

RECONCILING THE PARADOXICAL FINDINGS OF CHOICE OVERLOAD THROUGH AN ANALYTICAL LENS¹

Nan Zhang and Heng Xu

Department of Information Technology and Analytics, Kogod School of Business,
American University, Washington, DC, U.S.A. {nzhang@american.edu} {xu@american.edu}

Too much of a good thing can be harmful. Choice overload, a compelling paradox in consumer psychology, exemplifies this notion with the idea that offering more product options could impede rather than improve consumer satisfaction, even when consumers are free to ignore any available option. After attracting intense interest in the past decades from multiple disciplines, research on choice overload has produced voluminous yet paradoxical findings that are widely perceived as inconsistent even at the meta-analytic level. This paper launches an interdisciplinary inquiry to resolve the inconsistencies on both the conceptual and empirical fronts. Specifically, we identified a surprising but robust pattern among the existing empirical evidence for the choice-overload effect and demonstrated through mathematical analysis and extensive simulation studies that the pattern would only likely emerge from one specific type of latent mechanism underlying the moderated choice-overload effect. The paper discusses the research and practical implications of our findings—namely, the broad promise of analytical meta-analysis (an emerging area for the use of data analytics) and machine learning to address the widely recognized inconsistencies in social and behavioral sciences, and the unique and salient role of the information systems community in developing this new era of meta-analysis.

Keywords: Analytical meta-analysis, inverted-U relationship, curvilinear model, moderation analysis, choice overload

Introduction

The very essence of scientific progress is the systematic accumulation of knowledge, yet the viability of doing so is being questioned in many disciplines in social and behavioral sciences such as psychology (Open Science Collaboration 2015) and economics (Camerer et al. 2016). Critics point to many conflicting findings on the same research question in the social and behavioral sciences, echoing a famous quote by United States Senator Walter “Fritz” Mondale: “For every study that contains a recommendation, there is another, equally well documented study, challenging the conclusions of the first” (Bangert-Drowns and Rudner 1990, p. 1). Without the ability to collectively reason about such conflicting findings, we often see tens, even hundreds, of studies examining the same research question from different angles, only to make the picture even murkier. Choice overload is a “poster child” of such paradoxical research questions with notoriously inconsistent findings (Scheibehenne et al. 2010). This paper launches an

interdisciplinary inquiry into the feasibility of leveraging advanced analytical techniques to collectively reason about the conflicting findings for the choice-overload effect in the behavioral research literature. While most of the current paper focuses on choice overload as the case study, our broader goal, as elaborated in the discussion section, is to demonstrate the power of *analytical meta-analysis*, the use of data analytics and machine learning techniques in synthesizing the inconsistent findings of behavioral research, and to highlight the unique and salient role the information systems (IS) community may play in building and expanding its methodological arsenal.

In the rest of the introduction, we first review the remarkable inconsistencies in the literature of choice overload, before providing an overview of our novel analytical approach. We conclude the section with a summary of the intended contributions of the paper from both substantive and methodological perspectives.

¹ H. R. Rao was the accepting senior editor for this paper. / Xiao Fang served as the associate editor.

Inconsistencies in the Choice-Overload Literature

Choice overload is a compelling paradox in consumer psychology that has attracted intense research interest in the past decades. It captures the idea that offering more options could impede rather than improve consumer satisfaction, even when consumers are free to ignore any available option. Following the dramatic evidence of choice overload in the marketing context (Iyengar and Lepper 2000), a series of studies ensued in psychology (e.g., Inbar et al. 2011), marketing (e.g., Chernev 2003), and IS (e.g., He et al. 2019), only to present a complex picture with widely dispersed effect sizes and paradoxical findings even at the meta-analytic level. While some meta-analyses found strong evidence of the choice-overload effect and identified its moderators (Chernev et al. 2015; McShane and Böckenholt 2018), others contended a failure of replication and a general lack of empirical support for the effect (Scheibehenne et al. 2010; Simonsohn et al. 2014). These contradictory results leave the empirical understanding of choice overload in a fragmented state.

The conceptual underpinning of choice overload is similarly fragmented. On the one hand, the presence of choice overload naturally makes the option-satisfaction relation² into a curvilinear inverted-U model (Grant and Schwartz 2011; Reutskaja and Hogarth 2009) because offering more options would be beneficial to consumer satisfaction up to a point before becoming negative. On the other hand, most existing studies have conceptualized the choice-overload effect *no differently* than a linear relation and examined it through a classic two-group experimental design.³ Further, moderation in choice overload⁴ has been exclusively modeled using a linear mechanism without discerning the two theoretically distinct yet equally prevalent moderation types for a curvilinear relationship: whether the moderator steepens/flattens the curve, or shifts the turning point⁵ of the curve left/right⁶ (Haans et al. 2016). Like the fragmented view at the empirical level, the theoretical understanding of choice overload is similarly fragmented, representing a considerable gap in the literature.

Given the practical implications of choice overload in an e-commerce context (e.g., its importance in illuminating how consumer satisfaction may vary with the number of products displayed on an online browsing interface; Johnson et al. 2012), it has also received considerable attention in the IS literature. For example, cognitive overload in information processing has long been noted in IS research (Doll and Torkzadeh 1988; Driver and Streufert 1969), even before the emergence of choice overload in the marketing literature. Nonetheless, most existing IS studies have approached the issue from the perspective of information quality (Doll and Torkzadeh 1988; Geva et al. 2019; Jones et al. 2004), which is of paramount importance for the design and processing of complex information but does not afford direct guidance on the *number* of options to offer in an e-commerce context. Indicatively, some studies (e.g., Pan et al. 2013) contend that more options are always preferred in online browsing (i.e., choice overload does not exist), citing the inconsistent meta-analytic findings discussed earlier (Scheibehenne et al. 2010); while others assume a detrimental effect of having too many options displayed on the interface (e.g., Geva et al. 2019). In sum, the conflicting views of choice overload in consumer psychology are also present in IS research.

A Novel Analytical Approach

The fragmented views of choice overload directly result from the paradoxical combination of an obviously nonlinear (i.e., inverted-U) conceptualization of the effect and a methodological deficit at the empirical front to examine such an effect, most prominently the lack of a meta-analytic method that can synthesize the results of two-group experiments to probe the characteristics of a nonlinear relationship. Thus, we premise the reconciliation of the fragmented views on developing a novel analytical link between the conceptual underpinning of choice overload—i.e., the notion of the option-satisfaction relation being an inverted-U if choice overload exists and monotonic if it does not—with the unique analytical patterns discernible from the observed inconsistencies of the existing empirical findings. Ideally, such a link should explicate the

² That is, the relationship between the number of available options and consumer satisfaction.

³ With such a design, a researcher compares the satisfaction of two groups of individuals treated with two assortment sizes (small vs. large), respectively. All primary studies included in the recent meta-analyses of choice overload (e.g., McShane and Böckenholt 2018) belong to this category.

⁴ The existing literature has noted many potential moderators for choice overload. For example, (Chernev 2003) identified a consumer's familiarity with the available options as a moderator variable, as an increasing level of familiarity likely attenuates the choice-overload effect.

⁵ The turning point is the point at which the inverted-U curve reaches its maximum. It has also been referred to as the "inflection point" in the literature (Grant and Schwartz 2011). Since mathematically an inverted-U function reaches its peak at the turning point (where the first derivative equals zero), not the inflection point (where second derivative equals zero), we use the term "turning point" throughout this paper for the purpose of consistency.

⁶ This conceptual discrepancy is particularly perplexing in the context of choice overload, given that differentiating the two moderation types is salient for firms to assess the importance and feasibility of finding the "right" number of options in their product design. Specifically, the importance of choice overload is directly linked to the steepness of the curve, while the feasibility of finding a "right" assortment size is linked to the left/right variation of the turning point.

mechanisms through which the former entails the latter, enabling us to address the fragmented empirical view by elucidating how an underlying conceptual model could actually explain and account for the empirical inconsistencies. Similarly, we would reconcile the fragmented conceptual view by leaning on the coalescent existing empirical evidence to statistically unpack the underlying option-satisfaction relation, determining whether it is monotonic or an inverted-U and delineating the type(s) of moderation that should be expected.

Methodological research in adjacent fields has illustrated how curvilinearity can be linked to the paradoxical findings of correlational studies when range restrictions occlude the underlying effect (Pierce and Aguinis 2013). Still underdeveloped in the literature, however, is the analytical link between curvilinear relations and the outcomes of two-group experimental designs, like those used in the existing empirical studies of choice overload. We submit that the existence of this gap speaks to the interdisciplinary challenge facing the development of the analytical link. On the one hand, researchers in disciplines such as mathematics are oblivious to the gap as they rarely use behavior research methods like two-group experiment designs. On the other hand, behavioral researchers lack the analytical tools necessary to identify nonlinear effects from the findings of two-group experiments. In this paper, we tackle this interdisciplinary challenge by developing a novel *analytical pattern of inconsistencies* among the outcomes of two-group experiments that can only emerge if *both* of the following two conditions are met: First, the underlying option-satisfaction relation must be an inverted-U rather than monotonic (e.g., linear or curvilinear-yet-monotonic). That is, the choice-overload effect must exist. Second, the moderation of the option-satisfaction relation must substantially shift the turning point of the inverted U left/right, instead of merely steepening/flattening the curve. We prove the identifiability of the pattern mathematically and demonstrate its statistical power through extensive simulation studies, before confirming its presence in the empirical evidence for the choice-overload effect.

Summary of Contributions

For the specific phenomenon of choice overload, this paper makes substantive contributions by clarifying the basis for *expecting* the emergence of inconsistent findings from two-group experimental designs examining the choice-overload effect. We also offer strong evidence that the choice-overload effect *does* exist since a monotonic option-satisfaction relation could not have produced the pattern of inconsistencies in the existing empirical evidence. Further, we take care to specify how the two conceptually distinct moderation types manifest as clearly distinguishable patterns among the experimental findings. Given the voluminous evidence of moderation in

choice overload, the appreciation of how moderation operates is not only of practical pertinence but also crucial for the design of future research.

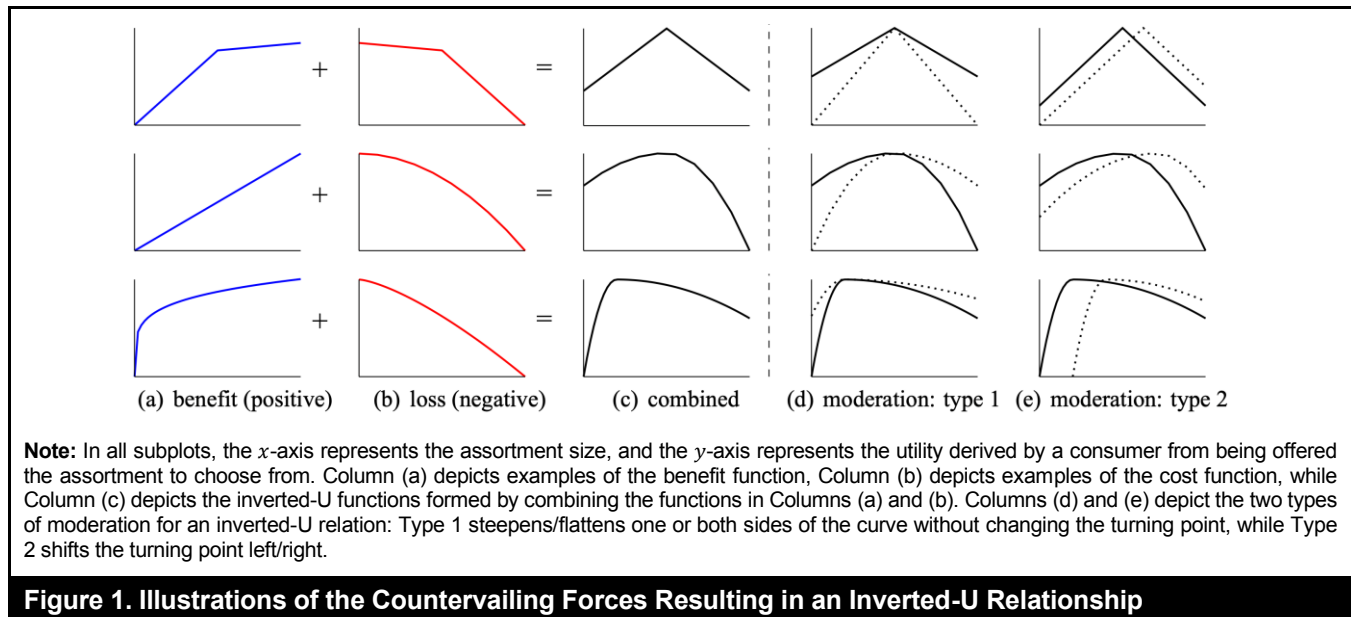
This paper also contributes methodologically to the broader research agenda of *analytical meta-analysis*, i.e., the use of data analytics and machine learning techniques to enable the collective reasoning of (potentially inconsistent) findings in the literature of social and behavioral sciences (Zhang et al. 2020). To this end, we provide one of the first analytical methods that can establish the presence of an inverted-U relationship and its moderation type based on experimental studies that were *not* specifically designed to test a curvilinear relation. This affords behavioral researchers an opportunity to leverage the existing empirical evidence in exploring, theorizing, and testing curvilinear relations before new experiments dedicated to the examination of curvilinearity can be conducted and accumulated over time (to allow for a comprehensive meta-analysis in the future). Further, the effectiveness of the analytical method developed in the paper demonstrates the promise of casting an analytics lens into the collective reasoning of conflicting findings in empirical studies. It is our belief that, as a scholarly community with a diverse set of methodological roots, IS researchers are uniquely positioned to contribute to the development of analytical meta-analysis and to accelerate the systematic accumulation of knowledge in a broad range of behavioral research.

Conceptual Development

In this section, we develop the conceptual foundation linking the theoretical underpinning of the choice-overload effect and its moderation to the empirical outcomes of two-group experiments. The mathematical model and simulation studies for the conceptual arguments will be presented in the sections that follow. Since the presence of choice overload would imply an inverted-U option-satisfaction relation, we first review the conceptualization of an inverted-U relationship, including its two distinct moderation types, in the literature. Then, we outline the alternative conceptualizations for choice overload, and link each of them to the empirical outcomes of two-group experiments, before developing a unique pattern of inconsistencies among the empirical outcomes that could have emerged from only one of the alternative conceptualizations.

Conceptualization of an Inverted-U Relationship

A relationship is an inverted-U if the dependent variable Y first increases with the independent variable X and then decreases once X reaches the “turning point.”



For the option-satisfaction relation, the assortment size and the satisfaction of consumers are the independent and dependent variables, respectively. Since the psychology literature has unequivocally confirmed people's preferences of having some options (e.g., $X = 2$) over no option at all (i.e., $X = 1$), the option-satisfaction relation would be destined to be an inverted-U should choice overload exist (Coombs and Avrunin 1977; Grant and Schwartz 2011; Reutskaja and Hogarth 2009).

An inverted-U relationship is often conceptualized as “the resultant of two opposed mediating processes, both monotonic” (McGuire 1997, p. 23) so as to explicate the two countervailing forces that jointly constrain and influence the relation. For the option-satisfaction relation, the monotonically increasing function captures the benefits associated with a larger assortment, such as the potential to offer consumers a better match with their personal preferences (Baumol and Ide 1956), to accommodate consumers' variety-seeking behavior (Levav and Zhu 2009), to create a perception of freedom of choice (Kahn et al. 1987), etc. The monotonically decreasing function, on the other hand, captures the negative implications of a larger assortment, e.g., by triggering “buyer's remorse” (Schwartz et al. 2002), by requiring additional cognitive costs for evaluating the alternatives (Roberts and Lattin 1991), by inflating consumers' (unreasonable) expectations of finding an “ideal” option (Diehl and Poynor 2010), etc. Figure 1(a)-(c) illustrates several examples of how two monotonic functions can be additively combined to form an inverted-U relationship. As

can be seen from the figure, the two monotonic functions could follow a variety of shapes, from linear functions to functions of diminishing returns or accelerating losses.

In terms of the moderation of an inverted-U relationship, Haans et al. (2016) developed a typology of two different types of moderating effects, which are illustrated in Figure 1(d) and (e), respectively. It is important to note that the two types of moderations often operate in tandem in practice. The first type steepens or flattens the curve, like how moderation in a linear mechanism increases or decreases the slope. More flexible than the linear case though, a moderating effect on an inverted-U relationship could steepen/flatten both sides of the inverted-U, or flatten one side while steepening the other, as can be seen from Figure 1(d). The second type of moderation shifts the inverted-U curve (and its turning point) left or right. Unlike the linear case in which a left/right shift has no influence on the observed effect,⁷ shifting an inverted-U left or right could vary the observed effect considerably due to the change of the turning point.

For the specific case of choice overload, the existing meta-analyses (Chernev et al. 2015; McShane and Böckenholt 2018) synthesized four types of moderator variables: decision goal, preference uncertainty, decision task difficulty, and choice set complexity. The first two types of moderators, decision goal and preference uncertainty, reflect the intrinsic, idiosyncratic, factors associated with individual decision makers. Specifically, *decision goal* captures whether consumers approach the assortments with the goal of making

⁷ Since shifting a linear relation left/right does not change the slope, its theoretical impact on the effect size is zero.

a choice (i.e., “purchasing”) or merely to understand the available options (i.e., “browsing”). In the case of browsing, the benefit (i.e., increasing) function is likely less steep, as it is now limited to the pleasure from the evaluation process (Kahn and Wansink 2004) rather than the gain from making the “right” purchase decision. Meanwhile, the pressure associated with decision-making also dissipates, flattening the cost function as well (Chernev and Hamilton 2009). Combining the two changes, we can expect the inverted-U curve to be flattened for browsing than for purchasing, despite the two curves potentially sharing a similar turning point. In other words, the first type of moderation (i.e., steepening/flattening) likely dominates the effects of moderators in the decision goal category.

Preference uncertainty refers to the extent to which consumers have articulated preferences with respect to the selection decision, e.g., whether they are familiar with the characteristics of the available product offerings (Chernev 2003). For consumers who are familiar with the available options, the cost function is likely pushed to the right, as their familiarity affords them the opportunity to assess a larger number of options before incurring an onerous cognitive rumination (Mogilner et al. 2008; Morrin et al. 2012). Corresponding to this delayed cost, the turning point is also shifted to the right, exerting the second type of moderation, i.e., a shift of the turning point.

The latter two types of moderators, decision task difficulty and choice set complexity, reflect the extrinsic factors associated with the decision task itself. Specifically, *decision task difficulty* captures the structural characteristics of the decision problem; while *choice set complexity* captures the complexity of the available choice options. Any time constraint on decision-making, for example, is a factor that belongs to the category of decision task difficulty. Whether there is a dominant option among the available choices, on the other hand, belongs to the category of choice set complexity. Both factors are likely to affect the steepness of the inverted-U curve. Time constraints are known to rapidly increase the cognitive load associated with decision-making (Bettman et al. 1998), steepening the cost function incurred by cognitive overload. On the flip side, having a dominant option can ease the decision-making process considerably (Huber et al. 1982), flattening the cost function instead. While this is consistent with the first type of moderation effect, it is important to note that both factors are also likely to shift the turning point. Time constraints have been shown to incur greater regret when an individual chooses from a larger assortment than a smaller one (Inbar et al. 2011), conceivably

shifting the turning point to the left. On the other hand, increasing the assortment size with the presence of a dominant option has been shown to enhance the dominance of the option and boost consumers’ satisfaction with choosing the dominant option (Dhar 1997), suggesting a shift of the turning point to the right when a dominant option is introduced. In sum, these latter two types of (extrinsic) moderators likely induce both types of moderating effects, steepening/flattening the curve while simultaneously shifting the turning point.

The two types of moderation have distinct practical implications for the choice-overload effect. The first type of moderation reveals the factors that make the effect more (or less) important in practice because a flat option-satisfaction curve renders the assortment size inconsequential for consumer satisfaction. With the aforementioned moderator of decision goal, choice overload is clearly more important when a firm is designing a website for direct purchase than a brochure for casual browsing. The second type of moderation, on the other hand, indicates the critical factors to consider in designing assortments for product offerings. For example, a firm may want to limit the number of product options when few consumers are familiar with the product, and gradually increase the assortment size when familiarity grows. Since the two types of moderation are theoretically and empirically distinct, they should be clearly distinguished in theoretical development and empirical examinations (Haans et al. 2016).

Linking Conceptualizations to Existing Empirical Evidence

To reconcile the fragmented views of choice overload, we seek to link its conceptual model with the effect sizes reported in the literature, in order to explain *what type of an option-satisfaction relation, combined with what type(s) of moderation,⁸ could have caused the observed heterogeneity among the existing empirical findings*. The emergence of heterogeneity is indeed a common issue in meta-analyses. In general, when the observed heterogeneity exceeds what could be explained by artifactual factors such as sampling error,⁹ researchers theorize and test moderator variables to account for the residual heterogeneity (Thompson 1994). The unique challenge for choice overload, as discussed earlier, is the dearth of meta-analytic methods that can probe the characteristics of a curvilinear relationship based on the reported findings of two-group experiments (Pierce and Aguinis 2013).

⁸ Throughout this section, we use the term “moderation” in a broad sense, encompassing not only theory-predicted moderator variables but also other study-level characteristics that could alter the option-satisfaction relationship.

⁹ The presence of such residual heterogeneity has been unequivocally established in the meta-analyses for choice overload (e.g., Chernev et al. 2015; McShane and Böckenholt 2018).

To address this challenge, we must extend beyond the simple aggregation of existing findings into statistical indicators, but develop analytical models that can glean *deeper* insights about how the moderation of a linear or curvilinear relationship manifests as the heterogeneity observed in the existing findings. In the passages that follow, we first outline the key conceptual question we aim to answer from the synthesized findings. Then, we develop three analytical models, corresponding to three possible answers to the conceptual question, respectively. Finally, we identify a unique pattern of inconsistencies among the reported effect sizes that could only have emerged from one of the three models. We acknowledge at the onset that, for the sake of clarity, it is necessary to gloss over several operationalization challenges in the conceptual development. At the end of this section, we summarize these challenges and how they are addressed later in the paper.

Key Question

The fragmented conceptual views of choice overload bring to the fore two pressing issues that are practically imperative: The first is whether the choice-overload effect exists at all (cf. Scheibehenne et al. 2010). The importance of this issue has long been touted in the literature (e.g., Iyengar and Lepper 2000), as it informs practitioners whether offering more options could adversely impact consumer satisfaction. If choice overload does exist, then the second issue that arises is whether the turning point of the option-satisfaction relationship may be shifted significantly by potential moderators. That is, whether the moderation operates only through the first type (i.e., steepening/flattening the curve), or whether the second type (i.e., turning-point shift) occurs *either instead of or in addition to* the first type. This issue is important because the turning point directly implicates the number of options one should offer to consumers in practice. In other words, by understanding whether the turning point shifts in moderation, we would be able to inform practitioners whether it is necessary to adjust the number of available options under different moderating conditions (e.g., for different customers, products, etc.). Considering the two issues in tandem, the key research question we aim to answer is whether the choice-overload effect exists and, if so, whether there is statistical evidence for the second type of moderation (i.e., a shift of the turning point).

Three Models

There are three possible answers to the key question: (1) the choice-overload effect does not exist, (2) the effect exists but moderation only operates through steepening/flattening of the curve, and (3) the effect exists and there are moderators that shift the turning point significantly. To offer an analytical link between the conceptual question and the existing empirical

evidence, we develop three models corresponding to the three possible answers, respectively.

We start with the case in which the choice-overload effect does not exist. For this possibility, numerous theories in social psychology (e.g., attribution theory, Kelley 1973; reactance theory, Brehm 1966) point to the option-satisfaction relation forming a monotonic function that may be linear or curvilinear-but-monotonic. This provides the underpinning for the first analytical model, an example of which is depicted in Figure 2(a). Similar to how moderation is commonly examined for a linear relationship, with a monotonic option-satisfaction relation, researchers often hypothesize that a change of moderator level triggers a slope change or even a sign reversal of the relation, e.g., from the increasing blue line to the decreasing red line in Figure 2(a). Indicatively, the effect size observed in a two-group experiment, as illustrated by the two solid line segments in the figure, would switch from positive to negative, like what has been widely reported in the literature (Chernev et al. 2015). As discussed earlier in the paper, while this model is obviously incompatible with an inverted-U option-satisfaction relation, it actually reflects how a moderating effect has been examined in most existing empirical studies for choice overload.

For the second possible answer (i.e., where the choice-overload effect is only moderated through the steepening/flattening of the curve), we consider a model of the option-satisfaction relationship as an inverted-U function, depicted in Figure 2(b). As can be seen from the figure, since the moderation does not shift the turning point significantly, the two inverted-U curves, representing the option-satisfaction relationship under two moderating conditions, always share the same turning point. Nonetheless, just like in the first model, the effect size observed in a two-group experiment could also switch from positive to negative under different moderating conditions. Thus, this second model could also entail a widely dispersed set of effect sizes, as reported in the literature.

The last possible answer to the conceptual question is that the moderation of choice overload involves a significant shift of the turning point. This answer gives rise to a wide range of possibilities in terms of how moderation operates, as moderation in this case may or may not involve a steepening/flattening of the inverted U. Figure 2(c) depicts one such possibility, where the moderating effect is limited to the horizontal shift of the turning point. As can be seen from the figure, even in this (more simplistic) case, the effect size observed in a two-group experiment could vary drastically, even flipping its sign with the shift of the turning point. In other words, all three models could plausibly produce the wide dispersion of effect sizes reported in the literature.

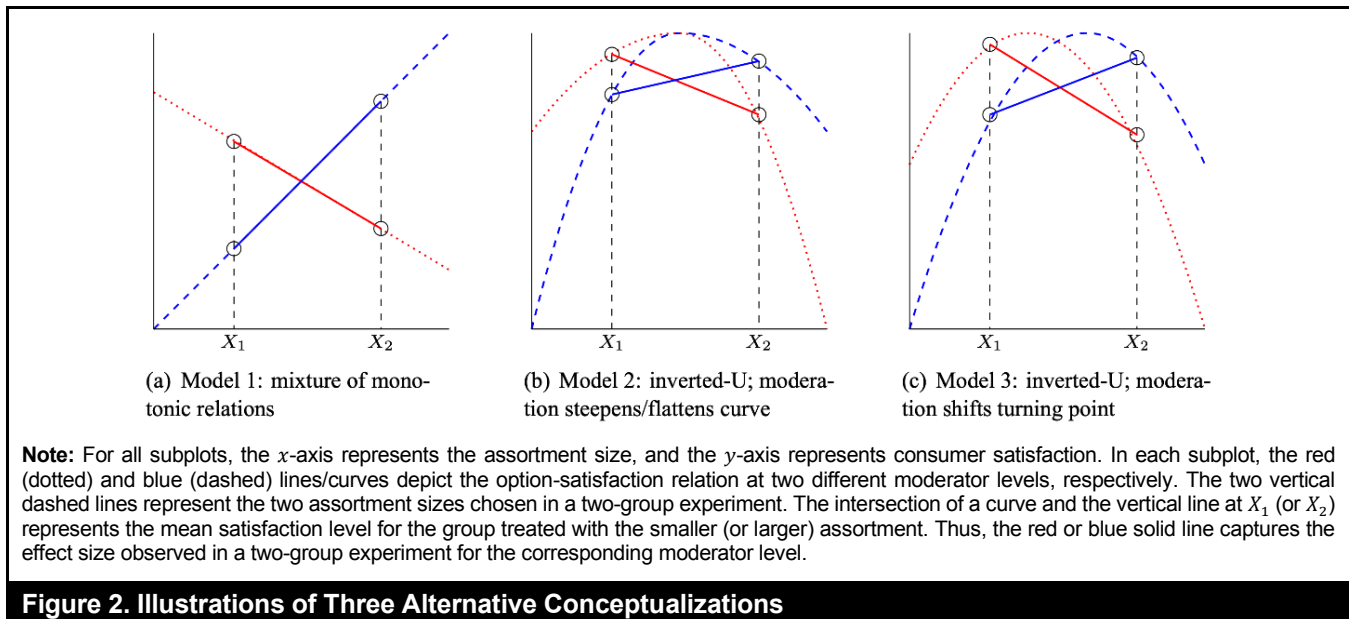


Figure 2. Illustrations of Three Alternative Conceptualizations

Unique Pattern of Inconsistencies

The discussions of the three models seemingly suggest that we have hit a dead end: As all three models could manifest as sign reversal in a two-group experiment, it appears untenable to distinguish between the three models based on the existing empirical evidence, which was mostly collected from two-group experiments. Interestingly, while this infeasibility is true for a *single* two-group experiment, it does not hold when we have access to the results of *multiple* such experiments with *different* assortment-size designs. The key reason can be explained as follows. Figure 2 suggests that, for a given pair of assortment sizes $\{X_1, X_2\}$, it is easy to find three models that all produce the same pair of effect sizes for the two moderator levels (i.e., blue and red). Nonetheless, if we now conduct another two-group experiment with another pair of assortment sizes, like X_3 and X_4 in Figure 3, then clearly the three models can no longer produce the same results.¹⁰ In other words, we can make a distinction between the three models by comparing and contrasting the results of two-group experiments with *different* assortment-size designs.

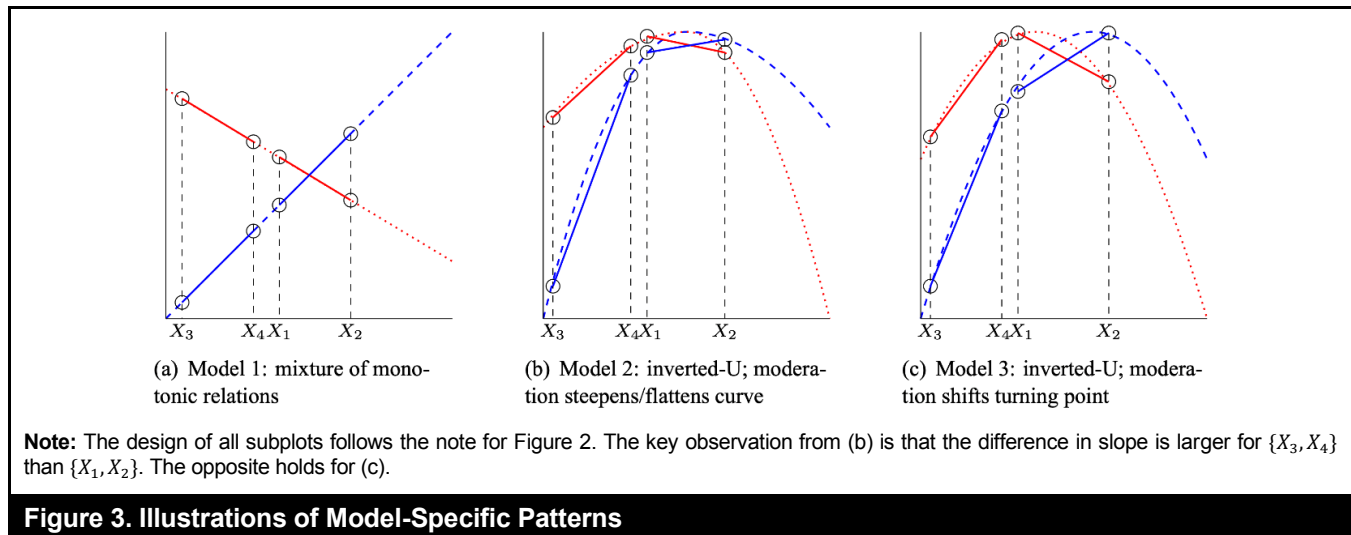
To develop a model-specific pattern, we start with a few observations from Figure 3. Consider two pairs of assortment sizes reported in two different experiments: $\{X_1, X_2\}$ and $\{X_3, X_4\}$ with equal distance (i.e., $X_2 - X_1 = X_4 - X_3$). When the underlying option-satisfaction relation

is an inverted-U (i.e., Models 2 and 3), we set X_1 and X_2 to be on different sides of the turning point X^* (i.e., $X_1 < X^* < X_2$), and X_3 and X_4 on the same side (i.e., $X_3 < X_4 < X^*$). Now, for each model, we examine how moderation affects the effect observed in a two-group experiment with assortment sizes being $\{X_1, X_2\}$ and $\{X_3, X_4\}$, respectively. Note that, since $X_2 - X_1 = X_4 - X_3$, the slope of each solid line segment in Figure 3 represents the effect size for the corresponding setup.¹¹

As can be seen from Figure 3, with the linear mechanism in Model 1, the observed effect sizes remain constant regardless of the assortment-size design. This is indeed consistent with the conventional understanding of study design in choice overload, as the numeric values of the assortment sizes were never included in any meta-analytic models examined in the literature. With Models 2 and 3, however, this conventional understanding belies the drastic change of effect sizes with varying assortment sizes. For example, with Model 2, the difference in effect size is more pronounced for $\{X_3, X_4\}$ than $\{X_1, X_2\}$, as evidenced by the large difference between the slopes of the red and blue line segments on the left side of Figure 3(b). The intuition behind this is simple. Based on the definition of a turning point, the slope (i.e., first-order derivative) of the curve must approach zero when the chosen assortment sizes are close to the turning point.

¹⁰ Mathematically, while $f(X_2) - f(X_1)$ alone cannot uniquely identify a function $f(\cdot)$ (e.g., a dotted line in Figure 2), the function can be clearly identified (from the three candidates) if we know both $f(X_2) - f(X_1)$ and $f(X_4) - f(X_3)$.

¹¹ As we elaborate in the mathematical formalism and results sections, a variety of effect sizes (e.g., Cohen's d , log odds ratio) have been reported in the literature, and can be fit into our conceptual model with proper transformations. For the purpose of conceptual development, it is appropriate to elucidate the effect size as the slope of a solid line segment.



Thus, even when a moderator changes the steepness of the inverted U, as long as the assortment sizes are still close to the (unchanged) turning point, the effect size observed in a two-group experiment must remain bounded within the vicinity of zero. In contrast, when the assortment sizes are far from the turning point, the observed effect could vary drastically, as it can now reflect the true steepening/flattening (hence the changed slopes) of the inverted U. This is exactly what can be observed in Figure 3(b).

The opposite may be true for Model 3. Consider the example depicted in Figure 3(c), where the moderating effect only entails a turning-point shift. As can be seen from the figure, the difference in effect size is now larger for $\{X_1, X_2\}$ than $\{X_3, X_4\}$. This is consistent with the conceptualization behind the moderating effect of the turning-point effect: When the assortment sizes are close to the turning point, like X_1 and X_2 in Figure 3(c), a turning-point shift could flip the sign of the observed effect with a pronounced change of slope. The slope change dissipates, however, once the assortment sizes are further away from the turning point—e.g., $\{X_3, X_4\}$ in Figure 3(c)—as the observed effect is now dominated by the steepness of the inverted-U rather than the position of the turning point. Comparing Model 3 with Model 2, a sharp contrast emerges. If moderation *only* operates through the steepening/flattening of the curve, we should expect a wider dispersion of effect sizes from experiments that select assortment sizes at either end of the spectrum. Only when the moderation entails a shift of the turning point, are we likely to observe a higher dispersion from experiments that feature assortment sizes in the middle, i.e., close to the turning point of the inverted U.

The distinct observations from the three subplots of Figure 3 give rise to the following model-specific pattern: Consider a set of existing two-group experiments with similar distance between assortment sizes, i.e., with a roughly constant $X_2 - X_1$. We sort the studies in increasing order of $\bar{X} = (X_1 + X_2)/2$, and then inspect a sliding window of k studies to assess how the dispersion (e.g., standard deviation) of the k reported effect sizes varies when $\bar{X}$ grows from small to large. If the linear mechanism (in Model 1) were true, we would observe no change of dispersion. If Model 2 were true, we would observe a U-shaped dispersion, as it is higher when $\bar{X}$ is at either end of the spectrum than when $\bar{X}$ is close to the turning point. Only if Model 3 were true would it be possible for us to observe an inverted-U-shaped dispersion, which reaches the maximum when $\bar{X}$ approaches the turning point of the option-satisfaction relation.

The above conceptual development leaves a few limitations and operationalization challenges to be addressed through mathematical formalism. First, the dispersion in Model 1 is constant only when the underlying relation is linear. In the next section, we generalize the result from linear to monotonic polynomial functions and prove that the dispersion may be monotonic or U-shaped, but cannot be an inverted-U. While the possibility of a U-shaped dispersion potentially confounds the generalized Model 1 with Model 2, both remain clearly distinguishable from Model 3, which is the only possible model that could entail an inverted-U dispersion. Second, the above conceptualization focuses on experiments with similar distance between assortment sizes. With an analytical correction developed in the next section, we expand the pattern to involve all existing studies with arbitrary distance between assortment sizes.

Finally, we would like to note again that the two moderation mechanisms are likely to occur in tandem in practice (Haans et al. 2016). This is especially true for the choice-overload effect, given the likely existence of unknown moderators (McShane and Böckenholt 2018) and the potential for different moderators to function through different mechanisms.¹² To this end, it is important to note that, while the existing empirical evidence might not afford us the resolution required to pinpoint the moderation mechanism for each moderator separately, if we observe an inverted-U pattern of dispersion from the collection of *all* existing studies, two findings abound: First, the underlying option-satisfaction relation is likely an inverted-U rather than a monotonic function as in Model 1, because the latter could not have produced an inverted-U pattern of dispersion. This provides strong evidence for the presence of the choice-overload effect. Second, the moderation of the option-satisfaction relation must involve a substantial shift of the turning point because, were the moderation limited to steepening/flattening the curve like in Model 2, the pattern of dispersion would have been a U-shape rather than an inverted-U-shape.

Mathematical Formalism

We detail the mathematical formulations of the three models introduced in the conceptual development, before proving the identifiability of the model-specific pattern. We discuss the operationalization of the pattern at the end of the section, addressing the challenges brought by the limited number of primary studies in the existing literature.

Mathematical Formulation for The Three Models

Recall from the conceptual development that the first model treats the option-satisfaction relation as monotonic in studying its moderation. In contrast, the other two models treat the relation as an inverted-U. This divergence in treatment necessitates different mathematical formulations, which we detail respectively as follows.

Model 1 (Mixture of Monotonic Relations)

Consider a consumer faced with the decision to choose from an assortment of X items ($X \geq 1$). With the monotonic-

mixture model, we express the consumer's level of satisfaction from making the decision as

$$\mathcal{U}(X) = \beta_0 + \beta_1 X^\alpha + \beta_2 X^\alpha Z + \beta_3 Z, \quad (1)$$

where Z is the moderator variable.¹³ Regardless of the model parameters, $\mathcal{U}(X)$ is always monotonic with X . But its shape depends on α and $\beta_1 + \beta_2 Z$ in tandem. For example, it is linearly increasing when $\alpha = 1$ and $\beta_1 + \beta_2 Z > 0$, and linearly decreasing when $\alpha = 1$ and $\beta_1 + \beta_2 Z < 0$. When $\alpha < 1$ and $\alpha(\beta_1 + \beta_2 Z) > 0$, $\mathcal{U}(X)$ represents diminishing returns as $\mathcal{U}(X)$ is increasing with X yet its first-order derivative $\partial \mathcal{U}(X) / \partial X$ is decreasing with X . In contrast, when $\alpha > 1$ and $\beta_1 + \beta_2 Z < 0$, $\mathcal{U}(X)$ represents accelerating losses. The shape of $\mathcal{U}(X)$ is moderated by Z . For example, in the linear case where $\alpha = 1$, changing Z from 0 to 1 can flip $\mathcal{U}(X)$ from increasing to decreasing if $\beta_1 > 0$ and $\beta_2 < -\beta_1$. When $\beta_2 = 0$, changing Z does not affect the shape of $\mathcal{U}(X)$, but moves it up and down instead.

While this mathematical model is considerably more general than the linear mechanisms depicted in Figure 3(a), it is still limited in assuming that Z does not change the exponent α . This assumption is likely inconsequential in the specific case of choice overload, given the aforementioned dominance of the linear model (i.e., $\alpha = 1$) in the literature. Nonetheless, assuming a constant α could threaten the generalizability of our results. Thus, while we adopt Equation (1) in the mathematical formalism section, we also conducted extensive simulations (to be discussed later in the paper) while varying α , in order to examine the robustness of our results.

Models 2 and 3 (Inverted-U Conceptualizations)

In this case, consumer satisfaction is conceptualized as combining two countervailing forces, one increasing with the assortment size and the other decreasing. Since either force can be captured using the model in Equation (1), we naturally express $\mathcal{U}(X)$ as the algebraic sum of the two, i.e., $\mathcal{U}(X) = \mathcal{U}_0(X) + \mathcal{U}_1(X)$ where

$$\mathcal{U}_0(X) = \beta_0 + \beta_1 X^\alpha + \beta_2 X^\alpha Z + \beta_3 Z, \quad (2)$$

$$\mathcal{U}_1(X) = \beta_4 + \beta_5 X^\gamma + \beta_6 X^\gamma Z + \beta_7 Z. \quad (3)$$

In order for $\mathcal{U}(X)$ to match the inverted-U conceptualization, the following four inequalities must all be satisfied:

¹² As a case in point, see the earlier discussions of the moderating effects of decision goal and consumer expertise.

¹³ Note we can directly extend Z to be a vector of multiple moderator variables. Correspondingly, β_2 and β_3 will become vectors of the coefficients of these variables

$$\alpha(\beta_1 + \beta_2 Z) > 0 \quad (4)$$

$$\gamma(\beta_5 + \beta_6 Z) < 0 \quad (5)$$

$$\alpha(\beta_1 + \beta_2 Z) + \gamma(\beta_5 + \beta_6 Z) > 0 \quad (6)$$

$$\gamma > \alpha \quad (7)$$

Specifically, (4) ensures $\mathcal{U}_0(X)$ is increasing with X ; (5) ensures $\mathcal{U}_1(X)$ is decreasing with X ; (6) ensures $\mathcal{U}(X)$ is increasing when $X = 1$; and (7) ensures $\mathcal{U}(X)$ is decreasing when $X \rightarrow \infty$.

Compared with the common way to model a curvilinear relationship between X and Y , which is to include in the regression equation the first-order X , the quadratic form of it, and their respective interactions with the moderator Z (Aiken et al. 1991; Lind and Mehlum 2010),

$$Y = b_0 + b_1 X + b_2 X^2 + b_3 XZ + b_4 X^2 Z + b_5 Z, \quad (8)$$

our model is considerably more general, as Equation 8 is a special case of our model when $\alpha = 1$, $\gamma = 2$, $\beta_0 + \beta_4 = b_0$, $\beta_1 = b_1$, $\beta_2 = b_2$, $\beta_3 + \beta_7 = b_3$, $\beta_5 = b_5$, $\beta_6 = b_6$, and $\beta_8 = b_4$. This added generalizability is essential for accommodating the nuanced theorizing of choice overload, e.g., the benefits derived from a larger assortment tend to follow a function of diminishing returns (Grant and Schwartz 2011). The traditional model, on the other hand, can only be interpreted as the algebraic sum of a linear relation and an accelerating loss function (Haans et al. 2016).

Taking the first derivative of $\mathcal{U}(X)$ with respect to X , we obtain the curvature of $\mathcal{U}(X)$:

$$\frac{\partial \mathcal{U}(X)}{\partial X} = \frac{\partial \mathcal{U}_0(X)}{\partial X} + \frac{\partial \mathcal{U}_1(X)}{\partial X} = (\beta_1 + \beta_2 Z)\alpha X^{\alpha-1} + (\beta_5 + \beta_6 Z)\gamma X^{\gamma-1}. \quad (9)$$

Setting it to zero, we derive the turning point

$$X^* = \left(-\frac{(\beta_5 + \beta_6 Z)\gamma}{(\beta_1 + \beta_2 Z)\alpha} \right)^{\frac{1}{\alpha-\gamma}}. \quad (10)$$

The two different types of moderation can be explicated by examining how different values of α , β_1 , β_2 , β_5 , β_6 , and γ yield distinct roles of the moderator Z in Equations (9) and (10). For example, when $\beta_6/\beta_2 = \beta_5/\beta_1$, changing Z has zero effect on the turning point X^* . Yet the curvature of $\mathcal{U}(X)$ could change considerably. For example, when $\beta_2 > 0$ and $\beta_6 < 0$, increasing Z from 0 to 1 steepens the curve, while decreasing Z flattens it. This reflects the first type of moderation (i.e., Model 2), where the moderating effect changes the curvature but not the turning point.

In contrast, Z could also shift X^* considerably without occasioning a large change of the curvature of $\mathcal{U}(X)$. For example, consider the case where α and γ are relatively close, e.g., $\alpha = 0.9$ and $\gamma = 1.1$. When $\beta_2 = -\beta_1/10$ and $\beta_6 = \beta_5/10$, increasing Z from 0 to 1 barely moves the curvature, whereas X^* is increased by a ratio of $(1.1/0.9)^5 = 2.73$, pushing the turning point far to the right. This reflects the second type of moderation (i.e., Model 3), where the moderating effect substantially shifts the turning point left or right.

Finally, we note that, while Equations (2) and (3) derive from the conceptual understanding of consumer satisfaction (i.e., the combination of two countervailing forces, Grant and Schwartz 2011), it is not the only analytical formulation that reflects an inverted-U relationship. Another commonly used analytical formulation for inverted U is the spline formulation (i.e., piecewise linear, Simonsohn 2018), which concatenates two linear functions, one increasing and one decreasing, together at the turning point to form an inverted-U (more precisely, an inverted-“V”) shape. Even though this formulation does not align with the conceptual underpinning of choice overload, we will demonstrate in the next subsection that the analytical results derived in the paper readily generalize to the spline formulation, testifying to the robustness of our results.

Model-Specific Patterns in the Dispersion of Reported Effect Sizes

Reported Effect Sizes and Their Dispersion

We start by linking $\mathcal{U}(X)$ to the effect sizes reported in the literature. In most existing studies, the subjects were randomly partitioned into two groups treated with two assortment sizes X_1 and X_2 ($X_1 < X_2$), respectively. The effect size was then computed based on the difference of an outcome variable (e.g., satisfaction) between the two groups. For example, Cohen's d is defined as the mean difference divided by their pooled standard deviation. The pooled standard deviation is commonly assumed to be constant to the independent variable (i.e., assortment size X) in meta-analysis, as a violation of the assumption suggests the existence of a treatment-by-subject interaction (Hunter and Schmidt 2004, p. 283) that has never been reported in the choice-overload literature. We also empirically tested this assumption and found no supporting evidence for a significant correlation between the reported standard deviation and the independent variable.¹⁴

¹⁴ Specifically, we collected all the group-wise standard deviations reported in the literature, resulting in a collection of 120 standard deviations reported by 60 primary studies. Regressing the reported standard deviation over the independent

variable (i.e., assortment size) failed to reject the null hypothesis of the regression coefficient being zero ($p = .082$). Similarly, the Breusch-Pagan

As such, for a given $\mathcal{U}(X)$, the reported effect size d is proportional to the difference of $\mathcal{U}(X)$ between the two assortment sizes X_1 and X_2 :

$$d(X_1, X_2) \propto (\mathcal{U}(X_1) - \mathcal{U}(X_2)). \quad (11)$$

Note that Equation (11) readily applies to other effect-size measures. For example, when the outcome variable is binary and measured on the probability scale $p \in [0,1]$, the commonly used effect-size measure, the log odds ratio, is usually modeled as the difference between $\mathcal{U}(X) = \log(p/(1-p))$ for the two groups and (approximately) transformed to Cohen's d with a linear multiplicative factor of $\sqrt{3}/\pi$ (Sánchez-Meca et al. 2003).

To assess the dispersion of reported effect sizes, it is essential to examine how moderation—i.e., the study-level variations captured by the moderator variable Z —affects $d(X_1, X_2)$. A natural dispersion indicator is the first-order derivative of $d(X_1, X_2)$ with respect to Z , with one caveat. The dispersion imputed from the existing findings can only be nonnegative,¹⁵ yet $\partial d(X_1, X_2)/\partial Z$ can be on either side of zero. A simple rectification is to use the absolute value of the derivative as the dispersion indicator:

$$\Delta(X, \epsilon) = \left| \frac{\partial d(X - \epsilon, X + \epsilon)}{\partial Z} \right| \propto \left| \frac{\partial \mathcal{U}(X - \epsilon)}{\partial Z} - \frac{\partial \mathcal{U}(X + \epsilon)}{\partial Z} \right| \quad (12)$$

Without loss of generality, we replaced $d(X_1, X_2)$ with $d(X - \epsilon, X + \epsilon)$ in Equation (12), in order to highlight the change of Δ with $(X_1 + X_2)/2$. Figure 4 illustrates three examples for how $\Delta(X, \epsilon)$ exhibits unique patterns for different underlying models. In the passages that follow, we develop more generic analytical results for the pattern of the dispersion indicator over each of the three models, respectively.

Model 1

With the monotonic-mixture model, by taking Equation (1) into Equation (12), we obtain

$$\Delta(X, \epsilon) \propto |\beta_2((X - \epsilon)^\alpha - (X + \epsilon)^\alpha)|. \quad (13)$$

Consider the expression inside the absolute value operator in Equation (13), $\vec{\Delta}(X, \epsilon) = \beta_2((X - \epsilon)^\alpha - (X + \epsilon)^\alpha)$. Clearly, holding ϵ (i.e., the assortment-size difference between the two groups) constant, $\vec{\Delta}$ is always monotonic to X , with the direction determined by the combination of α and β_2 . For example, in the linear case where $\alpha = 1$, $\vec{\Delta}$ remains constant no matter how we

shift X . When $\beta_2 < 0$, $\vec{\Delta}$ increases with X_i if $\alpha > 1$ or $\alpha < 0$, and decreases if $\alpha \in [0,1]$. Since $\vec{\Delta}$ is monotonic and Δ is the absolute value of $\vec{\Delta}$, Δ may be monotonic or U-shaped, but cannot be an inverted-U.¹⁶ This is summarized by the following theorem, the proof of which is available in the appendix.

Theorem 1: When $\mathcal{U}(X) = \beta_0 + \beta_1 X^\alpha + \beta_2 X^\alpha Z + \beta_3 Z$, the dispersion indicator $\Delta(X, \epsilon)$ has no local maximum with respect to X .

Model 2

Next, consider the inverted-U model when the moderating effect only steepens/flattens the curve without changing the turning point. A key observation here is $\lim_{\epsilon \rightarrow 0} d(X^* - \epsilon, X^* + \epsilon) = 0$, where X^* is the turning point of the inverted-U. This means that, when ϵ is sufficiently small, $d(X^* - \epsilon, X^* + \epsilon) \approx 0$ regardless of Z . When $d(X^* - \epsilon, X^* + \epsilon) = 0$ for all Z , there must be $\Delta(X^*, \epsilon) = |\partial d(X^* - \epsilon, X^* + \epsilon)/\partial Z| = 0$, its minimum possible value. Formally,

$$\lim_{\epsilon \rightarrow 0} \frac{\Delta(X^*, \epsilon)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \left| \frac{\partial d(X^* - \epsilon, X^* + \epsilon)}{\epsilon \partial Z} \right| = 0. \quad (14)$$

In other words, the dispersion indicator $\Delta(X, \epsilon)$ is *minimized* at the turning point $X = X^*$, consistent with our earlier conceptual discussions. The following theorem formalizes this notion, with the proof included in the Appendix.

Theorem 2: When ϵ is sufficiently small, the dispersion indicator $\Delta(X, \epsilon)$ reaches its minimum at X^* ($X^* > 1$) if for all Z ,

$$\frac{\partial \mathcal{U}(X^*)}{\partial X} = 0, \quad (15)$$

which also implies that X^* is always the turning point for $\mathcal{U}(X)$ regardless of the moderator Z .

Model 3

Theorems 1 and 2 are already sufficient for establishing the model specificity of an inverted-U shaped dispersion indicator $\Delta(X, \epsilon)$ with respect to X . Since neither Model 1 nor Model 2 was able to produce an inverted-U-shaped $\Delta(X, \epsilon)$, if we observe such an inverted-U dispersion, the only possible explanation is Model 3, i.e., the option-satisfaction relation being an inverted-U, whose turning point is shifted significantly through moderation.

heteroskedasticity test (Breusch and Pagan 1979), which regresses the variance over the independent variable, also fails to reject the null hypothesis of homoscedasticity ($p = .073$). Overall, our empirical examination identified no supporting evidence for heteroskedasticity or treatment-by-subject interaction.

¹⁵ The reason is that we do not have a perfect understanding of all moderator variables for all existing studies.

¹⁶ Mathematically, if f is monotonic, then for all x in the domain of f , there must be $|f(x)| \leq \max(|f(x - \epsilon)|, |f(x + \epsilon)|)$.

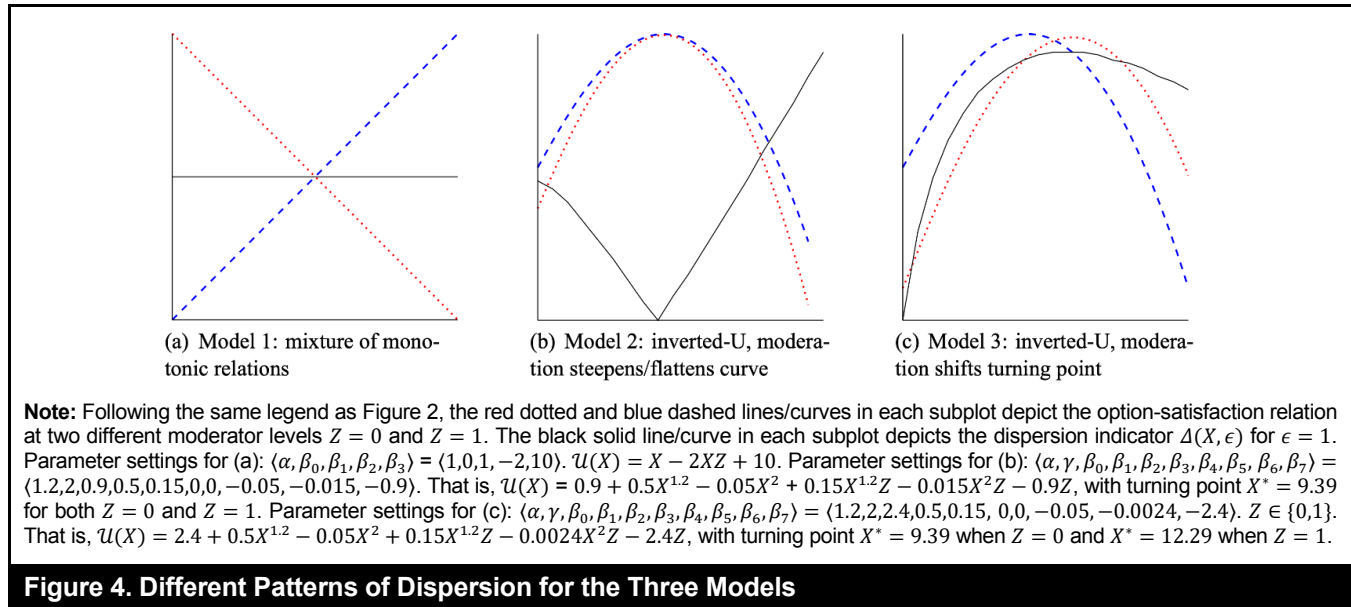


Figure 4. Different Patterns of Dispersion for the Three Models

Given the sufficiency of Theorems 1 and 2 on proving the uniqueness of the pattern, our remaining goal is to demonstrate that an inverted-U dispersion is indeed a common occurrence¹⁷ for Model 3. In other words, the shape of the dispersion indicator $\Delta(X, \epsilon)$ has a sufficiently high statistical power to identify the presence of Model 3. Since the dynamics among the model parameters in Model 3 are too complex to allow for an analytical examination of the dispersion indicator, we defer the power analysis to the simulation studies later in the paper, and present here a simple example demonstrating the emergence of an inverted-U dispersion from the left/right shift of a famous inverted-U function, (a generalized version of) the Runge function¹⁸, $\mathcal{U}(X) = 1/(\xi + (X - c_1Z - c_2)^2)$, where $\xi > 0$ marks an upper bound of $1/\xi$ for $\mathcal{U}(X)$, and the offset $c_1Z + c_2$ indicates that increasing Z shifts $\mathcal{U}(X)$ (and the turning point X^*) to the right by an offset of c_1Z . In this case, consider the change of $d(X - c_1, X)$ when Z varies from 0 to 1. We have

$$d_{Z=0}(X - c_1, X) - d_{Z=1}(X - c_1, X) = \mathcal{U}_{Z=0}(X) - \mathcal{U}_{Z=0}(X - c_1) - \mathcal{U}_{Z=1}(X) + \mathcal{U}_{Z=1}(X - c_1) \quad (16)$$

$$= \mathcal{U}_{Z=1}(X + c_1) - 2\mathcal{U}_{Z=1}(X) + \mathcal{U}_{Z=1}(X - c_1) \quad (17)$$

$$\approx c_1^2 \cdot \frac{\partial^2 \mathcal{U}_{Z=1}(X)}{\partial X^2} \quad (18)$$

The reduction from Equation (17) to (18) follows directly from the notion of second symmetric derivative (Zygmund 2002) and is asymptotic when $c_1 \rightarrow 0$. Taking advantage of the simple analytical form of the second derivative of the (generalized) Runge function when ξ is small, i.e.,

$$\lim_{\xi \rightarrow 0} \frac{\partial^2 \mathcal{U}_{Z=1}(X)}{\partial X^2} = \frac{6}{(X - c_1 - c_2)^4}, \quad (19)$$

we know from Equation (18) that the dispersion indicator must also follow an inverted-U shape, with the exact same turning point (i.e., $c_1 + c_2$) as $\mathcal{U}_{Z=1}(X)$.

Generalizability to the Spline Formulation

Finally, we demonstrate that the unique analytical pattern derived for the polynomial formulations also holds for the spline formulation discussed before, which concatenates a linearly increasing function with a linearly decreasing one at the turning point. Like how Theorems 1 and 2 show that the dispersion indicator $\Delta(X, \epsilon)$ cannot form a local maximum under the first two models, the following theorem shows that, with the spline formulation, the dispersion indicator $\Delta(X, \epsilon)$ also cannot form a local maximum if moderation is limited to the steepening/flattening effect but not a shift of

¹⁷ Note that an inverted-U dispersion is a sufficient but not necessary condition for Model 3. That is, not all inverted-U functions produce an inverted-U dispersion when being shifted left/right. An example is a function shaped like “_ □ _” (e.g., $\mathcal{U}(X) = 1$ if $X \in [5, 10]$ and 0 otherwise). Clearly, the dispersion indicator in this case is more complex

than an inverted-U, as it first increases, then decreases, before increasing and then decreasing again.

¹⁸ The Runge function is $f(x) = 1/(1 + 25x^2)$ (Cheney and Light 2009). We generalized the factor of 25 to $1/\xi$, and introduced an offset $c_1Z + c_2$ to capture the shifting effect of moderation.

the turning point. In other words, the analytical pattern we discovered readily generalizes to the spline formulation.

Theorem 3: When $\mathcal{U}(X)$ follows the spline model

$$\mathcal{U}(X) = \begin{cases} \beta_1 X + \beta_2 XZ, & \text{if } X \leq X^* \\ \beta_3 X + \beta_4 XZ + (\beta_1 - \beta_3)X^* + (\beta_2 - \beta_4)X^*Z, & \text{if } X > X^* \end{cases} \quad (20)$$

where $\beta_1 + \beta_2 Z > 0$ and $\beta_3 + \beta_4 Z < 0$ for all Z , the dispersion indicator $\Delta(X, \epsilon)$ is always monotonic.

Note that the theorem considers a spline model that concatenates line segments of the form $\mathcal{U}(X) = \beta_i X + \beta_j XZ$. Although this form is a simplification of Equation (2), it does not affect the generality of the theorem because a change of $\mathcal{U}(X)$ constant to X (e.g., by $\beta_0 + \beta_3 Z$ in Equation 2) has no impact on the observed effect size, given that the two segments must share the same value when $X = X^*$. Similarly, the last two items in the expression of $\mathcal{U}(X)$ when $X > X^*$ are also to ensure that the two segments feature the same value of $\mathcal{U}(X)$ when $X \rightarrow X^*$.

Operationalization for Pattern Recognition

There are two key challenges in operationalizing the dispersion index over the existing empirical evidence for choice overload. The first challenge is the latent nature of the study-level variations captured by Z . While multiple studies in one article may have otherwise identical designs, except for one explicitly marked moderator variable, such common ground is rare across articles. Thus, when analyzing empirical results from different articles, it is difficult to determine which (pairs of) studies feature more similar study-level designs than others. In other words, we cannot explicate the value of Z in practice, let alone computing $\Delta(X, \epsilon) = |\partial d(X - \epsilon, X + \epsilon) / \partial Z|$. Thus, to overcome this challenge, we have to *infer* the degree of effect-size variation with respect to Z without actually knowing the exact values of Z .

Fortunately, there is a simple inference if we have access to a sufficiently large body of empirical evidence. Specifically, suppose that we have a large number of existing studies featuring effect-size pair $X - \epsilon$ and $X + \epsilon$. We do not know the values of Z for any study. Instead, we assume that Z is randomly drawn from a probability distribution with domain $Z \in \Omega$ and density function $p(\cdot)$. A simple inference for $\Delta(X, \epsilon)$ is the standard deviation of the reported effect sizes:

$$\tilde{\Delta}(X, \epsilon) = \sqrt{\int_{\Omega} (d(X - \epsilon, X + \epsilon|Z_0) - \int_{\Omega} d(X - \epsilon, X + \epsilon|Z_1)p(Z_1)dZ_1)^2 p(Z_0)dZ_0} \quad (21)$$

$$= \sqrt{\int_{\Omega} \left(\int_{Z_1}^{Z_0} \frac{\partial d(X - \epsilon, X + \epsilon)}{\partial Z} p(Z_1)dZ_1 \right)^2 p(Z_0)dZ_0} \quad (22)$$

By taking the partial derivative of Equation (22) with respect to X , one can see that $\tilde{\Delta}(X, \epsilon)$ follows the same trend with respect to X as those specified for $\Delta(X, \epsilon)$ in Theorems 1 and 2. For example, since Model 3 is the only possible model for $|\partial d(X - \epsilon, X + \epsilon) / \partial Z|$ to reach a local maximum with respect to X for any Z , if we observe such a local maximum for $\tilde{\Delta}(X, \epsilon)$, we can still infer that Model 3 underlies the observed effects. Figure 5(a)-(c) demonstrates this operationalization for the three models depicted in Figure 4. As can be seen from the figure, the model-specific patterns are still distinctly identifiable from $\tilde{\Delta}$. Specifically, we can observe a constant dispersion for Model 1, a U-shaped change of dispersion (with respect to $(X_1 + X_2)/2$) for Model 2, and an inverted-U shaped dispersion for Model 3.

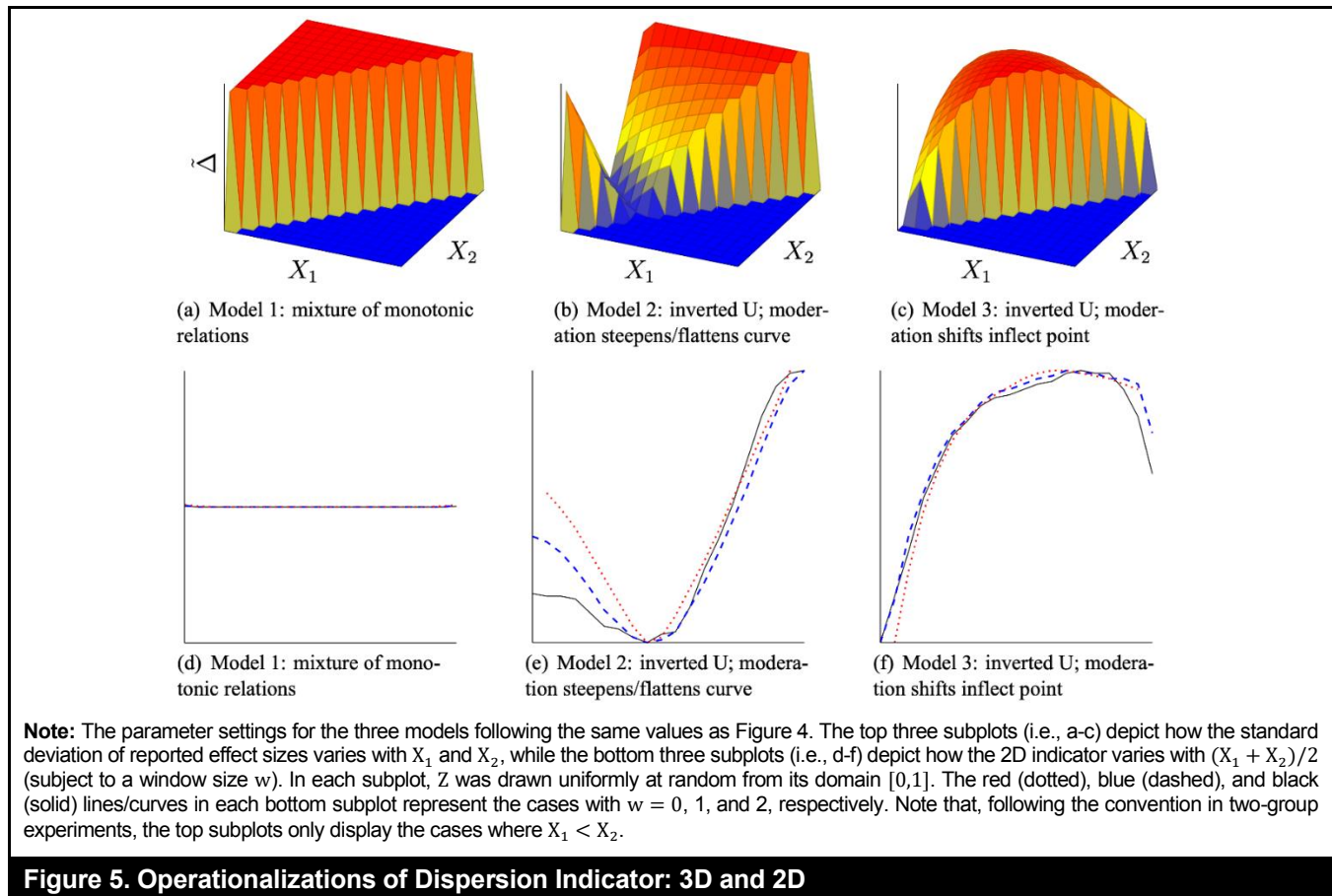
Unfortunately, this operationalization is blunted by the second key challenge in practice: The existing studies did not cover enough assortment-size pairs to enable the plotting of a 3D chart like those in Figure 5(a)-(c). When there are not enough data to properly support a 3D plot, a natural idea is to design a dimension-reduction projection of the 3D plot to a 2D one. In this case, instead of measuring the standard deviation for each pair of X_1, X_2 separately to estimate $\tilde{\Delta}((X_1 + X_2)/2, (X_2 - X_1)/2)$, we consider the measurement of standard deviation for all assortment-size pairs that share the same (or similar) $(X_1 + X_2)/2$ regardless of their difference (i.e., $(X_2 - X_1)/2$). In other words, we are projecting the 3D plot of $\tilde{\Delta}(X, \epsilon)$ to a 2D plot of $\tilde{\Delta}(X)$, based on the rationale that, if $\tilde{\Delta}(X, \epsilon)$ follows the same pattern regardless of ϵ , then we should observe the same pattern when grouping the studies with the same X (i.e., $(X_1 + X_2)/2$) but different ϵ (i.e., $(X_2 - X_1)/2$).

There is one remaining caveat in this projection, as grouping studies with different ϵ could entail the overweighting of studies with a large difference between their assortment-size choices. For example, consider how $\Delta(X, \epsilon)$ varies with ϵ for Models 2 and 3:

$$\begin{aligned} \Delta(X, \epsilon) &= \left| \frac{\partial d(X - \epsilon, X + \epsilon)}{\partial Z} \right| \\ &\propto |\beta_2((X - \epsilon)^\alpha - (X + \epsilon)^\alpha) + \beta_6((X - \epsilon)^\gamma - (X + \epsilon)^\gamma)| \end{aligned} \quad (23)$$

$$\approx |2\epsilon(\beta_2\alpha X^{\alpha-1} + \beta_6\gamma X^{\gamma-1})|. \quad (24)$$

Since $\Delta(X, \epsilon)$ is approximately proportional to ϵ (when ϵ is small), directly measuring the standard deviation of effect sizes with different ϵ could bear more weight on those studies with larger ϵ . To offset this bias, we introduce a multiplicative correction factor of $1/(X_2 - X_1)$ to the calculation.



Specifically, given a set of K primary studies with reported effect sizes being $d_1, \dots, d_K$ and the assortment-size pairs being $\langle X_{11}, X_{12} \rangle, \dots, \langle X_{K1}, X_{K2} \rangle$, respectively, our final operationalization of the dispersion indicator is:

$$\tilde{\Delta}(X) = SD \left(\left\{ \frac{d_i}{X_{i2} - X_{i1}} \mid X_{i1} + X_{i2} \in [2X - w, 2X + w] \right\} \right), \quad (25)$$

where SD represents the standard deviation, and w is a window size used to further address the cases where no existing study have $(X_{i1} + X_{i2})/2$ exactly equal to X . Figure 5(d)-(f) depicts the change of $\tilde{\Delta}(X)$ with respect to X for the three models when the window size w varies from 0 to 2. As can be seen from the figure, the model-specific patterns of the dispersion indicator remain clearly identifiable from our final operationalization. We will further elaborate on the setting of w and test the statistical power of $\tilde{\Delta}(X)$ -based model identification in the next section.

Results

In this section, we discuss the results of using our dispersion indicator to identify the latent mechanism of the underlying relation. First, we describe simulation studies that test the

accuracy of our method over an extensive set of linear, curvilinear-yet-monotonic, and curvilinear inverted-U relationships with both types of moderation. Then, we discuss the results of applying our method to the empirical evidence of choice overload reported in the literature. Finally, we present another case study of our method on the existing empirical evidence of choice overload in the specific context of online browsing, an important IS/IT artifact.

Design of Simulation Studies

Overview of Simulation Design

We conducted a total of four simulation studies. The first three tested the accuracy of our method with (1) monotonic relations, (2) inverted-U relations with moderation steepening/flattening the curve, and (3) inverted-U relations with moderation shifting the turning point, respectively, while the last study examined the robustness of our method in the presence of random noises in effect sizes. In each study, we computed the dispersion indicator as operationalized by $\tilde{\Delta}(X)$ in Equation (25), and tested whether $\tilde{\Delta}(X)$ is an inverted-U with respect to X . While the design of our method is agnostic to the specific method used

for testing the inverted-U, for the purpose of the implementation, we adopted the three-step procedure developed by Lind and Mehlum (2010), which is a popular choice for inverted-U testing in the IS literature (e.g., Scherer et al. 2015; Singh et al.; 2011). Specifically, the Lind-Mehlum test involves fitting $\tilde{A}(X) = b_0 + b_1X + b_2X^2$ with least squares before using the fitted parameters b_0, b_1, b_2 to determine whether $\tilde{A}(X)$ forms an inverted-U¹⁹. If the test indicates an inverted-U shape for $\tilde{A}(X)$, we further estimate its turning point as $-b_1/(2b_2)$ (Haans et al. 2016). Besides the Lind-Mehlum method, we also attempted other methods for inverted-U testing (e.g., the planned contrast tests in Kamis et al. 2008 and Xu et al. 2014; the two-line test in Simonsohn 2018) but did not find qualitative differences in the results.

Since Theorems 1 and 2 indicate that an inverted-U $\tilde{A}(X)$ should only emerge when the underlying relation is an inverted-U moderated through shifting the turning point left/right, we should ideally only observe an inverted-U dispersion indicator in the third simulation study. Thus, the accuracy metric is the Type I error rate for the first two studies, and the statistical power for the third study. We implemented our method and produced all simulation results in MATLAB. In the passages that follow, we describe the simulation design in each study, respectively.

Study 1

For the first study, we created two levels for the moderating variable: $Z = 0$ and $Z = 1$. We set $\mathcal{U}(X) = X^{\alpha_0}$ when $Z = 0$ and $\mathcal{U}(X) = -X^{\alpha_1}$ when $Z = 1$. For either parameter (i.e., α_0 or α_1), we created five levels: {0.5, 0.8, 1.0, 1.5, 1.8} to capture a wide variety of monotonic functions, from linear (i.e., $\alpha_0 = 1$) to diminishing returns (e.g., $\alpha_0 = 0.5$) to accelerating losses (e.g., $\alpha_1 = 1.8$). Recall from earlier discussions that this wide coverage of α_0 and α_1 was designed to cover the cases beyond the reach of our analytical examination, i.e., when the moderator Z varies the exponent of X in $\mathcal{U}(X)$. To examine the accuracy of our method given different numbers of primary studies, we created three different levels: {40, 70, 100}. Since our method involves only one parameter, the window size w , we created two levels for it in the simulations: {1, 3}. Overall, the design of Study 1 consisted of 160 unique conditions or a $5 (\alpha_0) \times 5 (\alpha_1) \times 3$ (number of primary studies) $\times 2$ (window size w) factorial design. Under each condition, we repeated the test 100 times. In each test, we generated the assortment sizes X_1 and X_2 uniformly at random from [1, 40], before computing the effect size based on $\mathcal{U}(X_2) - \mathcal{U}(X_1)$.

Study 2

For the second study, we again set two levels for the moderating variable: $Z = 0$ and $Z = 1$. Note that the goal of Study 2 is to test inverted-U relations with moderation steepening/flattening the curve but not shifting the turning point. Thus, we set the turning point to be the same value of $X^* = 20$ for both moderator levels and varied the steepness of the inverted U between the two moderator levels. Like in the first simulation study, we adjusted the steepness by varying the exponent of X in $\mathcal{U}(X)$, as such variation was beyond the reach of our analytical examination. Unlike in the first study, we can now vary the exponent differently on different sides of the turning point, in order to examine how the steepness change on one or both sides affect the results. Specifically, we consider a piecewise polynomial inverted U that has $\mathcal{U}(X)$ following X^{α_0} (i.e., increasing) to the left of the turning point and $-X^{\alpha_1}$ (i.e., decreasing) to the right of the turning point, with both segments normalized to the same value at the turning point $X^* = 20$. In numerical analysis (Cheney and Light 2009), such a piecewise polynomial function is well known to be a more generalized approximation of inverted-U functions than the polynomial approximations discussed in the mathematical formalism section²⁰. When $Z = 0$, we set $\alpha_0 = \alpha_1 = 1.5$. When $Z = 1$, we created four levels each for α_0 and α_1 : {0.5, 0.8, 1, 1.8} (1.5 is excluded as it represents no change from the case where $Z = 0$). Clearly, the parameter combinations captured a wide variety of moderating effects, including the flattening of the curve on both sides (e.g., $\alpha_0 = \alpha_1 = 0.5$ for $Z = 1$), steepening on both sides (e.g., $\alpha_0 = \alpha_1 = 1.8$ for $Z = 1$), or flattening on one side but steepening on the other (e.g., $\alpha_0 = 0.5, \alpha_1 = 1.8$ for $Z = 1$). Again, we varied the number of primary studies in {40, 70, 100}, and the window size w in {1, 3}. Overall, the second simulation study consisted of 96 unique conditions or a $4 (\alpha_0 \text{ when } Z = 1) \times 4 (\alpha_1 \text{ when } Z = 1) \times 3$ (number of primary studies) $\times 2$ (window size w) factorial design.

Study 3

For the third study, we followed the same design of an inverted-U function as in the second simulation study, i.e., by having $\mathcal{U}(X)$ follow X^{α_0} (i.e., increasing) to the left of the turning point and $-X^{\alpha_1}$ (i.e., decreasing) to the right of the turning point, with both segments normalized to the same value at the turning point X^* . We again created five levels for both α_0 and α_1 : {0.5, 0.8, 1.0, 1.5, 1.8}. Unlike in Study 2, where we varied α_0 and α_1 across the moderator levels, in Study 3, we varied the turning point X^* but kept the values of α_0 (or α_1) the same

¹⁹ We used a significance level of 0.05 when applying Lind and Mehlum (2010) procedure.

²⁰ For example, piecewise polynomials address the Runge's phenomenon, with which even a high-degree polynomial cannot properly approximate

certain inverted-U functions under equispaced interpolation (Dahlquist and Bjork 1974).

between the two moderator levels. Specifically, we set the turning point for $Z = 1$ at $X^* = 24$, and created two levels for the turning point for $Z = 0$: $\{16, 20\}$. Like in the previous two studies, we varied the number of primary studies in $\{40, 70, 100\}$, and the window size w in $\{1, 3\}$. Overall, the third simulation study consisted of 300 unique conditions or a $5 (\alpha_0) \times 5 (\alpha_1) \times 2 (X^* \text{ when } Z = 0) \times 3 (\text{number of primary studies}) \times 2 (\text{window size } w)$ factorial design.

Study 4

The last study was designed to test whether the presence of random noise in the reported effect sizes could substantially degrade the efficacy of our method. For this purpose, we focused on the *worst-case* simulation conditions for our method in the first three studies—i.e., the pairs of (α_0, α_1) under which our method produced the highest Type I error rate or the lowest statistical power. For each of these conditions, we tested the accuracy of our method when random noises are inserted into the input effect sizes. Specifically, for each condition, we added to each input effect size an independent and identically distributed (i.i.d.) Gaussian random noise $\mathcal{N}(0, (\epsilon M)^2)$, where M is the mean absolute effect size reported under the condition, and ϵ is a simulation factor controlling the magnitude of the inserted random noise. We created three levels for ϵ : 0.01, 0.05, and 0.1. Overall, the fourth simulation study consisted of 54 unique conditions or a $1 (\text{worst-case } (\alpha_0, \alpha_1) \text{ for Type I error}) \times 3 (\text{number of primary studies}) \times 2 (\text{window size } w) \times 3 (\text{noise level } \epsilon) + 1 (\text{worst-case } (\alpha_0, \alpha_1) \text{ for Type II error}) \times 2 (X^* \text{ when } Z = 0) \times 3 (\text{number of primary studies}) \times 2 (\text{window size } w) \times 3 (\text{noise level } \epsilon)$ factorial design.

Simulation Results

For the first simulation study, remarkably, there was not a single Type I error in any of the 16,000 runs. That is, the Type I error rate stays at zero for *all* 160 simulation conditions. This demonstrates the robustness of our method on ruling out the underlying relation being monotonic once observing an inverted-U dispersion indicator.

Table 1 depicts the Type I error rates in Study 2. Since swapping α_0 and α_1 does not change the result, we only listed in the table the cases where $\alpha_0 \leq \alpha_1$. The Type I error rates across all simulation conditions have mean $M = 0.0085$ ($SD = 0.0274$). Indeed, there is not a single error in 83.33% of the simulation conditions (80 out of 96). Table 1 shows that the Type I errors mostly occurred when α_0 and α_1 differ drastically (e.g., when

$\alpha_0 = 0.5$ and $\alpha_1 = 1.8$). A key reason is that the difference makes $\mathcal{U}(X)$ no longer differentiable at the turning point. As a result, when the assortment sizes are close to the turning point, the observed effect size could deviate from zero by a greater degree than in the case of a differentiable $\mathcal{U}(X)$ (e.g., as in Theorem 2). Nonetheless, even in these worst-case scenarios, the Type I error rate never exceeded 0.06 for the case of 100 primary studies, demonstrating the robustness of our method.

Table 2 depicts the statistical power of our method in the third simulation study. The statistical power across all simulation conditions had a mean of $M = 0.9387$ ($SD = 0.1415$). When the window size was $w = 3$, the *lowest* statistical power in all conditions was 0.92 when there were 100 primary studies. Even when the number of primary studies was limited to just 40, our method achieved a statistical power of at least 0.98 in the majority of simulation conditions (31 out of 50, 62.0%). These results demonstrate the power of our method for identifying the inverted-U nature of the underlying relation and its moderation mechanism (i.e., a left/right shift). Further, the turning points of the inverted-U dispersion indicator $\hat{D}(X)$ match closely with the true turning point (i.e., $X^* = 24$) of the underlying relation when $Z = 1$, as predicted in our earlier discussions (e.g., Equation 19). For example, the estimated turning point of the dispersion indicator had a mean of $M = 23.93$ ($SD = 2.63$) when the true turning points are $\{20, 24\}$ for the two moderator levels, and a mean of $M = 21.55$ ($SD = 2.17$) when the true turning points were $\{16, 24\}$.

Table 3 depicts the impacts of random noise on the accuracy of our method. Recall from Tables 1 and 2 that the worst-case settings of α_0, α_1 are (0.5, 1.8) for Type I error rates (Table 1) and (0.5, 1) for statistical power (Table 2), hence their inclusion in this fourth simulation study. As can be seen from the left half of the table, the Type I error rates incurred by our method in the noise-ridden cases are on par with or even lower than the noiseless case (i.e., $\epsilon = 0$), especially when the number of primary studies is large (e.g., $K = 100$, like in the choice-overload case). While this is ostensibly surprising, the reason is the same as the above-discussed reason why $\alpha_0 = 0.5$ and $\alpha_1 = 1.8$ became the worst-case setting in Study 2—i.e., the first-order discontinuity²¹ of $\mathcal{U}(X)$ at the turning point X^* . Specifically, the insertion of random noise reduced the drastic difference between the one-sided limits of $\partial \mathcal{U}(X) / \partial X$ at point X^* from the positive and negative directions (i.e., when $X \rightarrow (X^*)^+$ and $X \rightarrow (X^*)^-$), therefore lessening the discontinuity that likely caused the Type I errors in the noiseless case. A similar observation can be made from the statistical powers reported in the right half of Table 3.

²¹ A function $f(x)$ has a first-order discontinuity at point x_0 if it is continuous but not differentiable at $x = x_0$.

Table 1. Type I Error Rates in Study 2

K	α_0, α_1									
	0.5,0.5	0.5,0.8	0.5,1	0.5,1.8	0.8,0.8	0.8,1	0.8,1.8	1,1	1,1.8	1.8,1.8
$w = 1$										
40	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.00
70	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00
100	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
$w = 3$										
40	0.00	0.00	0.00	0.15	0.00	0.00	0.05	0.00	0.01	0.00
70	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00
100	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00

Note: K = number of primary studies; w = window size; α_0, α_1 = exponent settings in $\mathcal{U}(X)$ when $Z = 1$.

Table 2. Statistical Power in Study 3

K	α_0, α_1														
	0.5,0.5	0.5,0.8	0.5,1	0.5,1.5	0.5,1.8	0.8,0.8	0.8,1	0.8,1.5	0.8,1.8	1,1	1,1.5	1,1.8	1.5,1.5	1.5,1.8	1.8,1.8
$c = 4, w = 1$															
40	0.95	0.97	0.22	0.89	0.85	1.00	0.20	0.98	0.85	0.56	0.96	0.95	0.99	0.95	0.60
70	1.00	1.00	0.35	1.00	0.98	1.00	0.53	1.00	0.99	0.80	1.00	0.99	1.00	1.00	0.72
100	1.00	1.00	0.46	1.00	1.00	1.00	0.76	1.00	0.99	0.90	1.00	1.00	1.00	1.00	0.64
$c = 4, w = 3$															
40	1.00	1.00	0.64	1.00	0.99	1.00	0.64	1.00	1.00	0.82	1.00	0.99	1.00	0.98	0.82
70	1.00	1.00	0.83	1.00	1.00	1.00	0.88	0.99	1.00	0.98	1.00	1.00	1.00	1.00	0.98
100	1.00	1.00	0.92	1.00	1.00	1.00	0.95	1.00	1.00	0.96	1.00	1.00	1.00	1.00	0.99
$c = 8, w = 1$															
40	0.92	1.00	0.39	0.93	0.88	1.00	0.80	0.99	0.97	0.93	1.00	1.00	1.00	1.00	0.71
70	1.00	1.00	0.66	0.99	0.98	1.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.74
100	1.00	1.00	0.91	0.99	0.99	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.70
$c = 8, w = 3$															
40	1.00	1.00	0.83	1.00	0.98	1.00	0.99	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.75
70	1.00	1.00	0.96	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.78
100	1.00	1.00	0.98	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.92

Note: K = number of primary studies. C = difference between turning points of the two moderator levels (i.e., when $c = 4$, $X^* = 20$ when $Z = 0$; when $c = 8$, $X^* = 16$ when $Z = 0$). W = window size. α_0, α_1 = exponent settings in $\mathcal{U}(X)$ for both moderator levels.

Table 3. Impacts of Random Noise in Study 4

	Type I error rate (when $\alpha_0, \alpha_1 = 0.5, 1.8$)						Statistical power (when $\alpha_0, \alpha_1 = 0.5, 1$)					
	w, K						w, K					
	1,40	1,70	1,100	3,40	3,70	3,100	1,40	1,70	1,100	3,40	3,70	3,100
ϵ												
$c = 4$												
0.00	0.11	0.03	0.02	0.15	0.08	0.06	0.22	0.35	0.46	0.64	0.83	0.92
0.01	0.16	0.09	0.05	0.09	0.04	0.02	0.96	1.00	1.00	0.99	1.00	1.00
0.05	0.14	0.10	0.04	0.09	0.02	0.01	0.84	0.91	0.96	0.93	0.94	0.96
0.10	0.12	0.10	0.04	0.09	0.02	0.00	0.53	0.68	0.58	0.68	0.72	0.66
$c = 8$												
0.00							0.39	0.66	0.91	0.83	0.96	0.98
0.01							1.00	1.00	1.00	1.00	1.00	1.00
0.05							0.95	0.99	0.99	0.98	1.00	0.99
0.10							0.77	0.80	0.81	0.82	0.80	0.79

Note: α_0, α_1 = exponent settings in $\mathcal{U}(X)$. ϵ = magnitude of random noise. K = number of primary studies. w = window size. c = difference between turning points of the two moderator levels. The left part of the table reports the Type I error rates for different levels of ϵ under the worst-case conditions in Study 2 (i.e., $\alpha_0 = 0.5, \alpha_1 = 1.8$). The right part of the table reports the statistical power for different levels of ϵ under the worst-case conditions in Study 3 (i.e., $\alpha_0 = 0.5, \alpha_1 = 1$). Note that parameter c only applies to the case of statistical power (i.e., Study 3), hence the four additional rows in the right half.

Once again, our method surprisingly attained a considerably *higher* statistical power in most of the noise-ridden cases than in the noiseless case. The reason here again stems from a discontinuity in the noiseless case. Note that when $\alpha_1 = 1$, any two-group experiment with both assortment sizes falling to the right of the turning point would report the same effect size (after the weighting-based correction in Equation 24), entailing a sharp drop of the dispersion indicator $\tilde{D}(X)$ to zero once X grows past X^* . This jump in discontinuity is fundamentally incompatible with the existing methods for inverted-U testing.²² The resulting anomalies explain the lower statistical power for the simulation conditions with either $\alpha_0 = 1$ or $\alpha_1 = 1$ (or both) in Table 2. Like in the Type I error case, the insertion of noise alleviates this discontinuity and therefore increases the statistical power in many of these settings. Overall, as can be seen from Table 3, the accuracy of our method is robust to the presence of random noise in the reported effect sizes.

Empirical Evidence for Choice Overload

The intense research interests on choice overload have led to multiple recent meta-analyses (Chernev et al. 2015; McShane and Böckenholt 2018; Scheibehenne et al. 2010; Simonsohn et al. 2014), with the most recent two using the same dataset. We believe the best way to study the insights unveiled by our novel analytical method is to reuse the existing dataset to the extent possible (and therefore sharing the same inclusion/exclusion criteria, weighting design, etc., as the two existing meta-analyses) in order to minimize the chance of differences in outcomes stemming from differences in data.

The dataset contains 21 articles in which 99 observed effect sizes were reported. In terms of the study-level designs, the existing meta-analyses, Chernev et al. (2015) and McShane and Böckenholt (2018) synthesized four types of moderator variables—choice set complexity, decision task difficulty, preference uncertainty, and decision goal—and six types of operationalizations for the dependable variable: satisfaction /confidence, regret, choice deferral, switching likelihood, assortment choice, and option selection. In terms of the empirical evidence reported, while most reported effect sizes are in the form of Cohen's d , some studies operationalized the dependent variable with a binary decision outcome, such as whether an individual ultimately decided to make a purchase. In this case, the “mean” of the dependent variable for each assortment size becomes the proportion of individuals who made a purchase. For these cases, we followed the same standard transformations as the existing meta-analysis (Chernev et al. 2015).

Nonetheless, there are a few existing studies that operationalized the dependent variable in a way that is inherently incompatible with our purpose of examining the choice-satisfaction function. This operationalization, namely “assortment choice” (Chernev et al. 2015), requires each individual to make a binary selection between small and large assortments, and then measure the percentage of individuals who chose the small (or large) one. Given a choice-satisfaction function $\mathcal{U}(X)$ and two assortment sizes X_1 and X_2 , while this operationalization captures the percentage of individuals with $\mathcal{U}(X_1) > \mathcal{U}(X_2)$, it reveals no additional information (other than the sign) of $\mathcal{U}(X_1) - \mathcal{U}(X_2)$ (Cohen 1988, p. 147), which is crucial for quantitatively assessing the effect of choice overload.

Out of the 21 articles included in the meta-analyses, three (Chernev 2006; Chernev and Hamilton 2009; Goodman and Malkoc 2012) used this operationalization and therefore had to be excluded from our study. We note that Chernev et al. (2015) excluded the exact same set of studies when examining the mean effect of choice overload, for the same reason as discussed. This exclusion led to two changes in the dataset: First, it removed “assortment choice” as a type of dependent-variable operationalization. Second, a moderator variable *decision goal* was also removed because, as noted by Chernev et al. (2015), nearly all existing studies examining the moderating effect of decision goal used assortment choice as the outcome measure. As only two effect sizes remain for decision goal variable after the exclusion, we could not properly examine decision goal as a distinctive type of moderator variable.

Analytical Results for Choice Overload

Table 4 depicts the results of applying our method over the effect sizes reported in the choice-overload literature. Since a minimum window size of $w = 3$ is required to provide a consecutive coverage²³ of X , we tested our method with $w = 3, 4$, and 5. As can be seen from the table, our dispersion indicator $\tilde{D}(X)$ clearly follows an inverted-U shape regardless of the window size. Specifically, in all three settings, the dispersion reached its three maximum values when X is 10, 11, and 12. Combined with our mathematical analysis earlier in the paper, the results in Table 4 provide strong evidence that the option-satisfaction relation is an inverted U with moderation shifting the curve left and right.

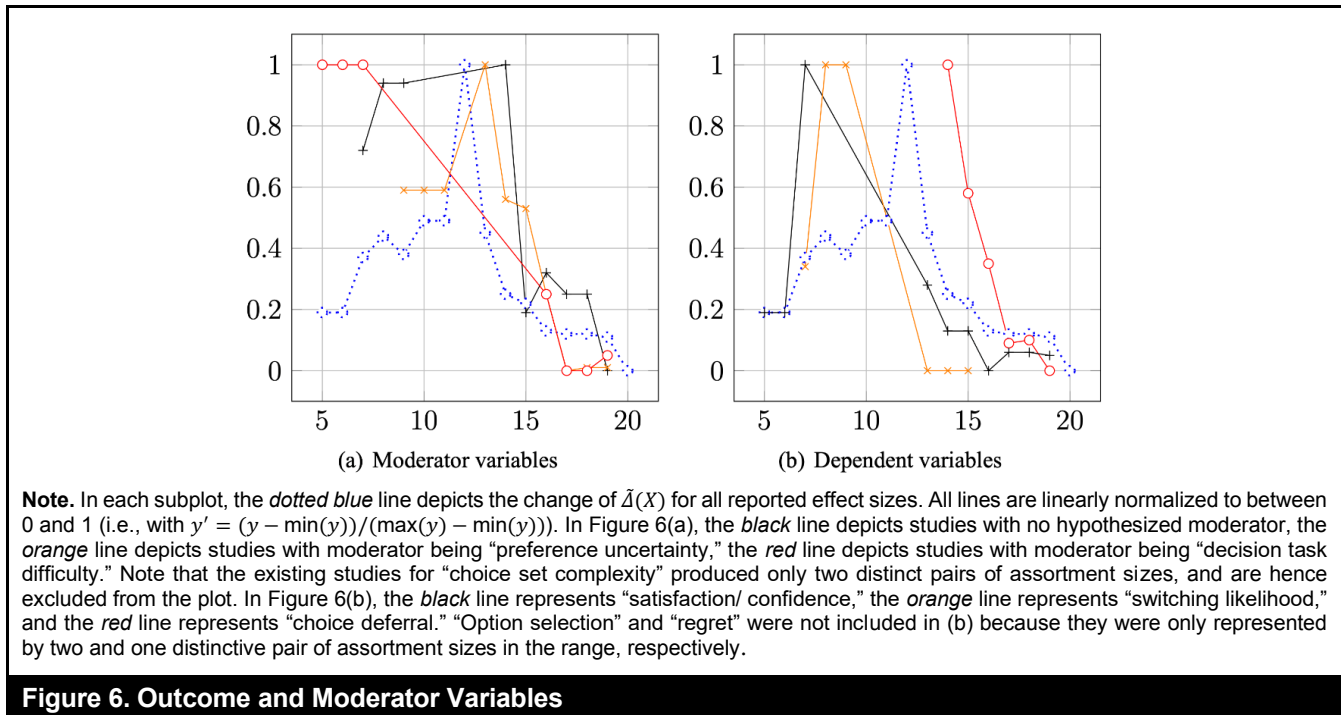
²² For example, the Lind-Mehlum method, which we used to determine if the dispersion indicator is an inverted U, assumes a point-wise continuous and differentiable quadratic function (Lind and Mehlum 2010).

²³ Let $C(X)$ be the number of existing studies with assortment sizes $X_1 + X_2 \in [2X - w, 2X + w]$. When $w < 3$, the number of consecutive values of X with $C(X) \geq 2$ never exceeds 5.

Table 4. Change of Dispersion Indicator $\tilde{D}(X)$ with X

X	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\tilde{D}_3$	.035	.035	.069	.083	.071	.091	.091	.188	.084	.047	.041	.025	.023	.022	.020	N/A
$\tilde{D}_4$	.035	.069	.069	.065	.084	.091	.095	.104	.047	.046	.041	.023	.023	.022	.022	N/A
$\tilde{D}_5$	.035	.069	.069	.063	.084	.084	.095	.084	.062	.046	.043	.034	.023	.022	.022	.020
σ	0.25	0.25	0.49	0.60	0.57	0.96	0.96	1.74	0.92	0.56	0.54	0.46	0.54	0.54	0.48	N/A
$\bar{\delta}_X$	7.2	7.2	7.5	7.8	9.2	11.6	11.6	9.0	11.2	15.3	16.5	18.8	23.7	23.4	24.0	24.0

Note: $\tilde{D}_i$ = value of $\tilde{D}(X)$ given a window size of $w = i$. The bold font cells indicate the top-3 values of $\tilde{D}(X)$ for the corresponding window size. σ = raw standard deviation of effect sizes without the assortment-size-difference correction. $\bar{\delta}_X$ = the average difference between assortment sizes (used for correction).

**Figure 6. Outcome and Moderator Variables**

As a robustness check, we examined how the dispersion indicator is affected by the assortment-size-difference correction in its operationalization (i.e., the denominator $X_{i2} - X_{i1}$ for computing $\tilde{D}(X)$ in Equation 25). As can be seen from Table 4, the correction indeed had no qualitative effect on the nature of dispersion. Specifically, even when we directly measured the standard deviation of reported effect sizes in each window (i.e., the row of σ), the dispersion was still maximized when X ranged from 10 to 12. Although not included in the table, we also conducted a variety of other robustness checks for the dispersion metric—e.g., by weighting an effect size with its corresponding sample size in computing the standard deviation (Hunter and Schmidt 2004), through a leave-one-out analysis on the included studies, by excluding studies with too wide (e.g., ≥ 20) or too narrow (e.g., ≤ 3) of an assortment-size

difference, etc.—but did not find notable changes on the shape of the dispersion indicator.

Finally, Figure 6 depicts how the dispersion indicator $\tilde{D}(X)$ varies with X when considering only a subset of the existing results featuring one type of moderator variable or one operationalization of the dependent variable. Most of these factorized studies returned an inverted-U shaped $\tilde{D}(X)$. For example, as predicted earlier in the paper, the type of moderator variables corresponding to consumer expertise, i.e., “preference uncertainty,” yielded a dispersion indicator that is clearly an inverted-U (i.e., the orange line in Figure 6a), reflecting our conceptual arguments that a change of consumer expertise likely moves the inverted-U option-satisfaction relation left/right. Nonetheless, note from the figure that there are also a few cases where the dispersion

indicator appears monotonic. We caution that the limited sample sizes in these individual cases likely constrain the insights one can derive from the computed $\hat{A}(X)$. For example, no existing study that examined the moderator type “decision task difficulty” (i.e., the red line in Figure 6a) chose a pair of assortment sizes with an average of between 6.5 and 16. This makes the decreasing line in the figure hard to interpret: It could be reflecting a monotonic (or U-shaped) dispersion indicator. Alternatively, the dispersion indicator could form an inverted-U with a turning point $X^* \in (6.5, 16)$, yet the lack of a sample in this range makes the inverted-U impossible to identify based on the existing evidence.

Case Study of Choice Overload in E-Commerce

The choice-overload effect has a strong potential to stimulate practical implications in an e-commerce context but the existing literature does not provide a sufficient number of primary studies (e.g., 10 or more; Sterne et al. 2011) to allow for a conventional meta-analysis. To address this limitation, we curated the existing empirical evidence in finer granularity than a typical meta-analysis. Specifically, we identified one of the largest randomized behavioral experiments (Moser et al. 2017), which examined the choice-overload effect in e-commerce platform design, and obtained the raw observations including 611 participants in the experiment. Then, we used the common subsampling method (e.g., the bag-of-little-bootstraps procedure; Kleiner et al. 2014) to randomly partition the raw observations into independent, simulated, primary studies, before using our method to perform a meta-analysis to assess the choice-overload effect in the context of e-commerce platform design.

More specifically, the participants of the experiment were assigned to one of six treatment conditions uniformly at random. All six treatment conditions feature the same e-commerce shopping web page (for chocolates). The only difference is the number of options (i.e., chocolates) displayed, which is 12, 24, 40, 50, 60, and 72²⁴ for the six conditions, respectively. After making a choice from the displayed options, the subjects were asked to rate their level of satisfaction with the choice on a 7-point Likert scale, from (1) *not at all satisfied* to (7) *extremely satisfied*. Figure 7(a) depicts the relationship between the observed satisfaction and the assortment size. While the observations do suggest an inverted-U option-satisfaction relationship, with the peak satisfaction level reached by an assortment size of between 24 and 50, Moser et al. (2017) showed that the traditional statistical tests for two-group experiments failed to reject the null hypothesis (i.e., choice overload does not exist). Note

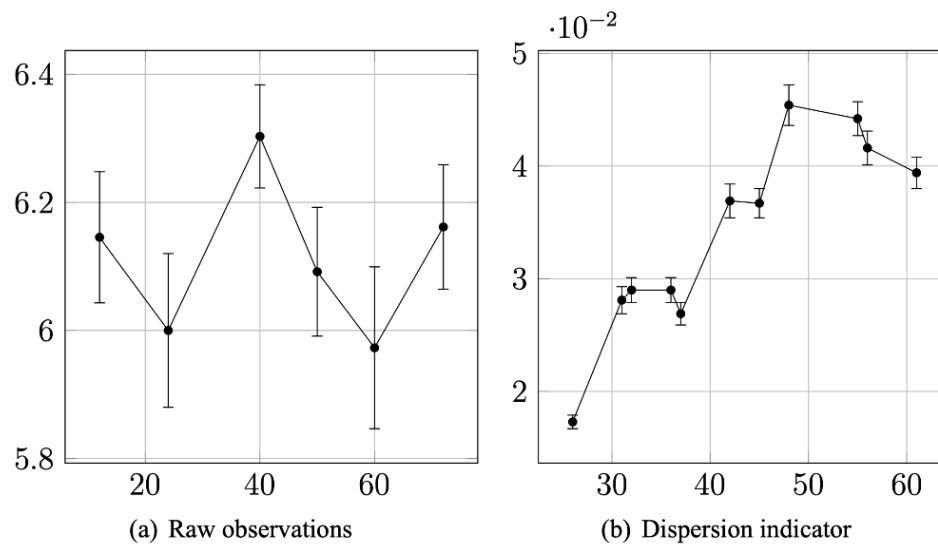
that, while there was no explicit manipulation of any moderating variable in the experiment design, Moser et al. (2017) duly noted the existence of many potential factors (e.g., individual differences in personality or experience) that could make the option-satisfaction relation idiosyncratic for different subjects. As discussed earlier, it is this idiosyncrasy of the option-satisfaction relation that enables the dispersion-based pattern leveraged by our method.

Theoretical Implications for Choice Overload

This paper contributes to the burgeoning literature on choice overload, which we move forward in several ways. First, we reconciled the fragmented empirical views of choice overload by explaining why the wide dispersion of reported effect sizes is not an indication of the malleability of the choice-overload effect but is instead *to be expected* when the underlying option-satisfaction relation is an inverted-U. We did so through the development of a novel conceptual and mathematical link between a moderated inverted-U relationship and the *dispersion* of effect sizes observed in two-group experiments. By identifying a unique pattern of dispersion among the existing empirical evidence, we provide strong evidence for the presence of the choice-overload effect. Despite the copious empirical studies about choice overload and the repeatedly posited theory of the option-satisfaction relation being an inverted-U (Grant and Schwartz 2011; Reutskaja and Hogarth 2009; Shah and Wolford 2007), our work represents the first method to use the former to test the latter and shows how the inverted-U conceptualization can account for the heterogeneity among the empirical evidence.

Second, we reconciled the fragmented views of choice overload by leveraging the existing empirical evidence to explicate the latent moderation mechanisms. The debate in the existing meta-analyses for choice overload is framed around the question of *what* variables moderate the choice-overload effect rather than *how* the moderators condition the option-satisfaction relationship. We identified and explained the two types of moderation for an inverted-U option-satisfaction relation, illuminated their distinctive implications in theory and practice, and leveraged our novel analytical method to demonstrate that the moderation in choice overload likely entails a considerable shift of the inverted-U curve to the left and right. This finding not only highlights the importance of choice overload for assortment design in practice but points to fruitful future research opportunities for further explicating and quantifying the effects of individual moderator variables on the inverted-U relation, which we will elaborate later in this section.

²⁴ These six assortment sizes were derived from an examination of the most popular e-commerce websites - see Moser et al. (2017) for details.



Note: The x -axis represents the assortment size in both subplots. The y -axis represents the choice satisfaction in the left subplot and the value of the dispersion indicator $\hat{\Delta}(X)$ in the right subplot. In both subplots, each dot represents the mean of the corresponding variable, while the length of the error bar is one standard error towards either direction.

Figure 7. Case Study for Online Browsing

Methodological Contributions

Our research findings have fundamental methodological contributions to experimental research on curvilinear relationships beyond the choice-overload effect. Specifically, our research answers the call for the development of advanced meta-analytic methods for examining the potential curvilinearity of a relationship (Johns 2006; Pierce and Aguinis 2013). The additional intricacies of curvilinear relationships (compared with linear ones) give rise to significant challenges in meta-analysis, where the notion of linearity has underpinned the development of most existing methods (Hunter et al. 2006). Indeed, with limited exceptions (e.g., Sturman 2003), there are not even adequate procedures to test for the presence of a curvilinear relationship in a meta-analysis of experimental studies. Our novel analytical method presented in this paper represents a first step towards bridging this gap. Specifically, our method of inferring the existence and moderation of an inverted-U relationship from two-group experiments is pertinent for at least three reasons:

First, it allows researchers to tap the gold mine of existing empirical evidence in theorizing and testing an inverted-U relationship, even though the existing research designs are all based on the linear mechanism (which is a common occurrence when examining curvilinear relationships; Aguinis et al. 2009). Second, our method enables the statistical disentanglement of the two types of inverted-U moderation, even though such moderation has never been theorized in the

literature nor brought to bear in research designs. Again, this facilitates the theoretical development and empirical examination of curvilinear relationships, even before new experiments dedicated to the study of curvilinearity can be conducted and accumulated over time. Finally, the findings of our method provide valuable guidelines for the design of future experimental studies. Specifically, when designing an experiment dedicated to the examination of an inverted-U relation, researchers may select a small set of “critical” independent-variable (i.e., X) values according to the estimations produced by our method, instead of having to examine the full spectrum of X . For example, if our method suggests a moderating effect that steepens/flattens the curve, researchers may want to select one value of X adjacent to the turning point and two values far away from the turning point to its left and right, respectively, in order to exemplify the change of steepness in the observed effect. On the other hand, if our method suggests a moderating effect that shifts the curve left/right, researchers may want to select multiple values of X around the potential range of the turning point in order to pinpoint the degree of shift conditioned by the moderation.

Broader Research Implications

More broadly, our research is premised on the notion of *analytical meta-analysis*, specifically the introduction of data analytics and machine learning techniques to the methodological arsenal of meta-analysis in behavioral

research. While still an emerging method, analytical meta-analysis has shown recent successes in several disciplines. For example, Zhang et al. (2020) leveraged the recent breakthroughs in theoretical machine learning on the decomposition of Gaussian mixtures (Kalai et al. 2012) to introduce mixture decomposition algorithms to the analysis of effect-size distribution in meta-analyses. Meager (2019), on the other hand, used graphical models to examine the latent factors driving the heterogeneity of existing findings.

The importance and promise of analytical meta-analysis is evidenced by the numerous and intense debates on reproducibility in behavioral sciences (e.g., Open Science Collaboration 2015), which speak to the importance of gleaning *deeper* insights from the synthesization of behavioral research findings than the simple aggregation of existing findings into statistical indicators (like what many existing meta-analytic methods do), especially when such findings are ostensibly inconsistent with each other. The rapid advancement of data analytics and machine learning techniques, on the other hand, makes such techniques ripe for adoption, exploration, and innovation in the space of meta-analysis. For example, our current work demonstrates that the use of data analytics techniques can not only reconcile the empirical inconsistencies of the choice-overload effect in behavioral research but can also develop a deeper understanding of its theoretical underpinning.

The research space of analytical meta-analysis is vast and its promise immense. Numerous research questions in the behavioral sciences have been blunted by the inconsistency of existing empirical evidence in the literature (Open Science Collaboration 2015). Some well-known ones include the impact of intrateam trust on team performance (Zhang et al. 2020), the effectiveness of fear appeals on persuasion (Tannenbaum et al. 2015), the link between happiness and success (Lyubomirsky et al. 2005), etc. All these research areas may stand to gain from the development of novel methods in analytical meta-analysis. We contend that, as a scholarly community with diverse methodological perspectives (e.g., employing behavioral, econometric, analytical, and computational methods; Rai 2016), the IS community is uniquely positioned to develop such fresh research perspectives and contribute to this emerging area of analytical meta-analysis. Indicatively, one of the first methods that introduced machine learning to meta-analysis was developed primarily by IS researchers (Zhang et al. 2020). From this perspective, it is our hope that the paper will inspire more “cross-paradigm connective IS research” (Rai 2018, p. v) in the near future. As demonstrated in this work, through the process of combining a behavioral theory with an analytical method, we can generate novel insights capable of connecting two disjointed worlds.

Limitations and Future Directions

It is important to note several limitations to our study. In terms of the generalizability of our method to other (potentially) curvilinear effects, we note that the independent variable of the option-satisfaction relation (i.e., assortment size) is an explicitly designed *absolute measure* (Cronbach et al. 1972) that has a common scale across studies. This stands in contrast to relationships that have independent variables relying on a normative or relative standard for measurement. Examples here include many psychological constructs such as personality traits, the measures for which tend to compare individuals against some norm groups rather than an absolute criterion (Cronbach et al. 1972). To apply our method to these relationships, future work needs to examine how to infer the change of the dispersion indicator with the independent variable when values of the independent variable are not directly comparable across studies.

With regard to the substantive examination of the choice-overload effect, the current method is limited in terms of the resolution it offers. For example, our method tests for the presence of a moderating effect that shifts the turning point left/right, but cannot pinpoint the exact amount of the left/right shift. Further, while we focused on distinguishing an inverted-U relationship (and its moderation) from a monotonic relationship, there may be practical needs to unpack further details of the inverted-U relationship, like the steepness on either side of the turning point. We elaborate on these limitations and the corresponding future research directions below.

High-Resolution Moderation Inference

Ultimately, inferring the moderation mechanism in a meta-analysis involves disentangling the multiple components that together form the distribution of the observed effect sizes. From a methodological perspective, Hunter and Schmidt (2004) noted that, when a moderator variable takes on a continuum of values, the effect-size distributions tends to be a mixture of many distributions, each having a different mean and variance. In the specific context of choice overload, Chernev et al. (2010) argued, and Simonsohn et al. (2014) concurred, that many existing experiments were “designed to document” how the sign of the effect can be reversed by a change of the moderator level, making the distribution of the reported effect sizes a mixture of two distributions, one having a positive mean effect size and the other negative. With this backdrop, the method developed in this paper can be construed as detecting whether the observed mixture distribution (of effect sizes) likely consists of two components produced by two inverted-U option-satisfaction relations with different turning points. Obviously, to enable the inference of the moderation mechanism with a higher resolution—e.g., to estimate the exact shift of the turning

point—future work needs to infer greater details about each of the components that together form the observed mixture distribution of the reported effect sizes.

To this end, a promising direction for future research stems from a recent breakthrough in theoretical machine learning on the decomposition of mixture distributions (Kalai et al. 2012). Traditionally, it was notoriously difficult to decompose an effect-size distribution for two reasons. The first reason is the limited sample size, as the number of reported effect sizes did not exceed a hundred even for an extensively studied effect like choice overload. The second, and more important, reason is that, because of the complexity of moderation, different components of the mixture decomposition likely overlap considerably with each other. The conventional methods for mixture decomposition, like the expectation maximization (EM) algorithm (Dempster et al. 1977) and its modern variations (e.g., Dasgupta 1999), are known to produce inaccurate results when two adjacent mixture components are not sufficiently separate from each other (e.g., in the case of the EM algorithm, when the mean difference between such two components is smaller than their standard deviations; Redner and Walker 1984). In 2010, a trio of breakthroughs in theoretical machine learning (Belkin and Sinha 2010; Kalai et al. 2010; Moitra and Valiant 2010) successfully addressed this challenge by developing algorithms that directly estimate the parameters of each mixture component without attempting to infer the component affiliation for each input sample (like the “M” step in the EM algorithm). With the more recent development of this idea, state-of-the-art algorithms can now accurately and efficiently decompose a mixture distribution even when the components overlap almost entirely with each other (e.g., Bandi et al. 2019). To further discern the moderation mechanisms of choice overload, future work could leverage these recent advances in mixture decomposition to first unpack the various components of the observed effect-size distribution before examining the role of moderator variables in forming these components.

Mathematical Model of Inverted U

In the mathematical analysis part of the paper, we adopted the conventional model of an inverted-U relation as a polynomial function (Lind and Mehlum 2010). The usage of this polynomial model provided valuable insights for our work and has proven fruitful for examining inverted-U relations in adjacent fields (e.g., see reviews in Haans et al. 2016; Pierce and Aguinis 2013). However, to fully appreciate the subtlety of an inverted-U relationship like the option-satisfaction relation, future work may need to examine the modeling of inverted-U relations beyond polynomial functions. Specifically, we note that while polynomials can often closely fit an inverted-U function within a limited range of the independent variable,

their fit could become questionable once the range is substantially expanded. For example, a quadratic model of inverted-U, as commonly used in the literature (Haans et al. 2016), would predict that consumer satisfaction decreases more when the assortment size changes from 100 to 110 than when it changes from 15 to 25. This is obviously counterintuitive, as one would reasonably expect consumers to barely notice a change in assortment size from 100 to 110 but not a change from 15 to 25. This counterintuitive phenomenon is unlikely to cause problems when analyzing the existing empirical results, as most of them focused on physical goods with a limited range of assortment sizes. However, it could become problematic for future research attempts to compare the choice-overload effect across physical and online settings, for which the range of the assortment size may have to be considerably expanded.

Fundamentally, this problem is rooted in a well-known challenge for polynomial interpolation in numerical analysis (Dahlquist and Bjork 1974). While the Weierstrass approximation theorem (Cheney and Light 2009, p. 151) ensures that any continuous function $F(X)$ can be approximated by a set of polynomial functions $P(X)$, the actual polynomial functions that can achieve such close approximations, even for a small closed range of X , can be extremely complex (e.g., Bernstein polynomials; Lorentz 2013). Further, as indicated by Runge’s phenomenon (Dahlquist and Bjork 1974), when fitting the polynomial function based on a limited number of observations, we may find a polynomial function $P(X)$ that perfectly fits an inverted-U $F(X)$ on all given observations yet produces unbounded errors on the unobserved values of X . To address these issues, future work may need to refine the model of the inverted-U function, e.g., by leveraging the techniques developed in numerical analysis to mitigate the errors of polynomial interpolation. Examples here include the development of alternative models, such as the use of spline interpolation with piece-wise polynomial functions (Hall and Meyer 1976) or the introduction of a regularization term (e.g., the ℓ_2 -norm of the first derivative of the polynomial function; De Boor et al. 1978).

Acknowledgments

The authors would like to thank the senior editor, associate editor, and reviewers for their helpful comments and suggestions, which greatly improved the quality of this work. The authors would also like to thank Eunice Kim and Echo Wen Wan as well as participants of the Research Seminar Series at the George Washington University’s School of Business for their valuable feedback. The authors were supported in part by the U.S. National Science Foundation (NSF) under Grant 1850605 and by the Defense Advanced Research Projects Agency (DARPA) under Grant HR00111920023. Any opinions, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF and DARPA.

References

- Aguinis, H., Pierce, C. A., Bosco, F. A., and Muslin, I. S. 2009. "First Decade of Organizational Research Methods: Trends in Design, Measurement, and Data-Analysis Topics," *Organizational Research Methods* (12:1), pp. 69-112.
- Aiken, L. S., West, S. G., and Reno, R. R. 1991. *Multiple Regression: Testing and Interpreting Interactions*, SAGE.
- Bandi, H., Bertsimas, D., and Mazumder, R. 2019. "Learning a Mixture of Gaussians via Mixed-Integer Optimization," *INFORMS Journal on Optimization* (1:3), pp. 221-240.
- Bangert-Drowns, R. L., and Rudner, L. M. 1990. Meta-analysis in Educational Research, *ERIC Digest*, Article# ED339748 (<https://files.eric.ed.gov/fulltext/ED339748.pdf>).
- Baumol, W. J., and Ide, E. A. 1956. "Variety in Retailing," *Management Science* (3:1), pp. 93-101.
- Belkin, M., and Sinha, K. 2010. "Toward Learning Gaussian Mixtures with Arbitrary Separation," in *Proceedings of the 23rd Annual Conference on Learning Theory*, pp. 407-419.
- Bettman, J. R., Luce, M. F., and Payne, J. W. 1998. "Constructive Consumer Choice Processes," *Journal of Consumer Research* (25:3), pp. 187-217.
- Brehm, J. W. 1966. *A Theory of Psychological Reactance*, Academic Press.
- Breusch, T. S., and Pagan, A. R. 1979. "A Simple Test for Heteroscedasticity and Random Coefficient Variation," *Econometrica* (47:5), pp. 1287-1294.
- Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmeld, A., Chan, T., and others. 2016. "Evaluating Replicability of Laboratory Experiments in Economics," *Science* (351:6280), pp. 1433-1436.
- Cheney, E. W., and Light, W. A. 2009. *A Course in Approximation Theory*, American Mathematical Society.
- Chernev, A. 2003. "When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice," *Journal of Consumer Research* (30:2), pp. 170-183.
- Chernev, A. 2006. "Decision Focus and Consumer Choice Among Assortments," *Journal of Consumer Research* (33:1), pp. 50-59.
- Chernev, A., Böckenholt, U., and Goodman, J. 2010. "Commentary on Scheibehenne, Greifeneder, and Todd Choice Overload: Is There Anything to It?" *Journal of Consumer Research* (37:3), pp. 426-428.
- Chernev, A., Böckenholt, U., and Goodman, J. 2015. "Choice Overload: A Conceptual Review and Meta-Analysis," *Journal of Consumer Psychology* (25:2), pp. 333-358.
- Chernev, A., and Hamilton, R. 2009. "Assortment Size and Option Attractiveness in Consumer Choice Among Retailers," *Journal of Marketing Research* (46:3), pp. 410-420.
- Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum.
- Coombs, C. H., and Avrunin, G. S. 1977. "Single-Peaked Functions and the Theory of Preference," *Psychological Review* (84:2), pp. 216-230.
- Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N. 1972. *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*, Wiley.
- Dahlquist, G., and Bjork, A. 1974. "Equidistant Interpolation and the Runge Phenomenon," in *Numerical Methods*, G Dahlquist, Å. Björk (eds.), Prentice Hall, pp. 101-103.
- Dasgupta, S. 1999. "Learning Mixtures of Gaussians," in *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*, pp. 634-644.
- De Boor, C., De Boor, C., Mathématicien, E.-U., De Boor, C., and De Boor, C. 1978. *A Practical Guide to Splines*, Springer.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)* (39:1), pp. 1-22.
- Dhar, R. 1997. "Consumer Preference for a No-Choice Option," *Journal of Consumer Research* (24:2), pp. 215-231.
- Diehl, K., and Poynor, C. 2010. "Great Expectations?! Assortment Size, Expectations, and Satisfaction," *Journal of Marketing Research* (47:2), pp. 312-322.
- Doll, W. J., and Torkzadeh, G. 1988. "The Measurement of End-User Computing Satisfaction," *MIS Quarterly* (12:2), pp. 259-274.
- Driver, M. J., and Streufert, S. 1969. "Integrative Complexity: An Approach to Individuals and Groups as Information-Processing Systems," *Administrative Science Quarterly* (14:2), pp. 272-285.
- Geva, H., Barzilay, O., and Oestreicher-Singer, G. 2019. "A Potato Salad with a Lemon Twist: Using a Supply-Side Shock to Study the Impact of Opportunistic Behavior on Crowdfunding Platforms," *MIS Quarterly* (43:4), pp. 1227-1248.
- Goodman, J. K., and Malkoc, S. A. 2012. "Choosing Here and Now Versus There and Later: The Moderating Role of Psychological Distance on Assortment Size Preferences," *Journal of Consumer Research* (39:4), pp. 751-768.
- Grant, A. M., and Schwartz, B. 2011. "Too Much of a Good Thing: The Challenge and Opportunity of the Inverted U," *Perspectives on Psychological Science* (6:1), pp. 61-76.
- Haans, R. F., Pieters, C., and He, Z.-L. 2016. "Thinking about U: Theorizing and Testing U- and Inverted U-Shaped Relationships in Strategy Research," *Strategic Management Journal* (37:7), pp. 1177-1195.
- Hall, C. A., and Meyer, W. W. 1976. "Optimal Error Bounds for Cubic Spline Interpolation," *Journal of Approximation Theory* (16:2), pp. 105-122.
- He, Y., Guo, X., and Chen, G. 2019. "Assortment Size and Performance of Online Sellers: An Inverted U-Shaped Relationship," *Journal of the Association for Information Systems* (20:10), 1503-1530.
- Huber, J., Payne, J. W., and Puto, C. 1982. "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research* (9:1), pp. 90-98.
- Hunter, J. E., and Schmidt, F. L. 2004. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, SAGE.
- Hunter, J. E., Schmidt, F. L., and Le, H. 2006. "Implications of Direct and Indirect Range Restriction for Meta-Analysis Methods and Findings," *Journal of Applied Psychology* (91:3), pp. 594-612.
- Inbar, Y., Botti, S., and Hanko, K. 2011. "Decision Speed and Choice Regret: When Haste Feels Like Waste," *Journal of Experimental Social Psychology* (47:3), pp. 533-540.
- Iyengar, S. S., and Lepper, M. R. 2000. "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of Personality and Social Psychology* (79:6), pp. 995-1006.

- Jachimowicz, J. M., Duncan, S., Weber, E. U., and Johnson, E. J. 2019. "When and Why Defaults Influence Decisions: A Meta-Analysis of Default Effects," *Behavioural Public Policy* (3:2), pp. 159-186.
- Johns, G. 2006. "The Essential Impact of Context on Organizational Behavior," *Academy of Management Review* (31:2), pp. 386-408.
- Johnson, E. J., Shu, S. B., Dellaert, B. G., Fox, C., Goldstein, D. G., Häubl, G., Larrick, R. P., Payne, J. W., Peters, E., Schkade, D., and others. 2012. "Beyond Nudges: Tools of a Choice Architecture," *Marketing Letters* (23:2), pp. 487-504.
- Jones, Q., Ravid, G., and Rafaeli, S. 2004. "Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration," *Information Systems Research* (15:2), pp. 194-210.
- Kahn, B. E., and Wansink, B. 2004. "The Influence of Assortment Structure on Perceived Variety and Consumption Quantities," *Journal of Consumer Research* (30:4), pp. 519-533.
- Kahn, B., Moore, W. L., and Glazer, R. 1987. "Experiments in Constrained Choice," *Journal of Consumer Research* (14:1), pp. 96-113.
- Kalai, A. T., Moitra, A., and Valiant, G. 2010. "Efficiently Learning Mixtures of Two Gaussians," in *Proceedings of the 42nd ACM Symposium on Theory of Computing*, 553-562.
- Kalai, A. T., Moitra, A., and Valiant, G. 2012. "Disentangling Gaussians," *Communications of the ACM* (55:2), pp. 113-120.
- Kamis, A., Koufaris, M., and Stern, T. 2008. "Using an Attribute-Based Decision Support System for User-Customized Products Online: An Experimental Investigation," *MIS Quarterly* (32:1), pp. 159-177.
- Kelley, H. H. 1973. "The Processes of Causal Attribution," *American Psychologist* (28:2), pp. 107-128.
- Kleiner, A., Talwalkar, A., Sarkar, P., and Jordan, M. I. 2014. "A Scalable Bootstrap for Massive Data," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* (76:4), pp. 795-816.
- Levav, J., and Zhu, R. 2009. "Seeking Freedom Through Variety," *Journal of Consumer Research* (36:4), pp. 600-610.
- Lind, J. T., and Mehlum, H. 2010. "With or Without u? The Appropriate Test for a U-Shaped Relationship," *Oxford Bulletin of Economics and Statistics* (72:1), pp. 109-118.
- Lorentz, G. G. 2013. *Bernstein Polynomials*, American Mathematical Society.
- Lyubomirsky, S., King, L., and Diener, E. 2005. "The Benefits of Frequent Positive Affect: Does Happiness Lead to Success?" *Psychological Bulletin* (131:6), pp. 803-855.
- McGuire, W. J. 1997. "Creative Hypothesis Generating in Psychology: Some Useful Heuristics," *Annual Review of Psychology* (48:1), pp. 1-30.
- McShane, B. B., and Böckenholt, U. 2018. "Multilevel Multivariate Meta-Analysis with Application to Choice Overload," *Psychometrika* (83:1), pp. 255-271.
- Meager, R. 2019. "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments," *American Economic Journal: Applied Economics* (11:1), pp. 57-91.
- Mogilner, C., Rudnick, T., and Iyengar, S. S. 2008. "The Mere Categorization Effect: How the Presence of Categories Increases Choosers' Perceptions of Assortment Variety and Outcome Satisfaction," *Journal of Consumer Research* (35:2), pp. 202-215.
- Moitra, A., and Valiant, G. 2010. "Settling the Polynomial Learnability of Mixtures of Gaussians," in *Proceedings of the 51st Annual Symposium on Foundations of Computer Science*, Las Vegas, Nevada, October 24-26, pp. 93-102.
- Morrin, M., Broniarczyk, S. M., and Inman, J. J. 2012. "Plan Format and Participation in 401 (k) Plans: The Moderating Role of Investor Knowledge," *Journal of Public Policy & Marketing* (31:2), pp. 254-268.
- Moser, C., Phelan, C., Resnick, P., Schoenebeck, S. Y., and Reinecke, K. 2017. "No Such Thing as Too Much Chocolate: Evidence Against Choice Overload in e-Commerce," in *Proceedings of the Annual Conference on Human Factors in Computing Systems*, pp. 4358-4369.
- Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science," *Science* (349:6251), Article aac4716.
- Pan, B., Zhang, L., and Law, R. 2013. "The Complex Matter of Online Hotel Choice," *Cornell Hospitality Quarterly* (54:1), pp. 74-83.
- Pierce, J. R., and Aguinis, H. 2013. "The Too-Much-of-a-Good-Thing Effect in Management," *Journal of Management* (39:2), pp. 313-338.
- Rai, A. 2016. "The MIS Quarterly Trifecta: Impact, Range, Speed," *MIS Quarterly* (40:1), pp. 3-10.
- Rai, A. 2018. "Beyond Outdated Labels: The Blending of IS Research Traditions," *MIS Quarterly* (42:1), pp. 3-6.
- Redner, R. A., and Walker, H. F. 1984. "Mixture Densities, Maximum Likelihood and the EM Algorithm," *SIAM Review* (26:2), pp. 195-239.
- Reutsckaja, E., and Hogarth, R. M. 2009. "Satisfaction in Choice as a Function of the Number of Alternatives: When 'Goods Siate,'" *Psychology & Marketing* (26:3), pp. 197-203.
- Roberts, J. H., and Lattin, J. M. 1991. "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research* (28:4), pp. 429-440.
- Sánchez-Meca, J., Marín-Martínez, F., and Chacón-Moscoso, S. 2003. "Effect-Size Indices for Dichotomized Outcomes in Meta-Analysis," *Psychological Methods* (8:4), pp. 448-467.
- Scheibehenne, B., Greifeneder, R., and Todd, P. M. 2010. "Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload," *Journal of Consumer Research* (37:3), pp. 409-425.
- Scherer, A., Wunderlich, N. V., and Wangenheim, F. V. 2015. "The Value of Self-Service: Long-Term Effects of Technology-Based Self-Service Usage on Customer Retention," *MIS Quarterly* (39:1), pp. 177-200.
- Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., and Lehman, D. R. 2002. "Maximizing Versus Satisficing: Happiness Is a Matter of Choice," *Journal of Personality and Social Psychology* (83:5), pp. 1178-1197.
- Shah, A. M., and Wolford, G. 2007. "Buying Behavior as a Function of Parametric Variation of Number of Choices," *Psychological Science* (18:5), pp. 369-270.
- Simonsohn, U. 2018. "Two Lines: A Valid Alternative to the Invalid Testing of U-Shaped Relationships with Quadratic Regressions," *Advances in Methods and Practices in Psychological Science* (1:4), pp. 538-555.

- Simonsohn, U., Nelson, L. D., and Simmons, J. P. 2014. "p-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results," *Perspectives on Psychological Science* (9:6), pp. 666–681.
- Singh, P. V., Tan, Y., and Mookerjee, V. 2011. "Network Effects: The Influence of Structural Capital on Open Source Project Success," *MIS Quarterly* (35:4), pp. 813–829.
- Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., and others. 2011. "Recommendations for Examining and Interpreting Funnel Plot Asymmetry in Meta-Analyses of Randomised Controlled Trials," *BMJ* (343), Article d4002.
- Sturman, M. C. 2003. "Searching for the Inverted U-Shaped Relationship Between Time and Performance: Meta-Analyses of the Experience/Performance, Tenure/Performance, and Age/Performance Relationships," *Journal of Management* (29:5), pp. 609–640.
- Tannenbaum, M. B., Hepler, J., Zimmerman, R. S., Saul, L., Jacobs, S., Wilson, K., & Albarracín, D. (2015). Appealing to Fear: A Meta-Analysis of Fear Appeal Effectiveness and Theories. *Psychological Bulletin* (141:6), pp. 1178–1204.
- Thompson, S. G. 1994. "Systematic Review: Why Sources of Heterogeneity in Meta-Analysis Should Be Investigated," *BMJ* (309:6965), pp. 1351–1355.
- Xu, J., Benbasat, I., and Cenfetelli, R. T. 2014. "The Nature and Consequences of Trade-Off Transparency in the Context of Recommendation Agents," *MIS Quarterly* (38:2), pp. 379–406.
- Zhang, N., Wang, M., and Xu, H. 2020. "Disentangling Effect Size Heterogeneity in Meta-Analysis: A Latent Mixture Approach," *Psychological Methods*, Advance online publication (<https://doi.org/10.1037/met0000368>)
- Zygmund, A. 2002. *Trigonometric Series*, Cambridge University Press.

About the Authors

Nan Zhang is a professor of information technology and analytics in the Kogod School of Business at the American University. Before joining Kogod, he was a professor of information sciences at Pennsylvania State University and a professor of computer science at George Washington University. His research focuses on data analytics, data privacy, and machine learning. His work has received many prestigious awards, including the Communications of the ACM Research Highlight in 2020, the ACM SIGMOD Research Highlight Award in 2019, the National Science Foundation's CAREER award in 2008, and several best paper awards from leading research conferences in computer science.

Heng Xu is a professor of information technology and analytics in the Kogod School of Business at the American University, where she also serves as the director of the Kogod Cybersecurity Governance Center. Her current research focuses on information privacy, data ethics, and algorithmic fairness. Her work has received many awards, including the Operational Research Society's Stafford Beer Medal in 2018, the National Science Foundation's CAREER award in 2010, and others. She has also served on a broad spectrum of national leadership committees including co-chairing the Federal Privacy R&D Inter-agency Working Group in 2016 and serving on the National Academies Committee on Open Science in 2017.

Appendix

Proof of Theorem 1

Theorem 1: When $\mathcal{U}(X) = \beta_0 + \beta_1 X^\alpha + \beta_2 X^\alpha Z + \beta_3 Z$, the dispersion indicator $\Delta(X, \epsilon)$ has no local maximum with respect to X .

Proof: Consider a simpler (and more generic) form of $\mathcal{U}(X)$ as

$$\mathcal{U}(X) = F(X) + G(X)Z. \quad (26)$$

To prove that no local maximum exists for $\Delta(X, \epsilon)$, it is sufficient to prove that the “directional” version of it (i.e., the value inside the absolute-value function in Δ), which we denote as $\vec{\Delta}(X, \epsilon)$, is monotonic with respect to X .

$$\vec{\Delta}(X, \epsilon) = \frac{\partial d(X-\epsilon, X+\epsilon)}{\partial Z} \propto \frac{\partial \mathcal{U}(X-\epsilon)}{\partial Z} - \frac{\partial \mathcal{U}(X+\epsilon)}{\partial Z}. \quad (27)$$

Consider how $\vec{\Delta}(X, \epsilon)$ varies with X and ϵ :

$$\frac{\partial^2 \vec{\Delta}(X, \epsilon)}{\partial X \partial \epsilon} = \frac{\partial^3 d(X-\epsilon, X+\epsilon)}{\partial Z \partial X \partial \epsilon} \quad (28)$$

$$\propto \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \cdot \left(\frac{\partial G(X-\epsilon)}{\partial X} - \frac{\partial G(X+\epsilon)}{\partial X} \right) = \frac{\partial^2 G(X)}{\partial X^2}. \quad (29)$$

Note that, so long as $G(X)$ is either concave or convex, $\partial^2 \vec{\Delta}(X, \epsilon) / \partial X \partial \epsilon$ always stays at the same side of zero for all X , meaning that $\vec{\Delta}(X, \epsilon)$ must be monotonic with respect to X . When $\mathcal{U}(X) = \beta_0 + \beta_1 X^\alpha + \beta_2 X^\alpha Z + \beta_3 Z$, we have $G(X) = \beta_2 X^\alpha + \beta_3$, which is either concave²⁵ if $\beta_2 \alpha(\alpha - 1) \leq 0$ or convex if $\beta_2 \alpha(\alpha - 1) \geq 0$. Thus, $\vec{\Delta}(X, \epsilon)$ must be monotonic with respect to X , completing the proof. $\square$

Proof of Theorem 2

Theorem 2. When ϵ is sufficiently small, the dispersion indicator $\Delta(X, \epsilon)$ reaches its minimum at X^* ($X^* > 1$) if for all Z ,

$$\frac{\partial \mathcal{U}(X^*)}{\partial X} = 0, \quad (30)$$

which also implies that X^* is always the turning point for $\mathcal{U}(X)$ regardless of the moderator Z .

Proof: When $\partial \mathcal{U}(X^*) / \partial X = 0$, there must be

$$\lim_{\epsilon \rightarrow 0} \frac{d(X^*-\epsilon, X^*+\epsilon)}{\epsilon} \propto \lim_{\epsilon \rightarrow 0} \frac{\mathcal{U}(X^*-\epsilon) - \mathcal{U}(X^*+\epsilon)}{\epsilon} = 0. \quad (31)$$

Since $\partial \mathcal{U}(X^*) / \partial X = 0$ for all Z , we have

$$\lim_{\epsilon \rightarrow 0} \frac{\Delta(X^*, \epsilon)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \left| \frac{\partial d(X^*-\epsilon, X^*+\epsilon)}{\epsilon \partial Z} \right| = 0. \quad (32)$$

Since $\Delta(X, \epsilon) \geq 0$, Equation (32) proves that, when $\epsilon \rightarrow 0$, the dispersion indicator $\Delta(X, \epsilon)$ reaches its minimum possible value at $X = X^*$. $\square$

²⁵ Note that $X \geq 1$ per the definition of $\mathcal{U}(X)$.

Proof of Theorem 3

Theorem 3. When $\mathcal{U}(X)$ follows the spline model

$$\mathcal{U}(X) = \begin{cases} \beta_1 X + \beta_2 XZ, & \text{if } X \leq X^* \\ \beta_3 X + \beta_4 XZ + (\beta_1 - \beta_3)X^* + (\beta_2 - \beta_4)X^*Z, & \text{otherwise} \end{cases} \quad (33)$$

where $\beta_1 + \beta_2 Z > 0$ and $\beta_3 + \beta_4 Z < 0$ for all Z , the dispersion indicator $\Delta(X, \epsilon)$ is always monotonic.

Proof: Consider the value of the dispersion indicator $\Delta(X, \epsilon)$. Equation 33 yields

$$\Delta(X, \epsilon) = \frac{\partial d(X - \epsilon, X + \epsilon)}{\partial Z} = \begin{cases} 2\beta_2 \epsilon, & \text{if } X + \epsilon \leq X^* \\ (\beta_2 + \beta_4) \epsilon, & \text{if } X \in (X^* - \epsilon, X^* + \epsilon) \\ 2\beta_4 \epsilon, & \text{if } X - \epsilon \geq X^* \end{cases} \quad (34)$$

The theorem directly follows. $\square$