

## MRP: A Model That Learns Over Time

MRP allows the Latana brand tracking solution to detect differences in brand KPIs for niche target audiences, even when the sample size is limited.

0 0



PUBLISHED BY Latana.com

## Welcome

Hello,

Latana is the world's first advanced brand tracker fuelled by AI technology. We focus on pushing boundaries to bring a new level of quality brand tracking to the world. The precise data our brand tracker provides accurately shows companies the real-word impact brought by brand campaigns and enables them to make better marketing decisions.

## Korbinian Kuusisto Fred DeVeaux

www.latana.com

We also regularly share our thoughts on marketing and brands.

www.latana.com/articles

### © 2021 Latana

We'll spare you the legal mumbo jumbo. But please don't share this book without giving us appropriate credit and a link. Got questions? Head over to latana.com and get in touch through our chatbot or content form.

Join Latana on Social Media



## Content

- P. 4 Executive Summary
- P. 5 Introduction
- P. 6 **A Brief History of MRP**

### P. 7 Case Study: Measuring Brand Awareness

- P. 9 Niche Audience: Respondents Aged 56-65
- P. 10 Niche Audience: The "Unproductive"
- P. 11 Niche Audience: Low Education

### P. 12 Conclusion

# Executive Summary

Multilevel Regression and Poststratification This whitepaper outlines a model for breaking past the limits of traditional survey research to build a more powerful brand tracking tool. This model is called **MRP (Multilevel Regression and Poststratification)**, a statistical technique that can be used to guarantee reliable and precise data.

The solution proposed in this whitepaper demonstrates just one key advantage that MRP provides. It will show that MRP is a predictive model that has the ability to learn from prior information accumulated over time through repeated surveys. To illustrate the benefit of this "learning" component of MRP, this whitepaper includes a case study based around Blinkist, a digital reading app. Using the data gathered by Latana for Blinkist, this whitepaper will show how it is now possible to detect differences in survey responses (i.e. brand awareness) for niche target audiences even when the sample size is limited.



## Introduction

## Imagine spending an hour on your balcony counting passing pedestrians in order to predict how many people pass by during the entire day.

It is fair to say that your guess would probably be pretty inaccurate because one hour of observation does not provide a lot of information. However, if you repeated this exercise every day for a month, your prediction by the end of the month would be a lot better. Why? Because as humans, we learn over time. We start to notice patterns, for example, how traffic increases during rush hour and decreases when it is raining. See, the more information we collect, the more we are able to build a general sense of how different factors contribute to what we are trying to measure.

This thinking process comes naturally to humans. However, in the world of brand tracking, it is almost entirely missing from traditional survey-research methodologies. This means that if we used a traditional brand tracking tool to measure a KPI every day for a month, the methodology would treat each day as a completely new day, making a new prediction from scratch every day. So by the end of the month, the prediction would be no more accurate than the prediction at the beginning of the month - no prior information was taken into account so nothing has been learned over time to improve the prediction.

### Why is this a problem?

Starting from scratch every day means that the prediction is vulnerable to daily fluctuations. Consider the balcony example again: on the thirtieth day the one hour of observation happened during a rainstorm and no one

### What about MRP?

However, if the brand tracking tool used information learned over time, there wouldn't be such variations: before making a new daily prediction, the brand tracking tool would already know that millennials have a certain level

walked by. Would you predict that no one walked down the street for the rest of the day, even during the sunny hours? No, because by the thirtieth day we would have a general sense that rain means fewer pedestrians. As a result, our prediction wouldn't depend entirely on what we happened to see in our limited, rainy hour of observation.

This is also a problem when it comes to brand tracking tools that rely on surveys to gather data. Why? Because it is quite common to have variations in the number or type of respondents that a survey reaches on a daily basis, especially when tracking niche audiences. For example, if a brand tracking tool wants to track brand awareness for a particular demographic group (i.e. millennials), the daily predictions might fluctuate depending on how many millennials the survey reached on a particular day. of brand awareness. So even if on a particular day the survey didn't reach many millennial respondents, the prediction would already have a general sense of what the level of brand awareness level should be. This accumulated information helps improve the prediction and guarantees that results don't depend entirely on a small amount of new information.

This whitepaper will demonstrate how the brand tracking tool Latana uses MRP to make better predictions over time. More specifically, it will show how MRP is able to use prior information to make more accurate predictions for brand-tracking KPIs in niche audience groups where the sample size is limited.

# A Brief History of MRP



**Thomas Bayes**, 1702 - 1761 English mathematician, statistician, philosopher and Presbyterian pastor.

#### **ELECTION POLLING**

**Thomas Bayes** was an English statistician and philosopher who formulated the Bayes' theorem upon which MRP is based. The theorem was first presented in the work "An Essay towards solving a Problem in the Doctrine of Chances", which was read to the Royal Society in 1763 after Bayes' death.

The Bayes' theorem determines the probability of an event. The main advantage is that it gives the full probability distribution i.e. beter uncertainty estimates. Therefore, the theorem allows for a more accurate risk analysis than a simple assumption that is typical of the population as a whole.

MRP was first used in the US to estimate US-state-level voter preference in 2009. However, the process gained popularity after it was used to estimate the outcome of the 2012 US presidential election based on a survey of Xbox users, where it was able to mitigate the problem of highly skewed data and provide good results despite some bad data being collected. MRP was also used to successfully predict the 2016 election victory of Donald Trump and the overall outcome of the 2017 UK general election.

While in the past research was a battle of finding a sample of people representative of the whole population, MRP records a lot of data about the respondents and uses it to create a model of how various groups of people are likely to vote. MRP enables you to make granular predictions by assuming that certain demographics in one segment have similar preferences to the same demographic in another segment, eliminating the need for large samples in every constituency.

#### MRP AND BRAND TRACKING

In 2019, Latana became the first brand tracker to apply MRP (Multilevel Regression and Poststratification) to brand tracking. MRP is a machine learning algorithm that enables dynamic modeling, which means it can control more variables and maintain higher levels of accuracy for niche audiences.

As a result, Latana is the first company to devise a brand tracking solution that guarantees deeper insights and more reliable results than quota sampling. That means brand tracking software which:

- Produces lower margins of error
- Allows <u>deep audience segmentation</u> so brands can track even the most niche demographics
- Provides better quality insights that are more closely aligned to real-world changes

CASE STUDY

# Measuring Brand Awareness













Niche Audience

## Respondents Aged 56-65

PAGE 9

Niche Audience

## The "Unproductive" Respondents

PAGE 10

Niche Audience

### Respondents with Low Education

PAGE 11

#### **CASE STUDY**

## **Measuring Brand** Awareness

## 🗕 Blinkist

non-fiction books into 15-minute audio summaries. Latana monitored Blinkist's levels of brand awareness in Germany before, during and after Blinkist's TV campaign by surveying 2000 people. They then used the MRP model to predict brand awareness levels for hundreds of niche target audiences.

Blinkist is an up-and-coming startup that has built a reading app that condenses

#### THE SOLUTION

For the sake of this case study, the focus is on the measure of brand awareness Blinkist reached in Germany after the campaign finished. To show the benefits of using prior information to make more accurate predictions, the following two approaches will be compared:

1. No prior knowledge: we use an MRP model to predict post-campaign brand awareness by simply relying on the respondents that we collected after the campaign.

2. Prior knowledge: we use an MRP model to predict post-campaign brand awareness that relies on the survey data collected before, during and after the campaign.

THE RESULTS

For both scenarios, the overall population estimate of post-campaign brand

Brand awareness results for the small group of respondents aged 56-65 with vs. without prior information.



awareness is 7.5% - meaning that around 7.5% of people in Germany were aware of an app called Blinkist in the time following Blinkist's TV campaign. In all following graphics, this average is shown as a dotted horizontal line.

Note that this estimate is the same regardless of whether or not we use prior knowledge. This is because the post-campaign sample size of 2000 respondents is large enough to generate a stable estimate for the general population and no additional information is required.

However, if we zoom into niche audiences for which the sample size is much smaller, the difference between the two approaches becomes apparent. On the one hand, when using no prior knowledge, the MRP model is not able to detect any difference in brand awareness between niche audiences and the 7.5% level of brand awareness for the general public. On the other hand, when the MRP model does use prior information, it is able to **detect differences in brand awareness for niche** audiences.

In other words, it is only by using prior information that it's possible to detect brand awareness for niche audience groups, particularly when the sample is so small. Keep reading to see some examples on the following pages.

#### NICHE AUDIENCE

# Respondents Aged 56-65



To estimate brand awareness for the small group of respondents aged 56-65 (approximately 11% of the sample / 220 people), using prior information from past surveys is crucial. In the graph below, it can be seen that if prior information is not used, the brand awareness estimate for this group is essentially the same as the overall brand awareness of 7.5%.

This happens because the MRP model doesn't have enough information from respondents aged 56-65 in the sample to find any differences between them and the rest of the sample. However, if the MRP model is allowed to use information from the past (i.e. the survey data that occurred before and during the campaign), then this helps the model find a stronger signal.

By using prior information, there comes a different result: the MRP model estimates that brand awareness for 56-65-year-olds is 5.5%. Therefore, without using prior information, MRP would not be able to detect a difference between the general population and 56-65-year-olds and would simply assign the niche audience the overall average of 7.5%, even if the full sample of 2000 respondents was used.

**Brand Awareness for Respondents Aged 56-65** 

NO PRIOR INFORMATION PRIOR INFORMATION



Sample Size

MRP: A MODEL THAT LEARNS OVER TIME

#### NICHE AUDIENCE

# The "Unproductive" Respondents



Consider another example: Blinkist is an app marketed to people who want to be productive during their free time. Therefore, it would be expected that these people would have a higher awareness of Blinkist than the people who choose not to be productive in their free time (i.e. the "unproductive"). The unproductive respondents are considered a niche audience because just 65% of the sample described themselves as "not wanting to be productive in their free time".

To predict brand awareness for this niche audience, it depends on whether or not the MRP model is allowed to use prior information. Because of the small sample sizes, **only the MRP model that uses prior information is able to detect that these unproductive people have a smaller level of brand awareness (~5%) than the average population (7.5%)**. The model that doesn't use prior information requires a sample size of at least 400 (~260 "unproductive" people vs ~140 "productive" people) to start detecting this difference. **At around a 1000 sample size, both approaches converge to 5.5%**, and the benefit of using past information becomes negligible. As opposed to the previous example with respondents aged 56-65, this example demonstrates how using prior knowledge is particularly useful for smaller sample sizes. To put it in simple terms, if a survey is run before, during, and after a campaign, and the sample sizes are on the smaller side, then using prior knowledge from past surveys would be the only way to detect lower brand awareness for the unproductive niche audience.

### **Brand Awareness for "Unproductive" Respondents**

NO PRIOR INFORMATION PRIOR INFORMATION



Sample Size

#### NICHE AUDIENCE

# Respondents with Low Education



In this case, this "low education" niche audience is considered as people who don't have a university degree. Again we see a similar pattern as the previous example.

The model that uses prior information helps detect a lower level of brand awareness, even with small sample sizes.

On the flip side, the model that doesn't use prior information only starts to detect the lower sample size at a sample size of 800 respondents or more.

### Brand Awareness for Respondents with Low Education



PRIOR INFORMATION



Sample Size

## Conclusion

When it comes to accurate and reliable brand tracking, it is more valuable to use as much information as possible. It was clearly shown in the three niche audiences of the Blinkist case study, that allowing MRP to use prior information from past surveys, it is possible to detect differences in brand KPIs for niche target audiences, even when the sample size is limited.

The results of the case study are particularly important for survey research that focuses on understanding niche audiences, and even more so where the sample size is limited and reaching respondents of a certain category may be challenging.

This is where using past information in an MRP model provides the biggest additional benefit. This is where using past information in an MRP model provides the biggest additional benefit.

Therefore, this whitepaper has demonstrated how Latana can improve the stability and accuracy of brand tracking predictions by using a key element of MRP: the ability to learn from past information.



MRP: A MODEL THAT LEARNS OVER TIME



## Al-Powered Brand Tracking Make Better Brand and Marketing Decisions

#### 

To start a conversation about the business value of MRP in your organization, contact:

For General Enquiries: <u>hello@latana.com</u>

For Sales: <a href="mailto:sales@latana.com">sales@latana.com</a>

For PR/Partnerships: marketing@latana.com

PUBLISHED BY Latana.com

Join Latana on Social Media

