ESTIMATING PRODUCT PREFERENCES: A METHODOLOGICAL EXCURSION

Author: Brian Cavoto, PhD

Survey Platform & Hosting Provided by:
decipher illuminate opportunity

YouGov | DefinitiveInsights
RESEARCH CURRENTS

Summary

In response to varying client requirements, several preference share estimation techniques have been developed. The focus here is to compare three of those approaches, including a traditional discrete choice task, a modified discrete choice task called a canned SKU task, and a simple static explicit question.

All three of the tasks were found to produce similar preference share estimates when parameters were matched, suggesting any one of the approaches is a viable estimation technique to use, with the choice of approach best driven by the requirements at hand, such as whether features need to be toggled as parameters, whether price points need to be varied, etc., as well as the survey real estate available. While additional research is needed, there do appear to be order effects operating when multiple discrete choice tasks are included in a survey, which should be taken into account in the design of future surveys.

Table of Contents

Background

One of the most common classes of business questions we are tasked with addressing involves estimating demand or preferences for products and services, often including how various features or capabilities — including price — impact demand.

The discrete choice approach has emerged as the preferred technique for predicting interest in products and services[1]. The task typically involves presenting a series of product or service options, and asking responders to select which one they would choose, if any. This approach is attractive for a variety of reasons:

- Given the requirement to make a choice, the approach forces discrimination (a welcome benefit for anyone who knows survey research) — implicitly providing a preference share type of currency for analysis.
- The approach has been proven to have tremendous validity in terms of the mapping between what participants are asked to do in the survey task and the choices they make in the real world.

The discrete choice approach involves the creation of statistical models, which can be used to move beyond static reports to create user-friendly simulators that afford "what-if" scenarios involving share and revenue. Clients find these simulators extremely actionable, and they have a longer shelf life than traditional static analyses and reports, particularly given the ability to dynamically toggle all relevant parameters.

As often happens, success leads to infatuation, leading to the over-application of the technique including pressure to "get one of those simulators". While we are big fans of leveraging discrete choice tasks in alternative contexts (beyond the typical product purchase scenarios), there are times when the specific business questions at hand do not lend themselves very well to a traditional discrete choice approach.

[1] Our focus here is discrete choice, or choice-based conjoint in some circles, to be distinguished from standard conjoint, where the task for participants involves rating or ranking their interest in products instead of making choices.

In traditional discrete choice a number of parameters are simultaneously and independently varied across a set number of products or SKUs (typically 3 to 5 choices at a time, Figure 1) with the specific makeup of each product option or SKU systematically varying across screens.

One common scenario that does not fit as well with traditional discrete choice is when a limited set of products or SKUs are available, with the research questions primarily around how to price the SKUs. This most often occurs in the later stages of the process, or because we are focused on getting a clean read on the uptake of SKUs (often set in a competitive context) without burdening the participant cognitively by asking them to digest and discriminate among numerous features.

Two frequent requests include:

- Systematically including each of those relatively static SKUs on each screen (e.g. always showing two of the clients' SKUs, and each of the key competitors' SKUs, on every screen), given the greater face validity that affords by virtue of better mimicking the real-world situation.
- Testing distinct price points per SKU that reflect more realistic price ranges per SKU, as they do or might exist in the marketplace.

	Option A	**Option B**	**Option C**	**None**
Processor	Fastest	Standard	Slower	
Memory	2 GB	3 GB	4 GB	
Hard Drive	500 GB	300 GB	150 GB	
Screen Size	10"	14"	17"	
Video Card	Discrete	Integrated	None	
Price	$100	$150	$125	
	☐	☐	☐	☒

Figure 1 Example Discrete Choice Task

We have developed a couple of modified versions of the traditional discrete choice task using a relatively static set of SKUs on each screen, with either limited or no variation in features, and with variation in price. Figure 2 depicts the simplest scenario, when several static products are offered at varying price points. The features in the solid red boxes remain static across screens, while price is varied (red dashed boxes). **Distinct price points** per SKU can also be accommodated with this approach.

The advantages of this approach are twofold:

- There is greater face validity for individual choices in the sense that the full set of options are shown on every screen, as would happen in the real world.
- The biggest advantage, is the reduced cognitive load put on the participant: once survey respondents digest the details of each of the SKUs, they just focus on the varying price points. This is a much easier task than having to determine the makeup of several new SKUs, with a varying price, as is the case with a traditional task. Given the level of engagement, and typical position of survey respondents on the speed/accuracy tradeoff, there are obvious advantages to reducing the cognitive load.

This modified approach, showing static features and varying price, is called a “canned SKU discrete choice”. While the traditional and canned flavors of discrete choice tasks have varying sweet spots, the question of which is “best” or most appropriate for a particular business is not always so black and white.

There have been opportunities to compare the results of discrete choice tasks and canned SKU tasks when multiple discrete choice tasks are included within a single study (e.g. a shorter canned SKU task set in competitive context, and a traditional task focused on detailed capabilities within a client’s SKUs). However, the tasks are usually designed and framed differently — on purpose, specifically to address different questions — limiting our ability to make comparisons. For example, the overall take rate for a similar SKU might be compared across discrete choice tasks, but competitors were included in one task but excluded from the other.

	Option A	Option B	Option C	None
Processor	Fastest	Standard	Slower	
Memory	2 GB	3 GB	4 GB	
Hard Drive	500 GB	300 GB	150 GB	
Screen Size	10”	14”	17”	
Video Card	Discrete	Integrated	None	
Price	$100	$150	$125	
	☐	☐	☐	X

Figure 2 Example Discrete Choice Task

Research Objectives

Traditional Discrete Choice	Canned SKU Discrete Choice	Single Explicit Question
• Independently varying features across cards • Single price band	• Static SKUs, with variation in just price across cards • Distinct price bands per individual SKU	• Static SKUs, with each set to a single price • Similar to single card from other tasks

Figure 3

The research described here compares traditional discrete choice and a canned discrete choice task, with the framing of the questions and other considerations purposely matched between tasks to afford as direct a comparison as possible. The tasks were matched in terms of the number of SKUs shown/screens, and the number of dimensions appearing within each SKU.

Also included was a third "task", with a single explicit question that includes SKUs set at static prices — in essence a single screen pulled from on of the other exercises. This third option is sometimes used when time/survey space is very limited and a quick high-level read on the uptake of various SKUs at anticipated set prices is needed.

Objectives

Key Objectives

1. How do results compare between a traditional and canned discrete choice that are otherwise matched content-wise?
2. Are there order effects resulting from the inclusion of multiple discrete choice tasks in a single survey?

The first and primary objective was to look at how the various tasks compare in terms of the resulting market share estimates for the otherwise comparable products/SKUs. There are a couple of ways to operationalize this comparison, including looking at the consistency between the tasks, consistency with other available metrics in the survey data, as well as comparing the results to external estimates. In this case, given the somewhat forward-looking framing and scope of the tasks, where the services tested are not yet offered in the market, the third validity check against external estimates is a difficult one to make.

The second objective related to our now fairly common practice of including multiple discrete choice tasks in a single study and, specifically, whether there are *order effects* operating across tasks when this is done. While there is some existing literature related to this question, most involves a comparison of estimates across individual screens within a single task (Johnson & Orme, 1996; Orme, 2010), as opposed to the effect of one task on another. One study did include multiple tasks, but that work focused on the impact of participating in a conjoint task, as gauged by short discrete choice tasks conducted before and after the conjoint task (Huber, Wittink, & Johnson, 1992).[2] In this research the order of the tasks were varied experimentally to determine how order should be considered in future studies. We focused on the impact of order on the pattern of preference share results, as well as the consistency of participants' responses within tasks - as a function of task position.

---

[2] While multiple discrete choice tasks have been leveraged in many client studies, there is often some logic behind the ordering of the tasks that has previously prevented us from doing this experimental manipulation. For example, we sometimes want participants to step through a detailed task prior to a higher-level task, so they have an appreciation of the nuances going into that higher-level task. This is similar conceptually to our standard use of sorting or other tasks as an educational prelude to critical tasks, to in effect force participants to digest the information, since empirical tests confirm that participants rarely read dense text that might otherwise be included up front before the task, or as mouse-over information available on an as needed basis.

Methodology

The research was part of a study conducted with small and medium-sized businesses (SMBs) in the U.S. during July, 2010. Commercial SMBs were targeted, defined here as companies that have 5-249 PCs, excluding home-based businesses. Other company-level criteria included:

- Commercial businesses (excluding Government and Education).
- Using email and productivity applications of some type.
- Using or at least be willing to consider hosted services.
- Purchased one of the following in past 12 months: servers, server operating systems, server management software, email software, web conferencing, collaboration tools, productivity applications, telephony systems, business applications, or file storage/backup.

Participants included IT professionals (ITDMs) and business professionals (BDMs) who are purchase decision-makers for software and services at their organization. Quotas were set by company size, including Small (5-24 PCs), Lower Mid-Market (25-49 PCs), and Core Mid-Market (50-249 PCs), with results weighted using D&B estimates. A separate paper is available that describes the rest of the research study, which was focused on the adoption of hosted services, the purchase process, and related dynamics (*see Cloud-Based Hosted Services in SMBs: Adoption, Purchase Process, and Players*).

For the core tasks, all of the product offerings tested consisted of some combination of the attributes below, which for the traditional discrete choice either varied in terms of being present/absent or in the nature of the technology that was always present — either on-premise or hosted.

- Email (hosted versus on-premise)
- VoIP (presence/absence)
- Collaboration Tools (presence/absence)
- File Storage (hosted versus on-premise)
- Productivity Apps (hosted versus client-based)
- Brand (Microsoft, Google, Oracle)
- Price

For the canned SKU and explicit scenario tasks all features (except price in the canned SKU task) are fixed, and thus levels had to be chosen a priori. We decided to focus on scenarios involving mostly hosted services, and productivity applications. So we tested SKUs involving all hosted solutions from Microsoft and Google, as well as solutions involving all hosted services, except productivity applications from Microsoft and Oracle. The explicit scenario screen that we tested is shown in Figure 4, including the specific price points that were tested. For the canned SKU and explicit scenario tasks the SKUs were not rotated within a screen. For the traditional and canned SKUs we approached pricing differently, again in accord with how those tasks are typically designed. The prices tested were in the currency annual per user prices.

SKU1	SKU2	SKU3	SKU4	None
• Hosted Email • VoIP • Hosted Collab Tools • Hosted File Storage • Hosted Productivity Apps	• Hosted Email • VoIP • Hosted Collab Tools • Hosted File Storage • Hosted Productivity Apps	• Hosted Email • VoIP • Hosted Collab Tools • Hosted File Storage • Client-Based Productivity Apps	• Hosted Email • VoIP • Hosted Collab Tools • Hosted File Storage • Client-Based Productivity Apps	
• Microsoft • $200	• Google • $75	• Oracle • $150	• Microsoft • $300	

Figure 4

The specific price points tested included:

- Traditional Task: $75, $150, $225, $300
- Canned SKU Task:
    - Microsoft w/ Web Apps: $100, $150, $200
    - Google: $50, $75, $100
    - Oracle: $100, $150, $200
    - Microsoft w/ Office Client: $200, $250, $300

The traditional discrete choice task consisted of a total of 12 screens per participant, while the canned SKU discrete choice consisted of a total of 9 screens per participant. Finally, these instructions were provided to participants, along with some definitions around the technology at hand (not shown here):

"For the next three tasks we would like you to imagine that in 6-12 months you are faced with the task of overhauling your company's technology ecosystem for one reason or another. Please answer the following questions in light of what you know about your organization's current ecosystem, commitments, staff resources, actual requirements, and cultural tendencies.

Next you will see different bundles of technology services consisting of different combinations of capabilities and features offered at different prices by different companies. Please select the one offering that you feel would be most appropriate for your organization."

Results 1: Comparing Tasks

The first objective related to how the results of the various tasks compare, with a focus on the pattern of preference share estimates produced in each case. While there are a number of scenarios that can be simulated using the results of the two discrete choice tasks, by toggling the various parameters, there is only one simple comparison that can be made across all three of the tasks examined, including the single explicit question, so we start there.

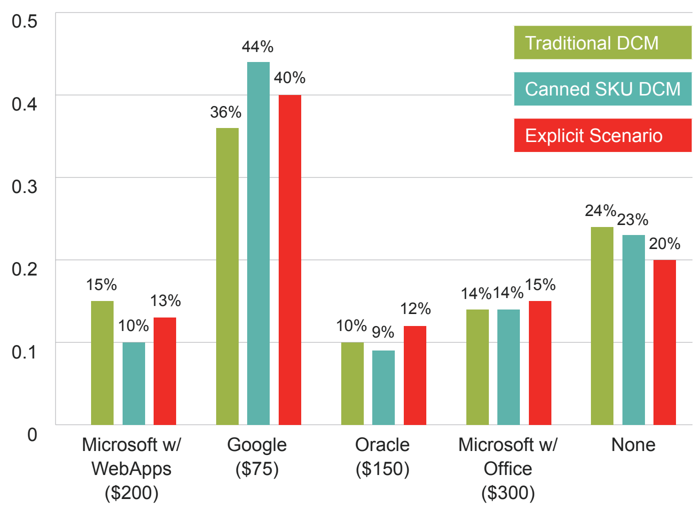


Figure 5

Recall that the explicit scenario task consisted of four specific solutions at four distinct static prices. The specific solutions compared involve all inclusive hosted versions from Microsoft and Google, and all hosted services except client-based productivity apps from Microsoft and Oracle. The graph depicts market share estimates for all the three tasks, with the parameters of the two discrete choice tasks matched to that of the explicit scenario.

The results for the traditional and canned DCM tasks represent the model-based estimates that were simulated. The results for the explicit scenario are simply the percentages of participants that chose each alternative. Despite the different paths to estimating this scenario, there is a reasonable amount of consistency in the pattern of results. In all tasks Google is the overall preferred vendor for such a solution, followed by Microsoft, which has two products in this scenario that collectively determine its share, followed by Oracle as a distant third.

The largest gap between tasks (and only one close to being statistically significant) is for Google, where there is an 8% difference in predictions between the traditional and canned tasks, with Google chosen more often in the context of the canned SKU; it appears to mostly swap share with the Microsoft With Web Apps solution, which is 5% higher for the Traditional task.

Additional comparisons can be made between the two discrete choice tasks by simply setting some of the parameters differently. We tweaked the prices of the two Microsoft offerings — Figure 6 depicting a scenario with the Microsoft with Web Apps priced $100 less, while Figure 7 depicts the Microsoft with Client Office solution priced $100 less. All other prices and parameters were set the same as in the first comparison. While the absolute numbers shift in accord with the changes, the pattern for the first scenario — as far as gaps between tasks — is similar to that observed for the first scenario, with Google higher for the canned SKU tasks, and the Microsoft With Web Apps solutions appearing to swap share.

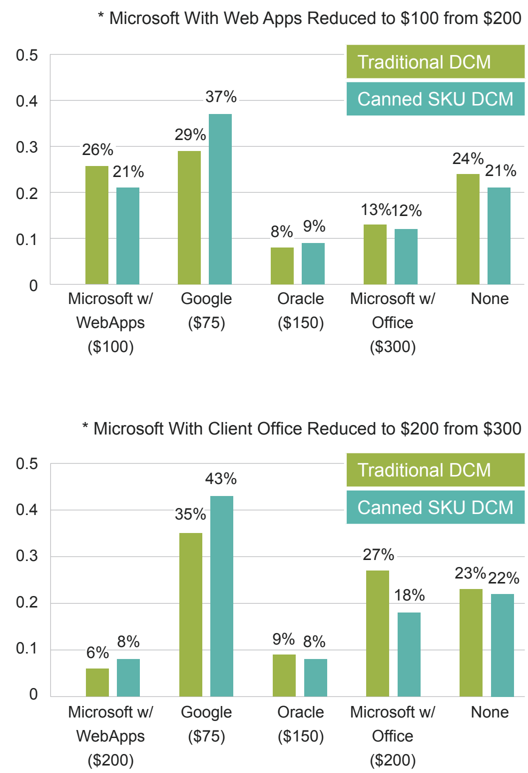


Figure 6 & 7

For the second scenario examined, the largest gap between tasks is for the Microsoft With Client Office SKU, whose price was tweaked here, with a larger percentage of participants choosing this SKU when presented in the context of the traditional discrete choice task. The general pattern is the same as seen with the original scenario we examined, with Google chosen more often in the context of the canned SKU task, with a resulting 8-9% gap in share estimates for the different tasks.

The question of which estimates are more *valid* is difficult to answer given the previously cited issues around the hypothetical framing of the task, and thus the lack of definitive external data to compare this data to. We did however ask one additional question in the survey that can serve as a comparison outside of the three key tasks themselves, providing at least an alternative check on internal consistency at least.

Toward the end of the survey we asked participants a simple high-level question about their one preferred provider for a suite of hosted services, divorced from the issue of price. Options included *Microsoft, Oracle, Google, Cisco, Apple, VMWare, HP, Dell, Yahoo, Amazon, IBM,* and *Local Provider*. The question that was posed was *"If each of the below companies offered a portfolio of hosted offerings that suited your organization's needs, which ONE company would you prefer to purchase the services from?"* The below graph depicts the percentage gaps between the vendors chosen in three tasks and the vendors chosen for that question — just for the three vendors that overlapped (Microsoft, Google, and Oracle).

For the purposes of this comparison, the two Microsoft SKUs tested in the key tasks are combined here to align with the high level survey question, and the explicit scenario settings are once again used in order to include all three tasks in the comparison. Given the larger list of providers for this question, each participant that chose anything other than Microsoft, Google, or Oracle would implicitly be part of the gaps, and thus the focus should be more around which task fares best — relatively speaking, not so much the absolute level of agreement.

As can be seen in Figure 8, each of the tasks varies from this other survey question in a very similar fashion — Google was chosen more frequently and Microsoft less frequently in each of the three key tasks. The key difference of course is that price implications are included in three core tasks, but not this high level survey question. The Microsoft solutions are much more expensive than the Google solution in those tasks, and thus the pattern (i.e. less Microsoft and more Google in tasks with price included) is as might be expected. In comparing the price points/bands used by task, the explicit scenario question paints Microsoft as most expensive, followed by the canned SKU, then followed by the traditional discrete choice (recall one of the advantages of the canned SKU approach is the ability to better mimic actual prices by using distinct price bands per SKU). The pattern observed is partially consistent with what we might expect, with the traditional task deviating the least from the high-level survey question. Thus, when compared to this other survey question at least, none of the tasks really stand out as better than the others — particularly once the price bands/ points consideration is taken into account.

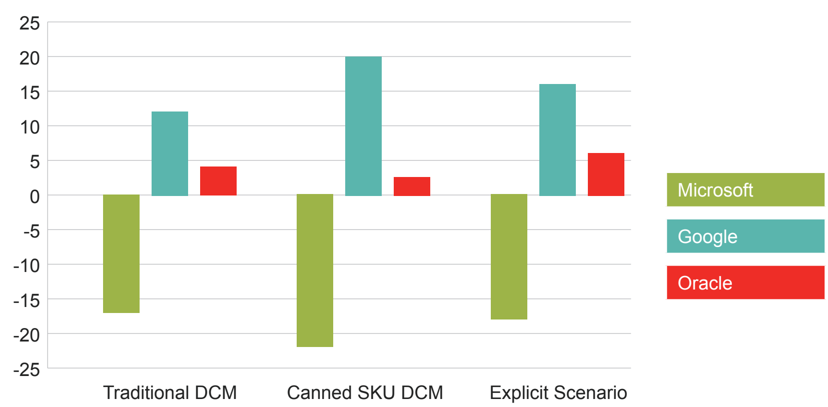


Figure 8 Gaps with Other Survey Question

Results 2: Order Effects

Two metrics are leveraged to look at order effects:

- The pattern of preference share estimates as a function of order.
- A gauge of the consistency of responding by task position, using a statistical measure we refer to as the *RLH* estimate[3].

We start with the latter. The logic of looking at this consistency of responding metric is that regardless of the exact pattern of preference share results, participants may behave more or less consistently in earlier or later tasks. Consistency of responding is of course interpreted as a sign of engagement/attentiveness to the task at hand, so seeing some gauge of this is useful in the context of our question about order effects. The RLH estimate provides this, with the higher the estimate the more consistent the pattern of responses.

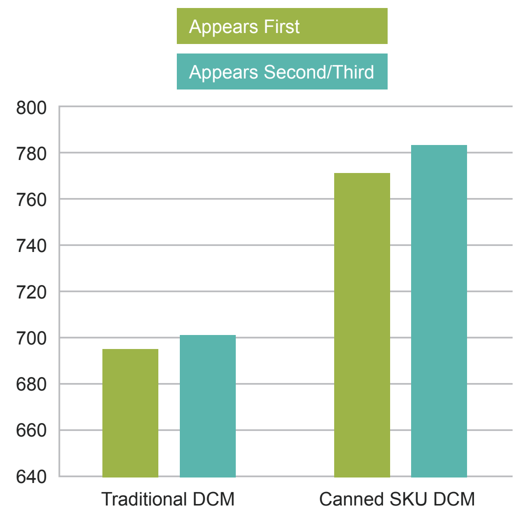


Figure 9 Consistency of Responding by Task Position (RLH Estimate)

The key question for our purposes is whether the RLH estimates for the discrete choice tasks vary as a function of task position. Figure 9 depicts the RLH estimates for the traditional and canned discrete choice tasks by task order, collapsing across the second and third positions given the minimal differences observed in general between the 2nd and 3rd positions. As can be seen, there are minimal differences in the RLH estimates by task position, with the larger difference between the task types. Statistical tests confirmed the lack of a difference by order. So this result suggests that at least at the general level of consistency of responding, the exact position of the task does not seem to have a big impact, at least in the context of survey with a reasonable amount of introductory screening and profiling questions appearing prior to any of the discrete choice tasks.

---

[3] RLH is short for "root likelihood" and measures the goodness of fit. To compute RLH we simply take the nth root of the likelihood, where n is the total number of choices made by all respondents in all tasks. RLH is therefore the geometric mean of the predicted probabilities. If there were k alternatives in each choice task and we had no information about part-worths, we would predict that each alternative would be chosen with probability 1/k, and the corresponding RLH would also be 1/k. RLH would be one if the fit were perfect.

We now look at how task position affects the pattern of preference share results (figures 10,11 and 12). The graphs below depict the preference share results by task position for the same focal scenario used previously (that used for the explicit scenario task), which again provides a common ground on which to compare all three tasks. There are two notable patterns:

- Results vary depending on whether the task appears first or not, while the difference between appearing second versus third is minimal. This pattern holds for all three tasks.
- There is a discrepant pattern of order effects by task, with the traditional discrete choice task showing one pattern, and the canned SKU discrete choice task and explicit scenario tasks showing another.

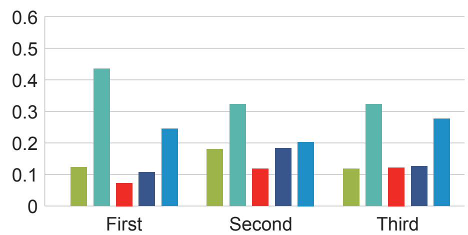


Figure 10 Traditional DCM

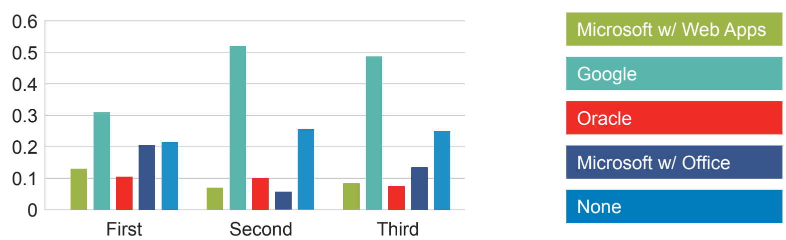


Figure 11 Canned SKU DCM

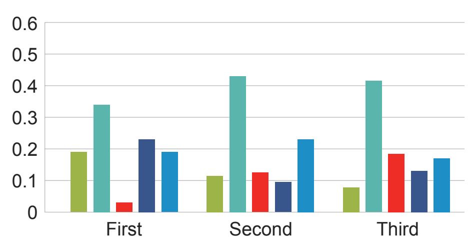


Figure 12 Explicit Scenario

For the traditional discrete choice task Google is chosen much more often when that task appeared first, mostly trading off with Microsoft's two products. The pattern is the opposite in essence for the canned SKU discrete choice and explicit scenario tasks, with Google chosen less often when that task appeared first, again with most of the tradeoffs involving the two Microsoft SKUs.

We have sought a concise explanation for the observed pattern, but it has been elusive. One hypothesis relates to the importance of the type of productivity applications available, and specifically whether they are hosted or client-based (since everyone in this population uses them). Recall that the traditional discrete choice task implicitly involves more variable feature sets and SKUs, due to the systematic testing of various permutations of attribute levels (producing, for example, client- based productivity apps from Google for some SKUs). We think that might set the tone for more openness to a solution from Google, which then carries through to the canned SKU and explicit scenario tasks, which show accordingly higher takes rates for Google when they appear after the traditional discrete choice task.

On the other hand, when the canned SKU discrete choice or explicit scenario tasks appear first, where we set the Google SKU to always include hosted productivity applications, participants for whom this is a big barrier may have “tuned out” the Google offer in favor of some other solution, which then also carried over to the traditional task when it appeared in the second or third position (where the Google take rate is also lower). While a plausible explanation, we feel that additional research that varies the specific content at hand is necessary to tease apart what is really happening (e.g. across a couple different product categories).

Regardless of the ultimate explanation, when using preference share estimates as the metric, the position of the task does appear to have an impact on the results of the tasks. In particular, the effect seems to be in the form of a first position versus later task form, which, while operating at the level of sets of screens, is interestingly reminiscent of the usual experimental finding of different behavior on the first versus other singular screens within a single discrete choice task.

Results 3: Segmentation Analyses

Our final questions relate to gaining a better understanding of what is happening behind the aggregate statistics that have been the focus thus far. Forty percent choosing Google could be driven by a 40% portion of the sample that always choose Google, or the entire sample sometimes choosing Google, on average, 40% of the time. Which is correct? How consistent were the participants across the various tasks? Are those who chose Google in the traditional task the same indivuals who chose Google in the canned SKU discrete choice?

To address the first question, two distinct sets of *latent-class segmentation* analyses were run on the data from the two discrete choice tasks. The goal was to see if there were clusters of participants responding similarly. We settled on to three-segment solutions as a comparison point.

Figures 13 and 14 depict the distribution of responses for the same focal scenario examined earlier, but broken out by the three segments that emerged in each respective analysis[4]. The traditional discrete choice task clusters have a more distributed pattern relative to canned SKU discrete choice, which has more homogeneous clusters. For the traditional discrete choice task, the most extreme segments only chose a particular alternative roughly half the time, while for two of the canned SKU task clusters, participants almost always chose the same alternative.

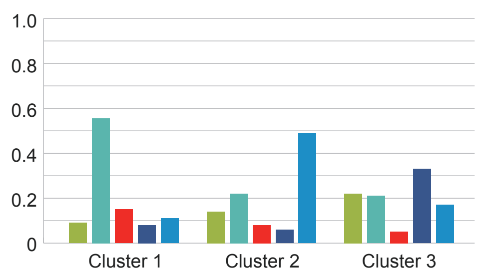
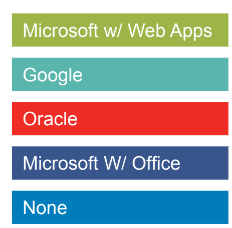


Figure 13 Traditional DCM: 3-Segment Solution



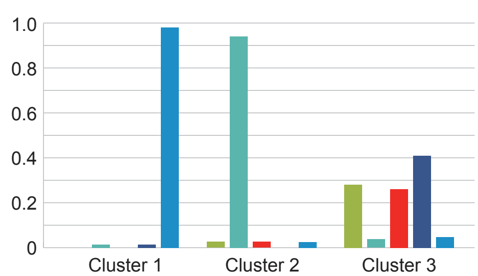


Figure 14 Canned SKU DCM: 3-Segment Solution

[4] There were two stable solutions for the traditional DCM (two and three cluster solutions), and four stable solutions for the canned SKU solution (two, three, four, and five cluster solutions) that emerged - a finding we will loop back to later in the *Discussion*.

The second key finding relates to the makeup of the segments. For both solutions, the two homogeneous segments involved Google and None, with the third segment consisting of a group of participants who had a more heterogeneous set of choices. As far as the size of the segments, the Google-biased groups are the largest, representing 41% and 45% of the sample, while the None biased groups represent 30% and 21%, with the heterogeneous third segment representing 29% and 34%, respectively, for the traditional and canned tasks.

In summary, roughly two-thirds of the participants with an affinity for Google or None chose those responses roughly half the time in the traditional task, and almost always in the canned SKU task, while the other third of participants had more distributed responses.

To address the question related to the overlap between segments, we look at the overlap between the two 3-cluster segmentation solutions that we described in the previous section. Figure 15 depicts the overlap between the three types of segments that emerged, including the Google (Traditional Cluster 1, Canned Cluster 2), None (Traditional Cluster 2, Canned Cluster 1, and Distributed (Traditional Cluster 3, Canned Cluster 3) segments[5].

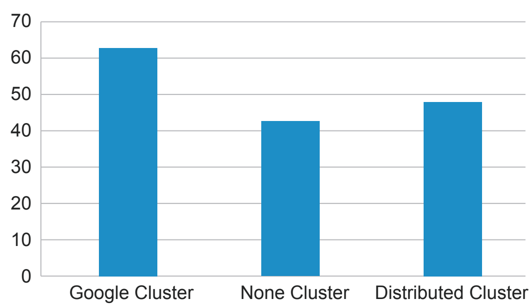


Figure 15 Percent Overlap Between Task Segmentation Solutions

About two-thirds of the "Google" segments overlapped, about 40% overlapped for the "None" segment, and about half overlapped for the "Distributed" segment. Once we take into account the more distributed pattern of responses for the "Google" and "None" in the traditional DCM segments, where participants in those segments only chose the brand that the segment was labeled after about half of the time, these overlaps don't seem so low. Thus while not perfectly overlapping there is a reasonable amount of consistency between the two discrete choice tasks in terms of which segment participants clustered into.

[5] In order to isolate the overlap, we had to make a decision about what direction to focus on (specifically, row or column percentages in the cross-tabulation). Given the consistency of responses for two of the canned SKU DCM segments, we decided to focus on the percentage overlap of the canned SKU DCM segments within each of the traditional DCM segments.

Summary

The first question posed relates to how the three tasks compare to one another when they are otherwise matched for content and framing. When it comes to overall preference share estimates, regardless of using a traditional discrete choice, canned SKU discrete choice, or explicit scenario approach to do the estimation, the results were pretty similar.

Although most attention was paid to the largest discrepancies between the tasks in the three results sections, the reality is that the traditional and canned SKU discrete choice tasks can be compared on fifteen data points in the scenarios we examined (five SKUs/options X three scenarios), with the average discrepancy across those fifteen data points being 3.6%. While that is not a nominal amount of discrepancy in some contexts (e.g. when attempting to do precise forecasting), the observed difference between tasks would probably not lead to different business decisions in most situations.

The additional comparison that was made between the tasks and another high-level survey question concerning their preferred provider for a similar hosted solution (that appeared outside the three core tasks and without price implications included) also did not suggest much difference between the three core tasks, particularly once you take into account the implicit price band difference among the three tasks. In all cases the pattern was as might be expected, with the more expensive solutions from Microsoft chosen less often when price was explicitly included. While we could have tried to match the price bands between the two discrete choice tasks more closely, we felt that the ability to leverage distinct bands in the canned SKU task is too much of a distinguishing characteristic of that approach to justify matching them exactly, so decided against doing so.

The second question relates to whether there are order effects operating when multiple discrete choice tasks are used in a single survey. Here the answer is a bit more complicated. There was no difference in the general consistency of responding by the order of tasks, at least as gauged by our RLH statistical metric, suggesting that participants didn't disengage in the later tasks in any broad sense. There was a difference in the qualitative pattern of preference share depending on order, with Google and the Microsoft solutions trading off share depending on the exact position of the task.

While we proposed a hypothesis to explain the observed pattern, regardless of the explanation, the presence and size of the effect is a little disconcerting, even if in aggregate we ended up with similar estimates. It could be that our attempts to kill two birds with one stone in this particular study (i.e. compare otherwise matched tasks AND examine order effects) actually exaggerated the influence of one task on the other because of the exact overlap of content. So perhaps this represents the extreme case, with our more usual scenario involving non-overlapping content being less of an issue.

Additional research is really needed where we vary the nature of the tasks involved to overlap in content or not, as well as the product category, to determine if this is a general effect or something more about the particular manipulations or context examined here.

The substantive differences observed for the segmentation analyses deserve comment. We saw a difference between the traditional and canned SKU segments in terms of how concentrated responses were by segment, with two of the three segments of the canned SKU solution choosing the same response most of the time. Thus while the aggregate-level preference share estimates were similar, the dynamics in terms of what participants were actually doing was very different.

Other facts speak to this same observed difference between tasks:

- The higher RLH estimate for the canned SKU task we saw earlier in the context of the order effects results is another expression of this consistency in participant responding.
- Although not examined in detail here, the fact that there were additional segmentation solutions of four and five clusters that emerged for the canned SKU task (with the implication that the distributed third segment we saw could be partitioned further into even more homogeneous segments), but not for the traditional discrete choice, is yet one other way that we see this trend operating.

There are several potential explanations for why participants might respond more consistently in the canned SKU task, but the one implicit difference between tasks that is likely the best explanation is that the SKUs in the canned SKU task are much more distinguished in terms of price — the most important consideration in many discrete choice tasks, including these (i.e. price accounts for about half the variance in choice behavior in both tasks). As alluded to earlier, this ability to map the prices more realistically against SKUs is one of the advantages and defining features of this approach, and thus we allowed price to vary by SKU. For the same cognitive load considerations, the location of the static SKUs is typically not varied across screens. So in some sense we are *encouraging* the consistency of response by virtue of how the task is implicitly structured. Additional research is needed to tease apart what role the distinct price bands play in this versus the other defining characteristics of the canned SKU task, such as the static features of the SKUs, and the static location of the SKUs.

Implications

- The similarity of the overall preference share estimates produced by the three different types of tasks suggests that any one of the approaches is a viable estimation technique, with the choice of approach best driven by the requirements at hand, such as whether features need to be toggled as parameters, whether price points need to be varied, etc., as well as the survey real estate available.
- While the observed order effects need to be confirmed across some additional contexts, the patterns suggest there is a cross-task influence that needs to be taken into account in the design and placement of survey tasks, with more research needed to tease apart the dynamics.

About the Author

Brian Cavoto, PhD, Managing Partner

Brian has over 20 years research experience. Prior to YouGov Definitive Insights he was Senior Partner at Momentum Market Intelligence, Head of custom research for Harte-Hanks Market Intelligence and Expert Consultant for SPSS.

Brian's BA is from University of Connecticut. His PhD was awarded at Tufts University where he studied cognitive psychology and published in multiple top-tier psychology journals.

Contact Us

Please contact us if you have any questions or interest in our research services:

Brian Cavoto
email: brian_cavoto@definitiveinsights.com
phone: 503-575-7630 ext. 803