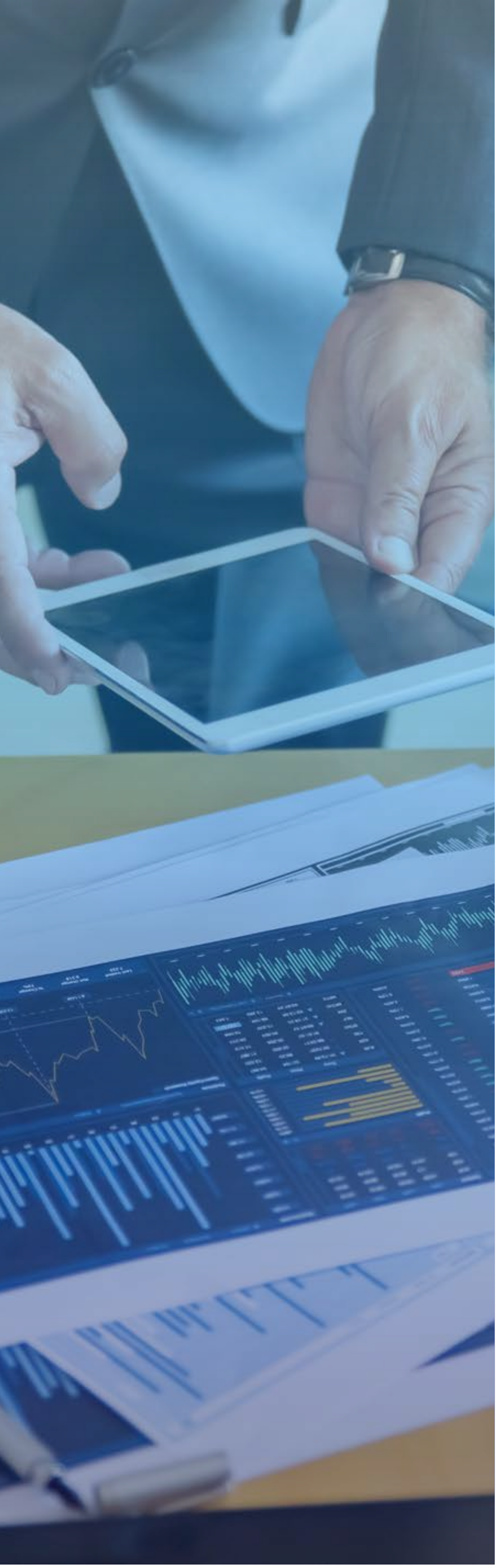
Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold	Threshold Value	Sample Size BELOW Threshold	Sample Size ABOVE Threshold
0.70	107,588	16,932	0.50	78,105	37,063	0.50	99,757	22,923
0.60	100,010	24,510	0.40	58,069	57,099	0.40	92,433	30,248
0.50	88,439	36,081	0.30	31,319	83,849	0.30	82,651	40,030
0.40	76,108	48,412	0.20	16,686	98,482	0.20	67,636	55,045

Appendix B : Rubinson Partners Credentials for Review of Truthset Methods

Joel Rubinson was retained to provide an independent assessment of the Truthset methodology. Joel's credentials for this review include the following: running a successful consulting business that has served 75 leading firms across AdTech and media (e.g., Oracle's Moat, Viant, NBC, AOL (now Verizon), and many marketers (e.g., General Mills, Coca-Cola, Unilever, J&J, MetLife, Verizon, Estee Lauder). Joel was the former Chief Research Officer of the Advertising Research Foundation. Among his consulting assignments, he has functioned since 2016 as the subject matter expert for Multi-Touch Attribution (MTA) approaches for the Mobile Marketing Association, interacting with over 50 leading marketers, media companies, and AdTech firms on advanced analytics topics. Joel's white paper, "The Persuadables", tested in partnership with Viant and Nielsen Catalina Services, is viewed as a definitive study of the value of targeting. Joel was the CRO for the NPD Group, leading efforts on creating and refining their weighting and projection systems for their sales currency data on the industries they track. As a faculty member at NYU, Joel created and taught their first grad course on social media marketing. He started his career as the head of analytics for Unilever in the US, and holds an MBA with concentrations in economics and statistics from the University of Chicago.

For more information about Joel Rubinson's professional credentials, please visit:

https://www.linkedin.com/in/joel-rubinson-a3a0763/



Appendix C : Glossary of Terms

3rd party provider: refers to companies that offer segments available for marketers to license for ad targeting purposes and other use cases.

Addressable advertising: advertising that is served to an individual ID selectively (others viewing the same media need not see the same ads), driven by either programmatic bidding or by algorithms applied by the media company whose content has attracted the consumer.

Assertion: a claim by a 3rd party provider that a particular HEM possesses a particular value of an attribute (e.g., that a particular HEM is a male.)

Attribute: the characteristics – demographic or otherwise -- associated with an individual ID (e.g., age, gender, household income, ethnicity, etc.)

Attribute value: the distinct values or levels within a given attribute (e.g., "male" and "female" for gender, "yes" and "no" for Hispanic, etc.)

HEMs: stands for hashed emails and is a common, anonymous way of matching lists of users across partnering AdTech companies, marketers, and publishers. Truthset's basic unit are consumers identified via HEMs.

IDs: generic term for identifier that is used to match a particular user across different datasets and across AdTech partners and publishers.

On-target ad delivery/impressions served (rate/percentage): refers to the percent of those receiving ad impressions who are truly in the marketers' defined demographic target.

Truthset: the name of the company offering Truthscores.

Truthscores: a numerical value ranging from 0 to 1, that reflects the probability as estimated by Truthset that a particular HEM possesses an asserted attribute value. Truthscores across attribute values within the same attribute (e.g., "male" and "female" for gender) sum to 1.0.

Validation data sets: a data asset that is comprised of a set of consumers who declare their own demographics when they are recruited to engage in survey panels.

Validated incidence: refers to the average incidence of a collection of IDs that is determined by direct match of a hold-out sample to a panel of consumers who have opted in and offered their own self-declared demographics.

The Wisdom of Crowds: a concept that was first enumerated by economist James Surowiecki in his book by the same name, purporting that groups of consumers providing independent estimates, when averaged, lead to highly accurate estimators.

A Review of Truthset's Truthscores™
by Rubinson Partners Inc.
Published: February 2021

truth{set}
RUBINSON PARTNERS, INC.