

**Choice-Based Conjoint for Designing Home Appliances:
Human Versus Silicon Samples of Respondents**

By Daniel Baier, Danilo Randazzo, and Maximilian Unger

Executive Summary

In marketing and market research, the support potential of Generative Artificial Intelligence (GenAI) was recognized very early and nowadays is wide-spread in practical use, e.g., for analyzing customers' feedback to products and services, for answering social media posts, for replacing respondents in surveys, for generating web content, or even for developing new ideas and concepts for products and services. However, some researchers have questioned the validity of this widespread practice, particularly when GenAI is used to replace samples of human respondents in surveys.

In this paper, we contribute to answering this question by exploring whether generated answers by GenAI in a Choice-Based Conjoint (CBC) market research study lead to similar partworth estimates as if estimated from collected answers from humans. Our application area for this comparison is the market-oriented design of a new luxurious refrigerator for German households by a major European home appliance manufacturer. We generate GenAI choices among sets of multi-attributed home appliances by simulating a representative sample of potential buyers (the so-called silicon sample) in the prompts and – alternatively – collect human choices from a sample of real-world respondents with the same distribution regarding age, gender, household size, income, and innovativeness. Based on both samples and their generated/collected choices, hierarchical Bayes multinomial logit partworths are estimated and compared. As an experimental factor, the GenAI answers are generated using prompts with a varying amount of background information, i.e., “without any additional information”, “with market-related information from company experts”, and “with information from a preliminary study”.

The results are encouraging: Even for a CBC study with innovative attribute-levels, partworth estimates from GenAI choices come close to partworth estimates from human choices, at least at the aggregate level and when market-related information of experts and/or from preliminary studies is provided. However, at the individual level and when additional information is not provided, GenAI choices lead to partworth estimates whose validity is almost equivalent to that of partworth estimates from random choices. The limited data that describe the synthetic respondent in the “without any additional information” version of the

GenAI prompts (the distributional information with respect to age, gender, household size, income, and innovativeness) seems to be the reason: This information is not sufficient to predict the individual preferences and purchasing behavior of the simulated respondent. If – additionally – answers from past surveys were available that describe the her or his preferences and past buying behavior in more detail, even in other markets, better results could be possible. One should test whether elaborated digital twins of online access panel members (with answers from past surveys) are superior. However, the General Data Protection Regulation defines clear limits for this endeavor.