



Research paper

Learning and probability matching in newsvendor decisions

Weiwei Han^a , Xun Wang^b ,*^a School of Economics and Management, Beijing University of Posts and Telecommunications, China^b Cardiff Business School, Cardiff University, UK

ARTICLE INFO

Dataset link: <https://doi.org/10.6084/m9.figshare.12937172>

Keywords:

Behavioral economics

Newsvendor

Bounded rationality

Electroencephalography

ABSTRACT

Human decision making in the newsvendor paradigm systematically deviates from the predictions of expected profit maximization, yet the cognitive mechanisms underlying these deviations remain contested. We investigate whether these biases arise from distorted utility evaluation or from adaptive learning induced by repeated feedback. Across three laboratory experiments ($N = 214$), we compare behavior under decision from description, decision from experience, and newsvendor environments while varying information availability. The first experiment combines behavioral and electroencephalographic (EEG) data, the second replicates the behavioral findings without EEG, and the third extends the analysis to a multi-alternative setting with continuous demand. We document two robust patterns in newsvendor environments: decision accuracy inferior to random guessing, and positive correlation between choice and realized demand. These patterns are absent in one-shot descriptive decisions and cannot be explained by standard utility-based models. Neural evidence shows that repeated environments engage parietal and frontal activity associated with choice deliberation and feedback-based learning, and that memory encoding differs systematically across demand outcomes. Motivated by these findings, we develop a model incorporating probability matching and asymmetric belief updating, which provides a substantially improved fit to observed choices. Our results demonstrate that repeated feedback fundamentally alters the mechanism of economic decision making, highlighting the central role of learning and belief dynamics in generating newsvendor biases.

1. Introduction

The newsvendor problem is a canonical model of economic decision making under uncertainty, in which agents must choose supply quantities before demand is realized. Its applications span a wide range of economic and managerial contexts, including perishable inventory management, capacity planning under demand uncertainty (Van Mieghem, 2003), and airline overbooking (Talluri and Van Ryzin, 2006). Under the classical formulation (Arrow et al., 1951), where the demand distribution and cost parameters are known, the optimal decision follows directly from expected profit maximization. Despite this analytical simplicity, a large experimental literature documents systematic and persistent deviations from optimal ordering (Schweitzer and Cachon, 2000; Zhang and Siemsen, 2019). Existing explanations tackle specific empirical deviations but fail to provide a unified account.

Prior research on decision making under uncertainty suggests that the origins of bounded rationality depend on how information is acquired. A central distinction is between decision from description (DFD) and decision from experience (DFE) (Stewart et al., 2006; Erev et al., 2010; Wulff et al., 2018). In DFD, agents are explicitly provided with the payoff distribution of each option. In contrast, in DFE, agents are not provided with any descriptive information and must learn the underlying outcome distributions

* Corresponding author.

E-mail addresses: hanweiw@bupt.edu.cn (W. Han), wangx46@cardiff.ac.uk (X. Wang).

Table 1
Design of newsvendor experiments.

Study	Repeated	Demand distribution	Cost parameters	Payoff distribution
Schweitzer and Cachon (2000)	✓	✓	✓	✗
Bolton and Katok (2008)	✓	✓	✓	✗
Bostian et al. (2008)	✓	✓	✓	✗
Benzion et al. (2010)	✓	✗	✓	✗
Gavimani and Isen (2010)	✗	✓	✓	✗
Kremer et al. (2010)	✓	✓	✓	✓/✗
Zhao et al. (2016)	✓	✗	✓	✗
D'Urso et al. (2017)	✓	✓/✗	✓	✗
Kirshner and Moritz (2023)	✗	✓	✓	✗
Niranjan et al. (2023)	✓/✗	✓	✓	✗

from feedback obtained through past decisions.¹ This distinction produces systematically different behavioral patterns, termed the description-experience gap (Wulff et al., 2018). Existing evidence suggests that identical economic problems may engage qualitatively different cognitive mechanisms, such that deviations from optimality in DFD primarily reflect distortions in the evaluation of stated payoffs, whereas in DFE they reflect adaptive processes such as reinforcement learning and memory constraints. However, it remains unclear how these mechanisms operate in the newsvendor problem.

A review of existing experimental designs highlights ambiguity in how the newsvendor task has been conceptualized (Table 1). Newsvendor experiments vary substantially in both the information provided to participants and whether decisions are made once or repeatedly with feedback. In most implementations, decisions are repeated and participants are given the demand distribution and cost parameters. As a result, the newsvendor task (DNV), as commonly implemented, does not fit neatly into either the standard DFD or DFE framework. Unlike canonical DFD, the payoff distribution is not explicitly presented in a direct and static form; unlike canonical DFE, both demand uncertainty and cost parameters are typically described, making the payoff distribution computable. This hybrid structure suggests that the newsvendor task may engage both utility evaluation based on description and adaptive learning based on experience. However, despite this dynamic feedback structure, most existing explanations of newsvendor bias rely on static utility-based models developed for one-shot descriptive choice (Schweitzer and Cachon, 2000; Su, 2008; Ho et al., 2010; Ockenfels and Selten, 2014; Uppari and Hasija, 2019, see Section 2 for details).

This study investigates whether and how information availability alters the cognitive mechanisms underlying newsvendor decision making. Although the economic structure of the problem remains unchanged, embedding decisions in a dynamic environment with feedback may shift behavior from static utility evaluation towards adaptive learning. We report three studies. Study 1 employs a binary-demand experiment that varies information availability while recording electroencephalography (EEG) during decision making, allowing us to identify neural signatures associated with choice selection and feedback-based learning. We use DFD and DFE as benchmark paradigms with well-established cognitive interpretations, enabling us to isolate behavioral and neural mechanisms specific to the newsvendor environment. Study 2 replicates the main behavioral results in an independent newsvendor experiment without EEG, establishing robustness beyond the neurophysiological setting. Study 3 extends the analysis to a multi-alternative newsvendor environment, testing the generality of the proposed mechanisms in a richer decision context.

The results reveal systematic differences across paradigms that are consistent with a shift in the underlying decision mechanism. Under asymmetric demand distributions, we document a pull-over-uniform effect, whereby decision performance, measured as the proportion of optimal choices, falls below the benchmark implied by uniform random choice. This pattern arises specifically in hard decision problems, where the absolute difference between choice probabilities is small, and is observed when decisions are embedded in the DNV environment, but not in DFD. In addition, we find a strong demand-chasing effect, defined as a positive correlation between order quantities and the most recent demand realization, in both DFE and DNV, but not in DFD. These findings are robust across experimental conditions, including with and without EEG measurement and in both binary and continuous demand environments.

These findings provide converging evidence that demand information availability and repeated feedback engages adaptive learning processes that alter choice behavior relative to the predictions of static utility maximization. To identify the mechanisms underlying this shift, we examine behavioral and neural signatures of choice deliberation and belief updating. When the demand distribution is salient, choices systematically follow event probabilities rather than expected-profit maximization or random selection. Following feedback, beliefs are updated dynamically and asymmetrically across demand realizations. Neural evidence supports this interpretation: compared with the one-shot DFD paradigm, repeated paradigms of DFE and DNV show differential activity in parietal regions associated with choice deliberation and in frontal regions associated with feedback processing and learning. Motivated by these results, we develop a decision model tailored to the newsvendor environment that incorporates probability matching and asymmetric belief updating. The model accounts for the key newsvendor biases observed and provides

¹ Most DFD studies use a one-shot design in which each set of decision parameters is presented only once. However, repeated DFD designs are also possible, allowing subjects to encounter the same lotteries across multiple trials (see e.g. Lejarraga and Gonzalez, 2011). By construction, DFE designs involve repeated exposure to the same decision environment.

a substantially improved fit compared with standard utility-based benchmarks, demonstrating that learning-based mechanisms are necessary to explain newsvendor decisions in dynamic environments.

The remainder of this paper is organized as follows: Section 2 provides a brief literature review of past research on decision-from-experience and newsvendor decisions. Section 3 introduces the binary newsvendor task and presents a novel decision model that incorporates asymmetric learning and probability matching. The hypotheses are developed from the structural properties of the proposed model. Section 4 outlines the experiment design and data collection method employed in Study 1. The results of Study 1, derived from behavioral and EEG analyses, are presented in Section 5. Sections 6 and 7 detail two additional studies (Studies 2 and 3) that verify the findings in Study 1, conducted under conditions of non-EEG and multiple-alternative scenarios respectively. We conclude by presenting the theoretical and practical implications of this research in Section 8.

2. Literature review

In this section, we review relevant literature on empirical decision-making theory and behavioral studies on newsvendor decisions. In Section 2.1, we briefly survey theoretical and empirical research on decision-making from the description/experience perspective. Section 2.2 introduces the biases observed in newsvendor experiments. Section 2.3 summarizes the existing theories on newsvendor biases.

2.1. The description-experience gap

For DFD tasks where the payoff information is provided, the most prominent theory to date is prospect theory, and its variation, cumulative prospect theory (Kahneman and Tversky, 1979). Originally devised to address inconsistencies inherent in the expected utility theory, such as the Allais paradox (Allais, 1953), prospect theory incorporates elements of the reference effect, loss aversion, risk aversion, and probability distortion. The model offers flexibility by omitting certain components. However, concerns have been raised regarding its comparatively low performance in alternative forms of decision tasks (Erev et al., 2010). Specifically, although prospect theory has been successful in the DFD paradigm, its predictive power does not excel in DFE. Subjects' behavior in DFE shows a distinct difference from that in DFD, creating the so-called description-experience gap. For instance, whereas the subjects tend to overweight rare events in DFD tasks, they underweight rare events in DFE (Barron and Erev, 2003). In a model prediction competition reported by Erev et al. (2010), various models were compared in terms of their predicting ability in these tasks. The findings revealed that the best-performing model for DFE does not conform to the framework of prospect theory. Instead, it incorporates memory capacity and learning as essential features.

2.2. Newsvendor decision biases

In the newsvendor model (Arrow et al., 1951), given the demand distribution, unit price, cost, and salvage value, obtaining the optimal decision is straightforward. However, numerous studies have revealed consistent deviations between model prediction and actual decisions. Empirically, Fisher and Raman (1996) documented that the managers' order quantities consistently fall short of the expected profit-maximizing quantity. These findings were based on data collected from the fashion industry, where the newsvendor model finds application.

Since Schweitzer and Cachon (2000), researchers have increasingly adopted the laboratory experiment approach to investigate newsvendor behavior, leading to the identification of several bias patterns. The *pull-to-center effect*, where the average order quantity is pulled towards the mean demand from optimality, has been consistently observed across various experiment settings and subject categories, including sample size and feedback type (Bolton and Katok, 2008), demand distribution (Benzion et al., 2008), and work experience (Bolton et al., 2012). A recent meta-analysis conducted by Zhang and Siemsen (2019) supports the existence of pull-to-center effect. However, they concluded that *order asymmetry*, that the degree of pull-to-center is not symmetric in high/low profit margin scenarios, appears less as a systematic attribute of human behavior and more as a function of experiment design. The *demand chasing bias* has been confirmed by numerous studies (Schweitzer and Cachon, 2000; Bolton and Katok, 2008; Kremer et al., 2010; Kirshner and Moritz, 2021) through investigations into the correlation between choice and previous demand. Researchers have also observed the *order stability* phenomenon, which means the average order quantity remains remarkably stable across trials.

Limited knowledge exists on how information salience affects newsvendor bias. Contrary to common belief, salient demand distributions do not necessarily alleviate the pull-to-center effect (Benzion et al., 2008; Shinde et al., 2021), although they have been observed to reduce demand chasing behavior (D'Urso et al., 2017). Furthermore, there has been a lack of effort to comprehend these effects on a cognitive level.

2.3. Theoretical endeavors

Most theories dedicated to explaining the newsvendor bias fall into either heuristic or utility-based categories (Glöckner and Betsch, 2011), depending on whether a utility-maximizing mechanism is employed. In the heuristic category, Schweitzer and Cachon (2000) proposed two anchoring heuristics. The mean anchor heuristic posits that the decision-maker anchors on the mean demand and then adjusts towards the profit-maximizing order quantity. The chasing demand heuristic suggests that the decision-maker anchors on the most recent order quantity and partially adjusts toward the most recent demand observation. They found that only the former heuristic can account for the pull-to-center effect. Several studies have extended these heuristics to include the learning

processes. For instance, [Benzion et al. \(2008\)](#) identified a learning effect, whereby the order quantity eventually converges due to the learning effect. They proposed that the mean demand, theoretical optimal solution, and the most recent demand are key factors contributing to the pull-to-center effect. [Bostian et al. \(2008\)](#) developed an adaptive learning model where the subjective prospects are updated each trial based on received payoffs, yet this model fails to capture the order asymmetry effect. Forecast judgment is another heuristic explanation put forward by [Lee and Siemsen \(2017\)](#). They found that breaking down newsvendor decisions into forecast judgments and service-level decisions can improve decision accuracy.

In the utility-based category, researchers alter either the decision objective, the level of confidence, or the level of rationality to accommodate the biases. [Schweitzer and Cachon \(2000\)](#) concluded that the models based on risk aversion, loss aversion, waste aversion, stockout aversion, and prospect theory cannot adequately explain the pull-to-center effect, whereas the model of minimizing ex-post inventory error can account for it. Building upon the concept of minimizing ex-post inventory error, [Ho et al. \(2010\)](#) proposed a utility-based equilibrium model based on the psychological costs of leftovers and stockouts. This model posits that the psychological aversion to leftovers outweighs the disutility of stockouts. It encompasses the ex-post inventory error minimization model as a special case. Similarly, [Ockenfels and Selten \(2014\)](#) complemented the aforementioned model by proposing that the decision-makers seek to balance the ex-post stockout and holding costs (termed upward and downward impulses). These two theories introduce cost asymmetry, and can therefore account for both pull-to-center and order asymmetry. [Su \(2008\)](#) introduced a quantal choice model based on the Boltzmann equation to address pull-to-center. [Ren and Croson \(2013\)](#) presented an overprecision model, where decision-makers underestimate the probability of a rare event. However, these models fall short in accommodating order asymmetry and individual heterogeneity ([Nagarajan and Shechter, 2014; Uppari and Hasija, 2019](#)).

Prospect theory, as an integration of the above adjustments, has garnered attention in recent behavioral newsvendor research. [Schweitzer and Cachon \(2000\)](#) observed that a prospect-theoretic model with loss aversion and risk aversion can only predict the pull-to-center effect when the payoffs include both gains and losses, which contradicts empirical evidence that the pull-to-center effect exists even if all possible payoffs are positive. [Nagarajan and Shechter \(2014\)](#) added nonlinear probability distortion and removed loss aversion, yet they reached the same conclusion. [Long and Nasiry \(2015\)](#) emphasized the importance of the reference effect in prospect theory in explaining pull-to-center. [Uppari and Hasija \(2019\)](#) explored various reference points in newsvendor decision-making. They found that when setting the reference point as the payoff when the decision equals the mean demand, prospect theory can explain a variety of observed phenomena and provides a good fit to the empirical data. Two recent studies ([Vipin and Amit, 2019; Surti et al., 2020](#)) also underscored the significance of the reference effect.

While utility-based models are suitable for one-shot decision settings, their application to repeated decision environments is limited, mainly because the role of realized demand is often overlooked. First, these models cannot explain demand-chasing behavior, as they typically assume temporal independence across choices ([Regenwetter and Davis-Stober, 2018](#)). Second, many existing models (e.g., prospect theoretic; [Uppari and Hasija, 2019](#)) treat choices as the direct outcome of utility functions, which may underemphasize the role of demand distributions relative to cost parameters. Empirical evidence shows that the supply chain context exacerbates newsvendor bias ([Kremer et al., 2010](#)), and that decision makers place greater weight on realized demand than on cost information ([Gavirneni and Isen, 2010; Niranjan et al., 2023](#)).

3. Model and hypotheses

3.1. The binary newsvendor model

We first introduce the binary newsvendor model. All notation used in this paper is summarized in Table A.1 in the Online Appendix. The unit cost and unit price are fixed at c and p , respectively. The salvage value is assumed to be zero. We use d to denote the demand value and d' the counterfactual demand value. Their probabilities are x_d and $x_{d'}$. Similarly, the admissible order quantities are denoted by r and r' . In the experiment we set $\{d, d'\} = \{r, r'\} = \{1, 2\}$. The average demand is $\bar{d} = 2 - x_1$. We denote the payoff of ordered pair $\{d, r\}$ to be $V(d, r)$, which can be calculated in the traditional newsvendor fashion:

$$V(d, r) = \min(d, r)p - rc \quad (1)$$

The utility (expected payoff) of choice r is denoted by $U(r)$. When $r = 1$ (the safe choice), the payoff is $p - c$ regardless of the demand:

$$U(1) = x_1 V(1, 1) + x_2 V(2, 1) = p - c = pm,$$

where m is the profit margin, $m = (p - c)/p$. When $r = 2$ (the risky choice), the payoff is either $p - 2c$ or $2p - 2c$ depending on the demand realization. Therefore:

$$\begin{aligned} U(2) &= x_1 V(1, 2) + x_2 V(2, 2) \\ &= x_1(p - 2c) + x_2(2p - 2c) = p(2m - x_1) \end{aligned}$$

The absolute difference in utility between choices is $U_\delta = |U(1) - U(2)| = p|m - x_1|$. If $m < x_1$, then the optimal choice $r^* = 1$.

3.2. The proposed behavioral model and properties

Prior belief updating is essential to address the description-experience gap ([Aydogan, 2021](#)). We adopt a simple probability learning model that assumes that the posterior subjective probability changes linearly with the prediction error, defined as the

difference between the prior and the realized demand. The posterior of d increases if the realized demand is d and decreases otherwise. The change in the probability is proportional to the difference between the event indicator and the prior. Denote the subjective probability of d in trial t as $\tilde{x}_{d,t}$, the model can be expressed as

$$\tilde{x}_{d,t} = \begin{cases} \tilde{x}_{d,t-1} + \alpha_d(1 - \tilde{x}_{d,t-1}) & \text{if } d_t = d \\ \tilde{x}_{d,t-1} + \alpha_{d'}(0 - \tilde{x}_{d,t-1}) & \text{if } d_t = d' \end{cases} \quad (2)$$

The parameter α_d is the rate of learning, $0 < \alpha_d < 1$. The higher the value of α_d , the more the last conforming experience will affect the posterior. $\alpha_d \neq \alpha_{d'}$ means that the rate of learning is different for events d and d' . We term this mechanism *asymmetric encoding*. The initial values of the subjective probability vary by task. In the DFE task, we assume symmetric priors ($\tilde{x}_{d,0} = 0.5$, Aydogan, 2021) as subjects do not possess any information about the probability distribution. In DNV tasks, we assume descriptive priors ($\tilde{x}_{d,0} = x_d$) as the information of demand distribution is given *a priori* to the subjects.

The symmetric form of (2) has been widely used to model statistical learning processes (Bush and Mosteller, 1955). A linear learning model is adopted here, rather than a Bayesian framework, primarily for computational tractability in model fitting and to allow the incorporation of asymmetry. The asymmetric learning model posits that distinct learning rates are assigned to different classes of events (Daw et al., 2002; Gershman, 2015). Empirical research has shown that asymmetric reinforcement learning models provide a useful framework for explaining specific patterns of human behavior (Ciranka et al., 2022).

In selecting the error model, we propose that subjects' rationality depends on differences in subjective expected payoffs. An irrational but partially informed subject does not choose between alternatives with equal probability (and maximal entropy) but instead adopts a probability-matching strategy. We refer to this mechanism as *contingent rationality*. Probability matching describes the tendency of decision-makers to align their choice probabilities with the (subjective) probabilities of events (Vulkan, 2000). This phenomenon is commonly observed in stochastic decisions as a deviation from expected utility theory. Probability matching requires advanced cognitive abilities and considerable effort compared with pure randomization (Saldana et al., 2022). Consequently, it is less frequently observed in children and non-human primates (e.g., rhesus monkeys) than in adult human decision-makers (Wilson and Rollin, 1959; Derks and Paclisanu, 1967; Unturbe and Corominas, 2007). It is believed that whether decision makers adopt maximizing or matching strategy depends on relative mental efforts of these strategies (Schulze and Newell, 2016b).

We assume that at trial t , subjects adopt the probability matching strategy with probability y_t and the utility-maximization strategy with probability $1 - y_t$. The probability y_t exponentially decreases with the absolute difference in the subjective utility:

$$y_t = \exp \{ -\lambda |SU(r, t) - SU(r', t)| \} \quad (3)$$

where $SU(r, t)$ is subjective utility of choice r in trial t based on the subjective probability, $SU(r, t) = \tilde{x}_{d,t}V(d, r) + \tilde{x}_{d',t}V(d', r)$ and evolves with belief updating. The intuition is that the more indifferentiable the two choices are, the more likely for the subject to show a probability-matching behavior. The utility function here, as the expectation of payoffs, does not involve loss or risk propensity. A discussion on this assumption is offered in Section 8.

The probability of choosing r at trial t , $z_{r,t}$ is given as follows:

$$z_{r,t} = \tilde{x}_{r,t}y_t + \mathbf{1}\{\tilde{x}_{r,t} > m\}(1 - y_t) \quad (4)$$

$\mathbf{1}\{\tilde{x}_{r,t} > m\}$ is the indicator function that r is optimal under the current subjective probability. When $\lambda \rightarrow 0$, the choice distribution approaches the subjective event distribution. We name this model Contingent Rationality and Asymmetric Encoding (CRAE) to manifest the two important features.

The CRAE model possesses the following properties that allows it to accommodate the observed newsvendor biases. The formal mathematical proof of these properties can be found in Section A.2 of the Online Appendix. We denote $\gamma = \lambda p$, and $W_0(\cdot)$ as the Lambert W function.

Proposition 1 (Evolution of Subjective Probability). *The mean subjective probability is updated according to the iterative relation:*

$$E(\tilde{x}_{d,t}) = E(\tilde{x}_{d,\infty}) - (1 - x_d\alpha_d - x_{d'}\alpha_{d'})^\gamma [E(\tilde{x}_{d,\infty}) - \tilde{x}_{d,0}], \quad (5)$$

where $E(\tilde{x}_{d,\infty}) = x_d\alpha_d / (x_d\alpha_d + x_{d'}\alpha_{d'})$.

Remark. There exists an asymptote for the mean subjective probability and it approaches the asymptote monotonically from the initial prior. The value of the asymptote does not necessarily equal the real probability, and the deviation depends on α_d and $\alpha_{d'}$. Specifically, $E(\tilde{x}_{d,\infty})$ increases with $\alpha_d/\alpha_{d'}$. If $\alpha_d = \alpha_{d'}$, then $E(\tilde{x}_{d,\infty}) = x_d$.

Proposition 2 (Bounded Rationality). $0 < E(z_{r,t}) < 1$.

Remark. The strict inequality implies that the model does not predict constant choice. In other words, the mean response is pulled towards the direction of the center (the mean demand), but it may well be pulled over the midpoint (the mean of a uniform distribution). This can be seen from the fact that $E(z_{r,t})$ is bounded by 0 and 1 but not 0.5.

Proposition 3 (Pull-over-uniform). *If $E(\tilde{x}_{r,t}) < 1 - W_0(\gamma e^\gamma/2)/\gamma$, then $E(z_{r,t}) < 0.5$; if $E(\tilde{x}_{r,t}) > W_0(\gamma e^\gamma/2)/\gamma$, then $E(z_{r,t}) > 0.5$.*

Remark. Under the above circumstances, the expected choice probability $E(z)$ can be pulled over the uniform midpoint (0.5). When $\gamma = 1$, the critical values of subjective probability are 0.315 and 0.685 respectively. Thus the model predicts that the shape of demand distribution affects the choice distribution, and the bias is more prominent when the demand distribution is asymmetric e.g., Bernoulli, Poisson, or exponential, and when the mode of demand deviates from the optimal choice.

Proposition 4 (Asymmetry). Let A and B be two experiment settings with $x_d^A = x_{d'}^B$, $m^A = 1 - m^B$ and $\alpha_d > \alpha_{d'}$, then $E(z_{d,t}^A) > E(z_{d',t}^B)$ for all t .

Remark. Proposition 4 captures the asymmetry in either direction. Such asymmetry, as this model proposes, is caused by the asymmetry in learning rates and how fast the subjective probability updates.

Proposition 5 (Demand Chasing). Given $z_{r,t-1}$, if $d_{t-1} = r$, then $z_{r,t} > z_{r',t}$.

Remark. Proposition 5 indicates that the probability of choosing r increases if the last demand realization is r . Intuitively, it is caused by the updating mechanism of subjective probabilities and choice probabilities. This result directly leads to a positive correlation between d_{t-1} and r_t .

Proposition 6 (Stability). $Z_{\min} < E(z_{r,\infty}) - E(z_{r,0}) < Z_{\max}$, where

$$Z_{\min} = E(\tilde{x}_{r,\infty}) - \tilde{x}_{r,0} - \gamma[E(\tilde{x}_{r,\infty}) + \tilde{x}_{r,0} - E^2(\tilde{x}_{r,\infty}) - \tilde{x}_{r,0}^2],$$

$$Z_{\max} = E(\tilde{x}_{r,\infty}) - \tilde{x}_{r,0} + \gamma[E(\tilde{x}_{r,\infty}) + \tilde{x}_{r,0} - E^2(\tilde{x}_{r,\infty}) - \tilde{x}_{r,0}^2].$$

Remark. Proposition 6 indicates that the difference between the initial and final choice probabilities is bounded, and that the bounds are determined by the priors and posteriors. Thus, the order stability can be attributed to the change in subjective probability. During a given period, stable order quantity can be caused either by a small difference between priors and posteriors or by a slow change in the subjective probability update due to low learning rates.

3.3. Hypotheses development

We derive testable hypotheses from both the utility-based models and the proposed model. Each hypothesis reflects qualitative predictions designed to contrast these models. We focus on two mechanisms. The first concerns decision error, that is, how variation in choices arises. The second concerns learning, namely how individuals use past experience to guide future decisions. For each mechanism, we propose two hypotheses: one tested using behavioral choice data and the other using EEG evidence.

First, most utility-based theories are algebraic and predict identical choices when the same decision problem is repeated (Regenwetter and Davis-Stober, 2018). To account for stochastic choice, probabilistic extensions have been integrated into deterministic decision frameworks (Su, 2008; Erev et al., 2010), typically assuming that the probability of selecting an option depends on its utility. Under the logit (softmax) specification, where the error term in the utility function follows a Gumbel distribution, the probability of choosing option x from choice set A is given by $\exp(\lambda U(x)) / \sum_{y \in A} \exp(\lambda U(y))$. Here, x and y denote options, λ is a precision parameter capturing sensitivity to utility differences, and U is the utility function. Importantly, the quantal choice model predicts uniform, equiprobable choices when utility differences are small or when precision is low ($\lambda \rightarrow 0$). In contrast, Proposition 3 implies that in repeated environments, choice probabilities depend on subjective beliefs shaped by experienced event frequencies. Even when utility differences are small, choices are expected to shift towards the more likely event, resulting in probability matching rather than equiprobability. This leads to Hypothesis 1.

Hypothesis 1. Under small utility differences, subjects' choice frequencies approximate equiprobability in DFD, but approximate event probabilities in DFE and DNV.

To explore the cognitive mechanisms potentially underlying CRAE and standard discrete choice models, we examine neural activity, focusing on parietal P3 amplitudes and beta-band oscillations. We first consider the parietal P3 component of event-related potentials (ERP), which has frequently been associated with cognitive load and explicit numerical evaluation (Polich, 2007; Wang et al., 2015). In value-based decision making, P3 amplitude has been interpreted as reflecting the strength of accumulated decision evidence, with more difficult decisions often exhibiting attenuated amplitudes due to weaker differentiation between alternatives (Twomey et al., 2015; O'Connell et al., 2012).

Beta-band activity has been associated with action selection, deliberative control, and response selection in incentive-based choice (Meyniel and Pessiglione, 2014; Pornpattananangkul et al., 2019). In standard discrete choice models, including Luce and multinomial logit formulations, stochastic choice is commonly interpreted through evidence accumulation processes formalized by drift diffusion models (DDM; Krajbich et al., 2014). Within this framework, task difficulty primarily affects the rate of evidence accumulation, leading to longer response times when utility differences are small, while the underlying decision policy remains unchanged. Consistent with this interpretation, beta activity during the decision phase has been linked to decision thresholds and response caution rather than evidence accumulation rates (Solway et al., 2022; Kirschner et al., 2024). Consequently, canonical evidence-accumulation accounts do not naturally predict substantial modulation of beta activity across utility differences, even

when difficult decisions require more time to reach a choice. Supporting this interpretation, Pornpattananangkul et al. (2019) report similar beta suppression across easy and difficult gamble decisions.

CRAE suggests a different possibility. In repeated decision environments, task difficulty may influence not only the efficiency of decision making but also the strategy used to make choices. When utility differences are small, the expected gains from consistently selecting the superior option are limited, potentially increasing reliance on learned event frequencies and resulting in probability-matching behavior (Schulze and Newell, 2016b). As utility differences increase, choices may increasingly favor direct exploitation of the dominant option (Daw et al., 2005; Wyatt et al., 2024). If beta activity indexes the degree of deliberative control engaged during choice, then beta suppression may vary systematically with task difficulty in environments where CRAE predicts such strategic adaptation. Although this pattern would not uniquely identify probability matching or any specific decision strategy, it would be consistent with the greater engagement of deliberative processes under difficult choices implied by CRAE.

Hypothesis 2. Beta-band suppression during the decision phase is more strongly modulated by task difficulty in DFE and DNV than in DFD.

Second, most existing models of newsvendor decisions, particularly utility-based approaches, assume independence across choices, such that each decision is evaluated separately and does not depend on past outcomes or actions (Regenwetter and Davis-Stober, 2018; Wilcox, 2024). This assumption is consistent with one-shot DFD tasks, but does not reflect the repeated nature of DFE and DNV environments. Because these models do not incorporate learning or memory, specifically, how realized demand and experienced rewards shape beliefs and subsequent choices, they cannot account for temporal dependence in ordering behavior. Consequently, they do not predict systematic adjustments in response to recent demand, such as the demand-chasing effect widely observed in newsvendor experiments (Schweitzer and Cachon, 2000; Bolton and Katok, 2008). By contrast, in a stationary repeated environment, the CRAE model allows subjective beliefs to evolve through experience, causing past demand realizations to influence subsequent choices. This mechanism generates demand chasing as a natural consequence of adaptive learning (Proposition 5).

Hypothesis 3. Order quantities are positively correlated with the most recent demand realization in DFE and DNV, but are independent of past demand realizations in DFD.

The learning mechanism can also be examined from a neural perspective. Learning is central to DFE, but plays a more limited role in DFD. As noted by Erev et al. (2010), standard DFD tasks involve one-shot decisions in which task parameters change across trials, rendering memory of past outcomes largely irrelevant. In contrast, DFE involves repeated choices in a stationary environment, requiring decision makers to learn from feedback. If DNV were cognitively equivalent to DFD simply because descriptive information is provided, as assumed by static utility-based theories, neural activity in DNV should exhibit the same “reset” pattern observed in DFD. However, DNV shares the stationary and repeated structure of DFE, which allows learning from realized demand. Accordingly, the CRAE model predicts that DNV involves continuous belief updating driven by prediction errors. Unlike the non-stationary DFD paradigm, this process should recruit working memory to encode and update beliefs across trials. To evaluate this prediction, we examine frontal P2 amplitudes in event-related potentials and frontal gamma-band activity in time-frequency analyses, both of which have been linked to working memory processes, particularly memory encoding and updating (Chapman et al., 1978; Roux et al., 2012). Accordingly, we predict distinct neural signatures reflecting this learning process:

Hypothesis 4. Frontal P2 amplitudes and frontal gamma-band enhancement during the feedback phase are larger in DFE and DNV than in the non-repeated task (DFD).

4. Study 1: Binary demand with EEG

4.1. Subjects

We recruited 46 university students (26 females) ranging in age from 19 to 36 (mean = 22.7, s.d. = 4.73). Previous studies did not find a significant difference between students and practitioners in terms of newsvendor decisions (Bolton et al., 2012). The EEG data of four subjects were partly missing due to software malfunction during the experiment, therefore they were excluded from this study, and the resulting $n = 42$. The sample size is similar to previous studies that adopted the EEG technique. Each subject received a base participation fee of 7 USD plus their total payoff received in the experiment. All subjects were right-handed and did not report any history of neurological or mental illness. The study protocol was approved by the ethics committee of the first author's institution and adhered to the legal requirements of the country where the experiment was conducted.

4.2. Experiment design

The experiment consisted of three tasks: DFD, DFE, and DNV. A within-subjects design was employed, requiring each participant to complete all three tasks. At the start of the experiment, participants received an instruction page that outlined: (a) the newsvendor context, in which they assumed the role of a retailer responsible for ordering inventory under uncertain demand; (b) the parameter settings, which either remained constant (in DFE and DNV) or varied between trials (in DFD); (c) the demand values and admissible responses (1 and 2); and (d) the payoff calculation method. The complete instructions are provided in Section A.3 of the Online Appendix.

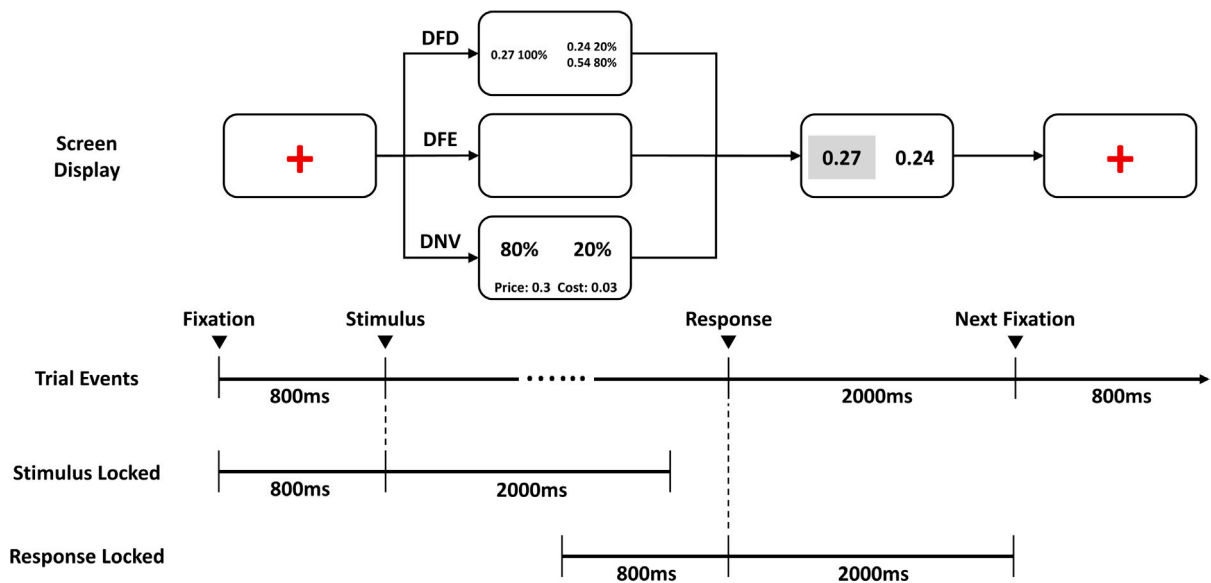


Fig. 1. The procedure of experiment.

Before starting each task, subjects were shown a brief page with the key information and sample screenshots to help them familiarize themselves with the task. Additionally, prior to each task, subjects had the option to practice in blocks of 5 trials, and they could continue practicing until they felt comfortable to proceed with the formal experiment.² This practice phase ensured that subjects were fully familiar with the task and its requirements. The order of the three tasks was randomized for each subject, and subjects were allowed a short break between tasks.

Due to the adoption of the EEG technique, each trial contained three phases: a fixation period, a stimulation period, and a feedback period. Each trial started with the fixation which is a red cross in the center of the screen. The fixation lasted for 800 ms and cannot be expedited. It is used as (a) a cue for subjects to focus on the screen, and (b) the baseline for correction in EEG analysis. The length of 800 ms is sufficient for these two purposes.

After the fixation, the stimulus was onset, which was used to present the experimental materials and for the subjects to make the decision. The stimulus differed by each task. In DFD tasks, the stimulus was the exact payoff distribution of the two choices. In DFE tasks, the stimulus did not contain any information about the choices. In DNV tasks, the stimulus contained the distribution of the demand, as well as the price and cost values, consistent with most previous newsvendor studies.

Our study involves complex and high-level cognitive decision-making, so it is not appropriate to give a time limit for the response. However, a long decision time will affect the quality of EEG data due to extra movement of the eyes and the body. It also affects the volume of the data, which linearly increases with the experiment duration. In many decision-making experiments that involve EEG, the stimulating interface is presented for less than 4 s (e.g., Han et al., 2019; Shen et al., 2022). Thus, in the experiment brief, we prompted the subjects to produce their decisions in 4 s but allowed them to take more time in the actual decision. Due to the restrictions imposed by EEG, subjects, when making decisions, were not allowed to review the history of decisions and outcomes, nor do calculations on paper.

After the subject made the response, the feedback was immediately shown with both the received and the foregone payoffs. This is to be consistent with the previous newsvendor experiments, where the payoff of all possible choices can be seen post-response. The feedback lasted for 2000 ms, sufficient for memory encoding to take place, while not too long to induce distraction. The consideration of data quality and data volume was also applied here. After the feedback phase, the fixation for the next trial was shown. We focus on stimulus-locked and response-locked phases with the notion that the cognitive tasks in these two phases can be differentiated. The stimulus-locked phase represents the decision process where the stimulus is the available information for the decision; the response-locked phase represents the post-response process where the stimuli are the realized demand and the earned and foregone payoffs. Fig. 1 shows the procedure of this experiment.

4.3. Experiment setting

The experiment varied the values of m , x , and p . The m parameter (profit margin) can be either 0.1 or 0.9. We label the setting of $m = 0.1$ as “mixed” (M) since the payoff can either be negative or positive under this treatment. Similarly, $m = 0.9$ is labeled as

² We directly tested whether the length of practice influenced task performance and no significant effect was observed.

Table 2
Parameters of the 2×2 treatments.

Treatment	m	x_1	r^*	U_δ	Task difficulty
GH	0.9	0.2	2	$0.7p$	Easy
ML	0.1	0.8	1	$0.7p$	Easy
GL	0.9	0.8	2	$0.1p$	Hard
MH	0.1	0.2	1	$0.1p$	Hard

Table 3
Summary of the experiment tasks.

Task	Description	Available Information	# Settings	# Trials
DFD	Trials in non-repeated settings	Payoff distribution	36	36×42
DFE	Trials in repeated settings	None	12	40×42
DNV	Trials in repeated settings	Demand distribution, price and cost	12	40×42

“gain” (G), as both admissible payoffs will be positive. The unattractive payoff $V(1, 2) = p - 2c = 2m - 1$, which is negative in M treatments and positive in G treatments. The demand probability x_1 can be either 0.2 or 0.8, which are labeled as “high” (H) and “low” (L) respectively, to represent high demand or low demand being the dominant event. Thus we label the four treatments as GH, ML, GL, and MH. The absolute difference between choice utilities, $U_\delta = |U(1) - U(2)|$, is either $0.1p$ (in GL and MH) or $0.7p$ (in GH and ML). The U_δ value defines the difficulty of the decision task. When $U_\delta = 0.1p$, it is “hard” for subjects to differentiate the utility of the choices; whereas when $U_\delta = 0.7p$, it is “easy” to do so. The task difficulty can also be understood as whether the optimality coincides with the demand mode. In GH and ML this is the case. Table 2 details the parameters used in the 2×2 treatments where the profit margin and demand distribution vary.

One difficulty in designing this experiment is that the treatments between trials should not repeat in DFD, but should repeat in DFE and DNV. This was dealt with by varying the price parameter p . In DFD tasks, p took nine possible values from 0.1 to 0.5 with a step of 0.05. This created $2 \times 2 \times 9 = 36$ parameter settings. In DFE and DNV tasks, p can take 0.15, 0.3, or 0.45, but it did not change between trials for each subject. This design created $2 \times 2 \times 3 = 12$ settings. We note that while the standard quantal choice model (e.g., logit) predicts that varying p should affect rationality through changes in absolute payoff differences, our experimental design held the cost-to-price ratio (c/p) constant while varying p . The fixed relative difference, which, according to models sensitive to relative comparisons (e.g., the Weibit model, Gu et al., 2022), may attenuate the effect of absolute price levels on choice consistency. An ANOVA analysis confirmed that the price value did not induce a significant difference in the behavior.

Consequently, each subject completed 36 trials in DFD with non-repeated settings, and 40 trials in DFE and DNV with repeated settings. This ensures the comparable sizes of the total number of trials between the three tasks. A similar design can also be found in previous research to compare DFD and DFE tasks, where the parameter settings used in DFE were a subset of those used in DFD, and each subject conducted more trials in DFE than in DFD (Erev et al., 2010).

Subjects were randomly assigned to the $\{m, x, p\}$ settings in both DFE and DNV. The G and M treatments were balanced across experiments; for example, a subject assigned to GH in DFE was also assigned to ML in DNV. This procedure controlled for payoff asymmetries, as the expected payoff in the Mixed treatments was substantially lower than in the Gain treatments. In DFE, the number of subjects in each condition was $n = 11, 12, 9, 10$ for GH, ML, GL, and MH, respectively. In DNV, the distribution was $n = 12, 11, 10, 9$. The p values were randomly assigned to subjects, and the corresponding c values were adjusted to maintain a fixed profit margin. The demand series across the 40 trials was randomly generated based on the assigned x_1 value. The average duration of one experiment session was 15.2 min for each subject (s.d. = 3.75; including the time for practicing). The average payment to subjects was USD 19.45 (s.d. = 1.02). Table 3 summarizes the three experiment tasks.

4.4. Data collection, EEG preprocessing and analysis

The experiment was conducted with one subject each time. During the experiment, the subject was seated in a comfortable armchair in a soundproofed room. They were instructed to avoid blinking, saccade, and body movement as much as possible. They placed their right hand on the numeric keyboard, the index finger on Key 1, and the middle finger on Key 2. Responses were made by pressing these two buttons. Stimuli were presented on a 20-inch computer monitor located one meter in front of the subjects (with a 10° visual angle). The stimuli were controlled by E-Prime 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA) which also recorded the behavioral data. It includes the demand, response, response time, and parameter values for each trial. The EEG data was recorded simultaneously using a 64-channel HydroCel Geodesic Sensor Net, Net Amps 300 amplifier, and Net Station software (Version 5.4, Electrical Geodesics Inc., Eugene, Oregon, USA).

The details of EEG data collection, preprocessing, and analysis can be found in Section A.4 of the Online Appendix.

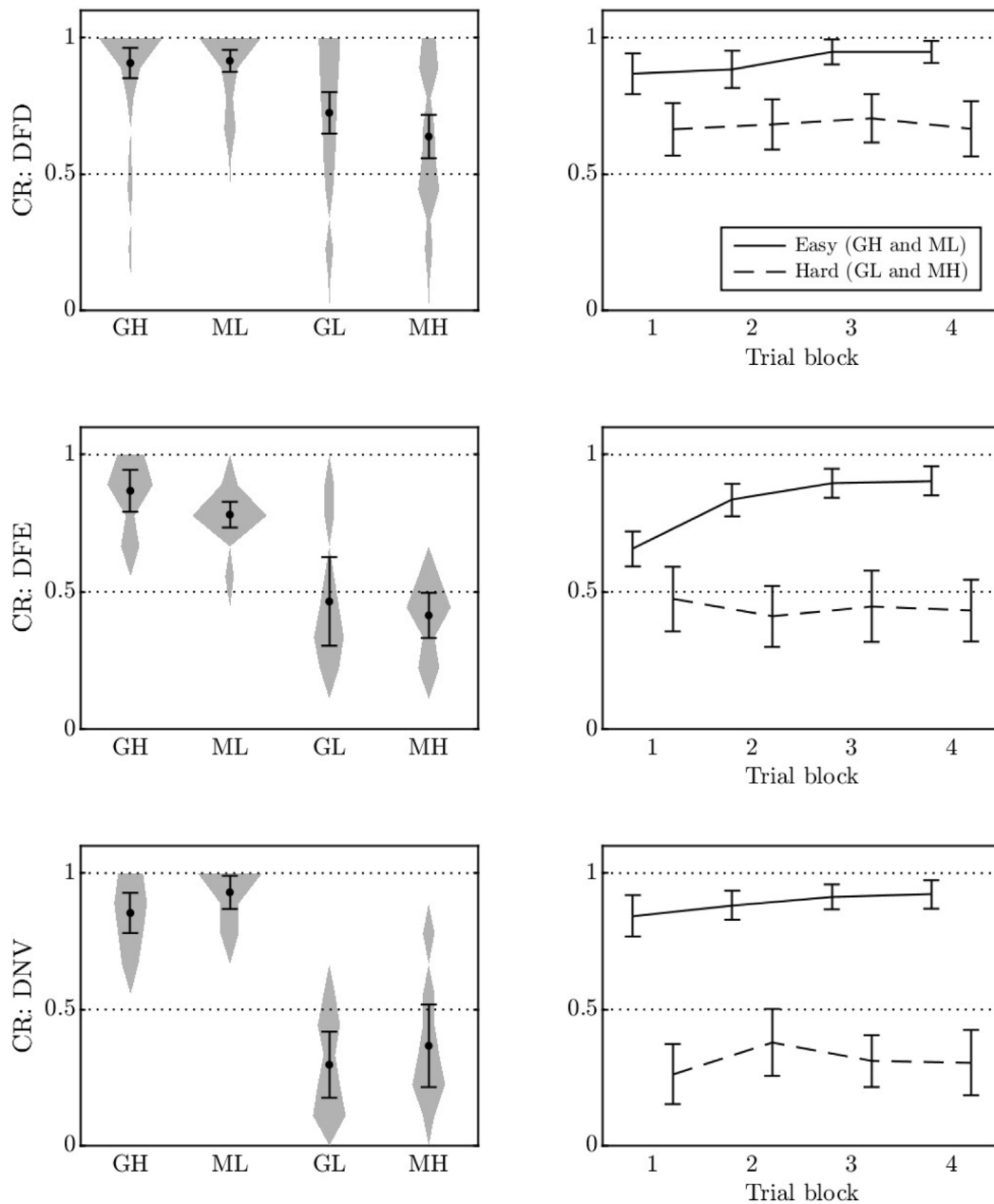


Fig. 2. Correctness ratio between tasks and treatments. Error bars: 95% CI. Shaded area: histogram.

5. Result of study 1

5.1. Evidence of newsvendor bias

Subjects' decision behavior is characterized by correctness ratio (CR), response time, and odds ratio. The correctness ratio measures the percentage of correct or optimal responses, which varies between zero and one.³ Response time measures the time

³ CR is directly linked to the commonly used measure of the pull-to-center (PTC) effect, defined as $PTC = (\bar{r} - \bar{d}) / (r^* - \bar{d})$, which captures the extent to which responses deviate from optimality towards the mean demand. The relationship between PTC and CR can be expressed as follows: if $U_\delta = 0.7p$, then $CR = 1 - 0.2PTC$; if $U_\delta = 0.1p$, then $CR = 1 - 0.8PTC$.

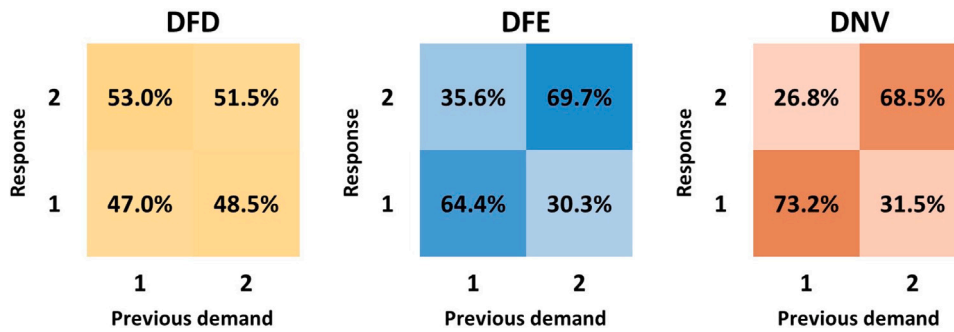


Fig. 3. Frequency between previous demand and latter response. Color lightness corresponds to cell values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

difference from the onset of the stimulating interface to the subjects' response. It serves as an important indicator of the underlying cognitive process (e.g. Clithero, 2018). The odds ratio, defined as

$$OR = \frac{z_{r,t}|(d_{t-1} = r) \cdot z_{r',t}|(d_{t-1} = r')}{z_{r,t}|(d_{t-1} = r') \cdot z_{r',t}|(d_{t-1} = r)}$$

measures the correlation between choice and preceding demand. We use the odds ratio to measure the demand-chasing effect since both demand and order quantity are categorical. For more discussions on measuring the demand-chasing effect in continuous cases, we refer interested readers to Lau and Bearden (2013) and Kirshner and Moritz (2020, 2021). A summary of the above statistics can be found in Table A.2 in the Online Appendix.

Fig. 2 shows the between-task and between-treatment comparisons. First, it is evident that task difficulty has a negative impact on correctness across all three tasks. In all paradigms, the difference in CR between easy and hard treatments is significant. Second, in easy treatments, there is CR improvement across trial blocks, which reaches about the same level in the fourth block. Specifically, there is a notable upward trend in DFE. In DFD and DNV, such trend is mild. The CR improvement is not observed in hard treatments. Third, in the hard treatments, correctness in DFE and DNV is significantly lower than in DFD. While choice probabilities in DFD are pulled towards equiprobability from the optimal benchmark, resulting in reduced but still above-chance correctness, the pull-over-uniform pattern emerges in DFE and is more pronounced in DNV. In these tasks, correctness falls below 0.5, indicating decision accuracy worse than uniform random choice, consistent with the probability matching prediction of Hypothesis 1. This leads to the overall performance in DNV to be significantly lower than DFD. Lastly, the *order asymmetry* can be observed in DFE and DNV. Under the same level of difficulty, in DFE, accuracy in G settings (where the optimal response is 2) is higher than in M settings (where the optimal response is 1). In DNV, accuracy in M settings is higher than in G settings. Such asymmetry is more significant in easy treatments.

Fig. 3 illustrates the correlation between the response and the preceding demand across the three decision tasks. There is a strong *demand-chasing* effect in DFE and DNV, but not in DFD. The odds ratios in DFD, DFE, and DNV tasks are 0.94, 4.15 and 5.94 respectively. All differences are significant on the 0.05 level. The prominent correlation between demand and choice strongly supports Hypothesis 3.

Fig. 4 shows the response time. We exclude one more subject here because the response time of this subject in the DFE task was not recorded, while their EEG data and other behavioral data are recorded. The response time in DFD and DFE are steadily high and low across trials, averaging at 5.01 s and 0.97 s respectively. For DNV, there is a significant decrease in response time from 3.58 s to 1.25 s. These results provide further insight into the cognitive processes underlying the three decision paradigms. In DFD, subjects rely on deliberate reasoning to calculate expected payoffs on each trial. In DFE, decisions are guided primarily by rapid sampling and heuristic strategies. In DNV, subjects initially attempt effortful computation of the optimal choice while sampling but eventually revert to heuristic-based decision-making. The CRAE model accounts for this interplay by incorporating both sampling and heuristic components.

We compare the maximum likelihood of the proposed model with four benchmark models from the literature, which were proposed to explain the newsvendor biases. They include adaptive learning (AL, Bostian et al., 2008), minimizing inventory error (MIE, Schweitzer and Cachon, 2000), reference dependence (RD, Ho et al., 2010), and prospect theory (PT, and its two variants, PT-R0 and PT-RV, Uppari and Hasija, 2019). Details of these models can be found in Section A.6 of the Online Appendix. Cross-validation is employed to prevent overfitting, and the particle swarm optimization algorithm is used to estimate the model parameters. For each task, the first three quarters of trials (i.e., 27 for DFD and 30 for DFE and DNV) serve as in-sample data for parameter estimation. The fitted parameters are then applied to the remaining one quarter of trials (out-of-sample data) to compute the maximum likelihood, which is subsequently converted into the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to account for model parsimony. Table 4 shows the minimum negative logarithmic likelihood, AIC, and BIC values of each model in the three tasks.

For the DFD task, as anticipated, the prospect theoretic model (with reference being zero) performs better than any other candidate. This is consistent with various previous literature that demonstrates the rigor of prospect theory in such decision tasks

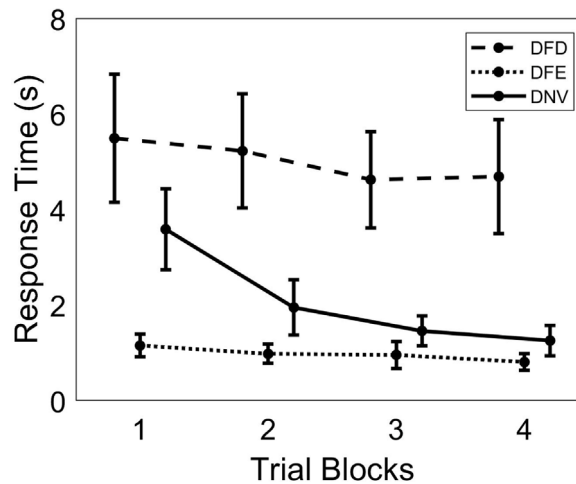


Fig. 4. Subjects' response time in trial blocks. Error bars: 95% CI.

Table 4

Fitness of the candidate models in Study 1.

Model type	Model	DFD			DFE			DNV		
		–log L	AIC	BIC	–log L	AIC	BIC	–log L	AIC	BIC
Utility-based	MIE	166.23	<u>336.46</u>	<u>346.53</u>	216.52	437.03	447.31	183.49	370.97	381.25
	RD	166.23	338.46	353.56	216.96	439.91	455.33	182.13	370.26	385.68
	PT	165.97	339.94	360.07	222.90	453.80	474.36	207.02	422.04	442.60
	PT-R0	<u>165.14</u>	338.28	358.41	216.84	441.68	462.24	198.68	405.35	425.91
	PT-RV	165.93	341.86	367.03	227.04	464.09	489.78	195.57	401.14	426.84
Learning-based	AL	262.01	530.02	545.12	225.91	457.82	473.24	214.89	435.77	451.19
	CRAE	196.53	399.07	414.17	<u>208.73</u>	<u>423.46</u>	<u>438.88</u>	<u>169.79</u>	<u>345.58</u>	<u>361.00</u>

(Erev et al., 2010; Ruggeri et al., 2020). MIE and RD also perform well in terms of AIC and BIC, due to model parsimony. Note that they all belong to the utility-based category where learning is not involved. In DFE and DNV, the CRAE model provides a substantially better fit than the benchmark models. This finding supports the view that belief updating and probability matching are central to explaining behavior in repeated decision environments with feedback.

In the next two sections, we present behavioral and EEG results that further evidence the underlying cognitive process in DNV.

5.2. Contingent rationality

Fig. 5 presents the ERP in the stimulus-locked phase in the parietal area. A significant difference in P3 triggered by information availability is observed (One-way ANOVA, $p < 0.05$, $\eta^2 = 0.370$), with DFD exhibiting the highest P3, followed by DNV and DFE. Furthermore, the ERSP analysis (Fig. 6) reveals significant beta suppression in DFD in both easy and hard treatments. However, in DFE and DNV, beta suppression is more pronounced in hard than in easy treatments. The main effect is significant to task difficulty (two-way ANOVA, $p < 0.001$), information availability ($p < 0.001$) and their interaction ($p < 0.001$).

The EEG results strongly support Hypothesis 2. In DFE and DNV, the P3 component is more attenuated than in DFD (reflecting lower evidence accumulation rates), and unlike DFD, task difficulty significantly affects beta suppression in DFE and DNV. Specifically, greater brain activity (stronger beta suppression) is associated with hard decisions in DFE and DNV. Thus, the seemingly “irrational” choices in these environments are accompanied by active, resource-consuming cognitive processes rather than by reduced deliberation. This pattern contradicts the prediction of standard utility-based diffusion models (DDM), which imply stable beta activity across difficulty levels. Instead, it is consistent with the CRAE framework, suggesting that when utility differences are small, subjects actively implement probability-matching strategies.

Could subjects be engaging in the maximizing strategy in hard treatments of DNV? Recall that the correctness ratio in the hard treatment of DNV is significantly lower than that of DFD and shows no signs of improvement (Fig. 2). If the maximizing strategy were applied, we would not expect to observe such a substantial difference in the correctness ratio, particularly when subjective utilities are adequately updated. Additionally, there is a significant difference in response time between easy and hard treatments in DFD (4.51 s vs. 5.50 s, $p < 0.001$), indicating that subjects adopting a maximization strategy expend greater deliberative effort when the decision is difficult. However, no such difference is observed in DNV (2.13 s vs. 1.98 s, $p = 0.29$). Given that the only difference between DFD and DNV is information availability, these findings strongly support the probability-matching account proposed by CRAE.

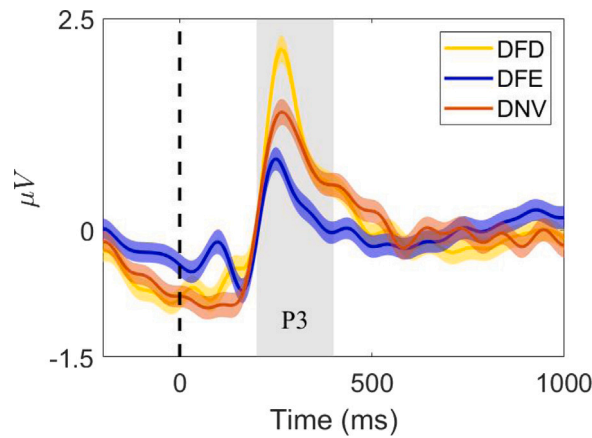


Fig. 5. The ERP plot in the stimulus-locked phase in the parietal area. Electrodes: Pz, PO3, PO4, POz. Shades: 95% CI. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

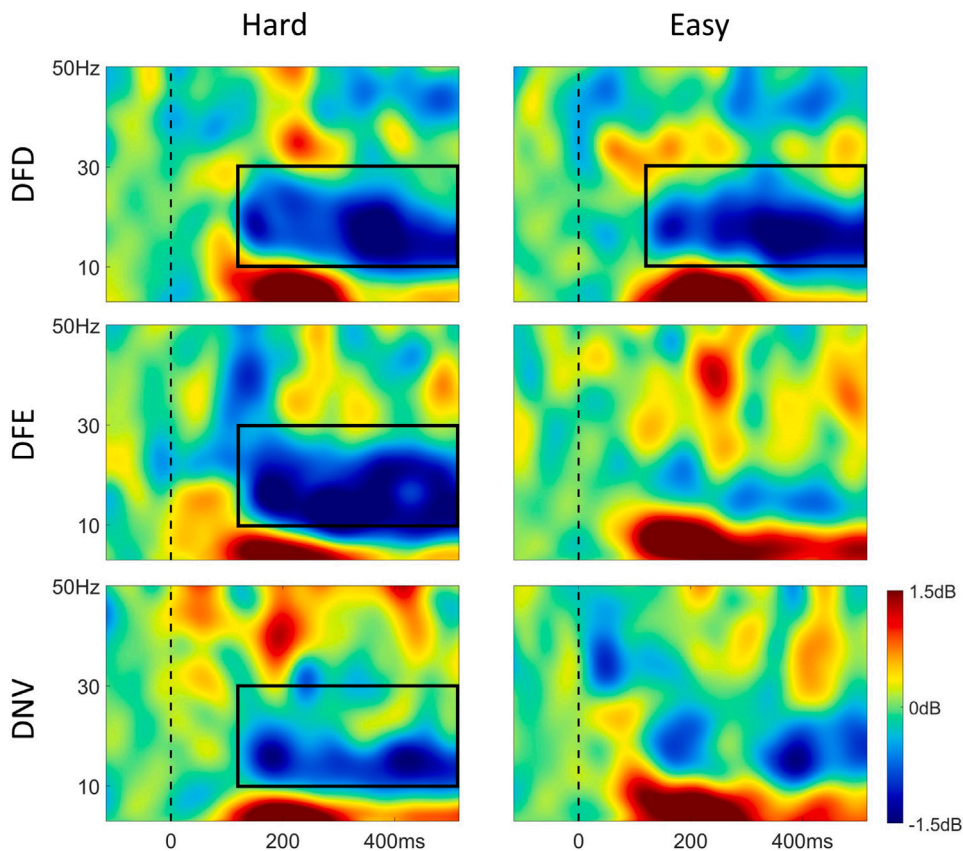


Fig. 6. The ERSF plot in the stimulus-locked phase in the parietal area. Electrodes: Pz, PO3, PO4, POz. Rectangle: 100–500 ms, 10–30 Hz. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The fitted parameter values of DFE and DNV also show the effect of information availability, task difficulty, and experience. The salience of demand distribution has a negative effect on rationality. On the task level, the rationality parameter $\lambda_{DFE} = 3.20$ and $\lambda_{DNV} = 2.88$, suggesting that subjects are more inclined towards utility-maximization in DFE than in DNV. Subjects also behaved more rationally towards the end of the experiment, with reduced response time. When fitting the CRAE model to the first and the

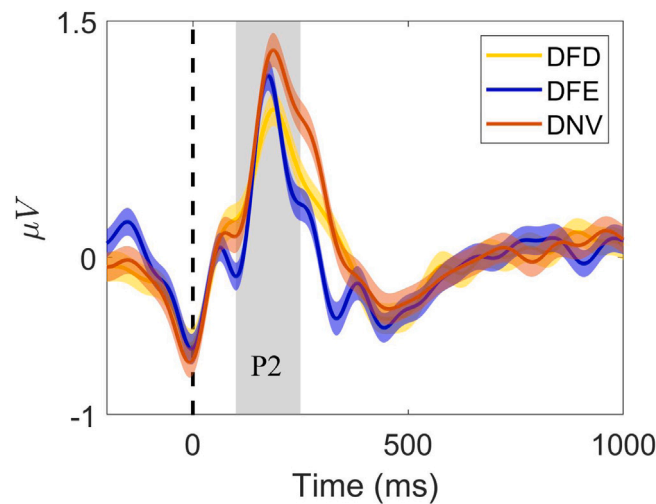


Fig. 7. The ERP plot in the response-locked phase in the frontal area. Electrodes: Fz, F1, F2, AFz. Shades: 95% CI. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

last two blocks, we derive the corresponding λ values to be 2.17 and 4.20 in DFE and 1.41 and 4.31 in DNV.⁴ Meanwhile, the decrease in response time across trials can be seen in Fig. 4. This suggests the accumulation of information and experience can alleviate the speed-accuracy trade-off, enabling higher decision accuracy.

5.3. Asymmetric learning

Our focus now shifts to the cognitive process in the response-locked phase, occurring after the subjects have made their responses and received feedback.

The ERP shows significant differences in the amplitude of P2 among the tasks (one-way ANOVA, $p < 0.05$, $\eta^2 = 0.023$). Specifically, stronger frontal P2 components are evident in DNV and DFE compared to DFD. In ERSP, we observe gamma enhancement in DFE and DNV but not in DFD. The gamma enhancement in the frontal area is sensitive to the preceding demand realization. Particularly, strong gamma enhancement is noted in DFE when the demand is 2, and in DNV when the demand is 1, but not in DFD ($p < 0.005$). A two-way ANOVA revealed a significant main effect of the information availability ($p < 0.001$), preceding demand realization ($p < 0.001$), and their interaction ($p < 0.001$) (see Figs. 7 and 8).

The EEG results strongly support Hypothesis 4. In DFD there is a milder P2 and a lack of gamma enhancement, indicating that probability learning mechanisms are not engaged. In contrast, in DFE and DNV, we observe a much stronger P2 and gamma enhancement, with the magnitude of gamma enhancement being influenced by previous demand realizations.⁵ Notably, the asymmetry in gamma enhancement, order quantity, and fitted α values align in the same direction (in DFE, $\alpha_1 = 0.08$, $\alpha_2 = 0.09$; in DNV, $\alpha_1 = 0.04$, $\alpha_2 = 0.03$). The significant demand-chasing effect in DFE and DNV (Fig. 3) serves as additional behavioral evidence of subjective probability update. On the other hand, the less significant correctness improvement in DNV can be attributed to subjects' awareness of the demand distribution, leading to stable order quantities.

To further test the significance of asymmetry in learning rate, we compare the performance of the CRAE model (full model) and its nested model where the learning parameters $\alpha_1 = \alpha_2$ in DFE and DNV. To test the null hypothesis that $\alpha_1 = \alpha_2$, we calculate $\chi^2_{\text{DFE}} = 2(\log L_{\text{full, DFE}} - \log L_{\text{nested, DFE}}) = 6.91$, and $\chi^2_{\text{DNV}} = 76.03$. Since the full model has one additional parameter, the test statistic follows a χ^2 -distribution with one degree of freedom, with the 95% critical value $\chi^2(0.95; 1) = 3.84$. Since χ^2_{DFE} and χ^2_{DNV} exceed this critical value, we reject the null hypothesis, indicating that the asymmetry in encoding ($\alpha_1 \neq \alpha_2$) is significant in DFE and DNV.

Subjects exhibit faster learning in DFE compared to DNV, as indicated by the rapid improvement in correctness (Fig. 2), and greater α values. Similarly, D'Urso et al. (2017) also found that when the demand information is not known to the subjects, the demand-chasing behavior is more prominent. This, coupled with higher levels of rationality, elucidates why subjects overall perform better in DFE than DNV. Also note that the odds ratio in DNV is higher than DFE, which might be due to the more stable response patterns in DNV.

⁴ The higher λ value in DNV compared to DFE for the last two blocks is likely to be caused by the parameter optimization algorithm, as the probability matching behavior is more prominent in DNV than DFE across all blocks.

⁵ It should be noted that the absence of, or a weak gamma enhancement in DFE and DNV does not imply that the memory encoding process was absent. The observed difference is best interpreted as stronger engagement of the process between the two scenarios, rather than its nonexistence (Poldrack, 2006).

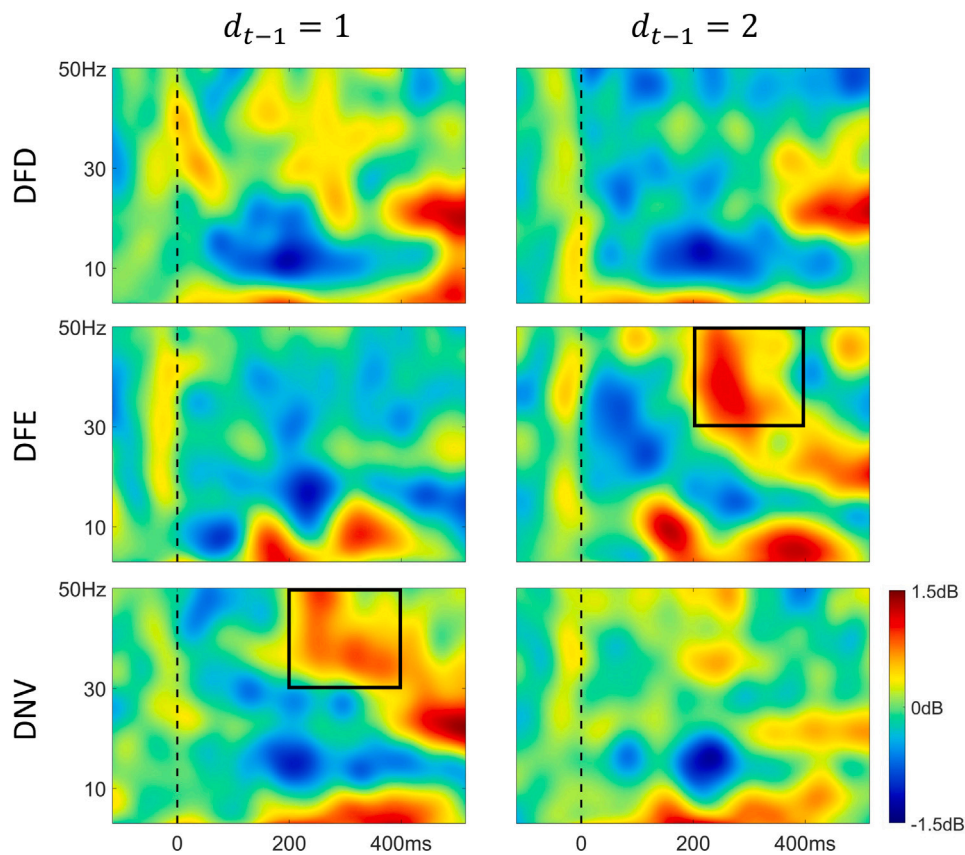


Fig. 8. The ERSP plot in the response-locked phase in the frontal-centrofrontal area. Electrodes: Fz, F1, F2, AFz. Rectangle: 200–400 ms, 30–50 Hz. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6. Study 2: Binary demand without EEG

Study 2 aims to investigate whether the findings observed in Study 1 are due to artifacts in the EEG data collection process. We designed an experiment mirroring the setting and procedures as Study 1 but excluded EEG measurements and the constraint on response time. Given our focus on understanding the mechanisms of newsvendor decisions, Study 2 only employs the DNV paradigm. The sample size of Study 2 ($n = 70$) was determined based on the effect size of model fit in Study 1 (individual maximum likelihood between the best and second best performing models, mean = 1.23, s.d. = 2.11). With the probability of Type I/II error being 0.01, the required sample size is 67. Subjects' ages range from 19 to 25 (mean = 20.9, s.d. = 1.16). The experiment was conducted using the E-Prime 3.0 platform.

Observations from Study 2 replicate those from Study 1. The phenomena of pull-to-center, demand chasing, and pull-over-uniform are evident (Fig. 9). In easy treatments (GH and ML), the correctness ratio remains high, ranging between 0.8 and 1, indicating a pull-to-center effect. Conversely, in hard treatments (GL and MH), the correctness ratio drops below 0.5, consistent with the notion of pull-over-uniform. With the odds ratio of 4.56, a strong demand-chasing effect is suggested.

Interestingly, the order asymmetry effect observed in Study 1 was not significant in Study 2. The difference in correctness ratio between ML and GH conditions was minimal ($p = 0.38$). To explore this discrepancy, we compared response times across the two studies. A two-tailed t -test revealed no significant difference, with Study 2 subjects responding slightly faster than those in Study 1 (1.65 s vs. 2.06 s, $p = 0.15$).

Because Study 2 subjects responded faster yet exhibited more symmetric learning, the presence of asymmetry in Study 1 cannot be attributed to insufficient deliberation time. One possible interpretation is that the difference reflects variation in *extraneous cognitive load* (Sweller et al., 1998). The EEG setup in Study 1 imposed physical constraints and monitoring requirements intended to minimize blinking and movement. Prior research indicates that such dual-task demands can consume attentional resources (Huxhold et al., 2006) which may in turn affect decision quality (Deck and Jahedi, 2015).

Importantly, the performance difference across studies was concentrated in difficult treatments. Performance in easy treatments remained high and comparable across studies (0.89 vs. 0.89, $p > 0.9$), whereas performance in hard treatments improved significantly in Study 2 relative to Study 1 (0.31 vs. 0.40, $p < 0.05$). This pattern is consistent with the possibility that the additional demand imposed by the EEG environment disproportionately affected difficult decisions, where cognitive resources may be more heavily

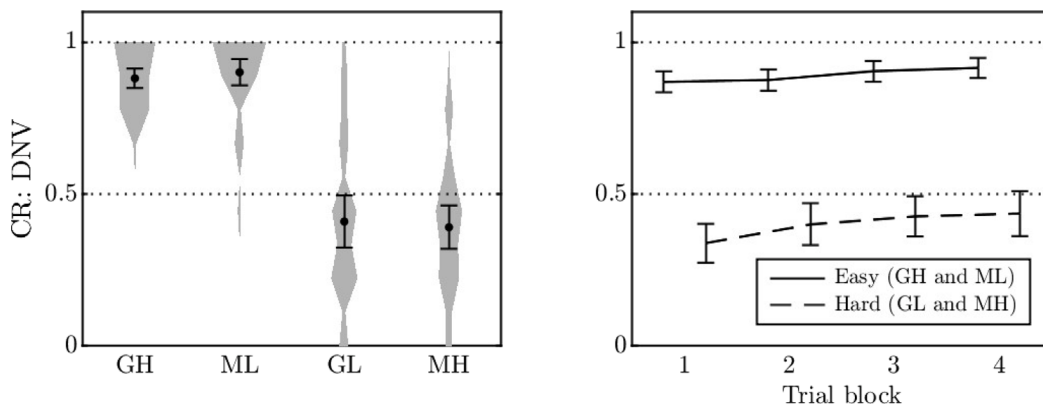


Fig. 9. Correctness ratio in Study 2. Error bars: 95% CI. Shaded area: histogram.

Table 5
Fitness of the candidate models in Study 2.

Model type	Model	DNV	AIC	BIC
		$-\log L$		
Utility-based	MIE	704.47	1412.94	1423.21
	RD	704.49	1414.98	1430.39
	PT	728.64	1465.28	1485.83
	PT-RO	703.29	1414.58	1435.13
	PT-RV	702.71	1415.42	1441.11
Learning-based	AL	735.54	1477.08	1492.50
	CRAE	694.43	1394.86	1410.27

taxed (Pocheptsova et al., 2009). Under this interpretation, subjects in Study 1 may have relied more heavily on simplifying or conservative decision rules when faced with difficult choices. Conversely, the absence of these additional demands in Study 2 may have allowed participants to devote greater attention to the task, contributing to higher accuracy and a reduction in the observed order asymmetry.

As shown in Table 5, the model-fitting results indicate that CRAE remains the best-performing model. This finding further supports the notion that the core cognitive mechanisms of learning and probability matching are robust, even in the more deliberative context of Study 2. The model's consistent performance across studies, despite differences in time pressure, highlights its reliability and its ability to capture variations in decision-making strategies under varying conditions.

7. Study 3: Exponentially distributed demand

Study 3 aims to replicate the findings of Studies 1 and 2 in a multi-alternative scenario. Since Studies 1 and 2 suggest that pull-over-uniform occurs when the demand distribution is heavily skewed and the optimal choice is distanced from the demand mode, we designed an experiment with an exponentially distributed demand process. Both demand and order quantity are constrained to integers within the range of $[0, 100]$. In the high-margin-right-skewed treatment (analogous to GL in Study 1), the demand d follows $\text{Exp}(30)$, yielding an average demand of 30. The profit margin is set at 0.9 (price = 10, cost = 1), and the optimal order quantity is approximately 70. Conversely, in the low-margin-left-skewed treatment (analogous to MH in Study 1), $100 - d$ follows $\text{Exp}(30)$, resulting in an average demand of 70. Here, the profit margin is set at 0.1 (price = 10, cost = 9), positioning the optimal order quantity at 30.

We recruited 102 subjects aged between 18 and 22 (mean = 19.5, s.d. = 0.85, 49% female), and each subject was asked to perform the two tasks described above. The sequence of the tasks was counterbalanced to mitigate order effects. The demand series in each treatment was predetermined and subjects were not aware of it. Similar to the DNV paradigm in Studies 1 and 2, subjects were informed about the demand distribution, price, and cost. They were asked to input their order quantity in a dialog box. Following the decision, their selected order quantity, actual demand, and received payoff were displayed on the feedback page. Each task consisted of 40 sequential trials, resulting in a total of $102 \times 40 \times 2 = 9160$ trials. In rare instances where subjects failed to input a number (0.3% of trials), we assume they made a decision of 50 (the average random choice). The experiment was conducted using the E-Prime 3.0 platform. Fig. 10 displays screenshots of the experiment interface.

Fig. 11 illustrates the average order quantity and its variation across trial blocks. The pull-over-uniform effect is evident (Fig. 11a), where the average order quantity in both tasks is pulled over the uniform or choice by chance (50). However, this observation does not contradict the pull-to-center prediction, as the average order still falls between optimality and mean demand.

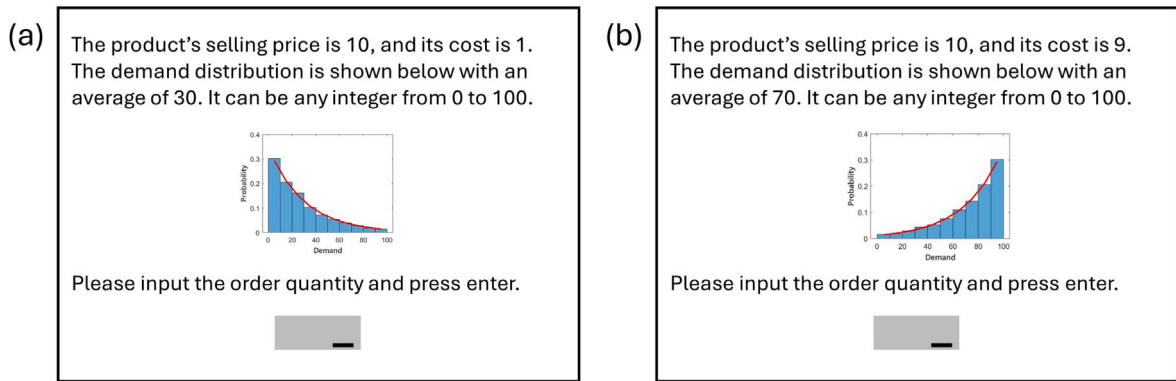


Fig. 10. Screenshot of Study 3. (a) high-margin-right-skewed; (b) low-margin-left-skewed.

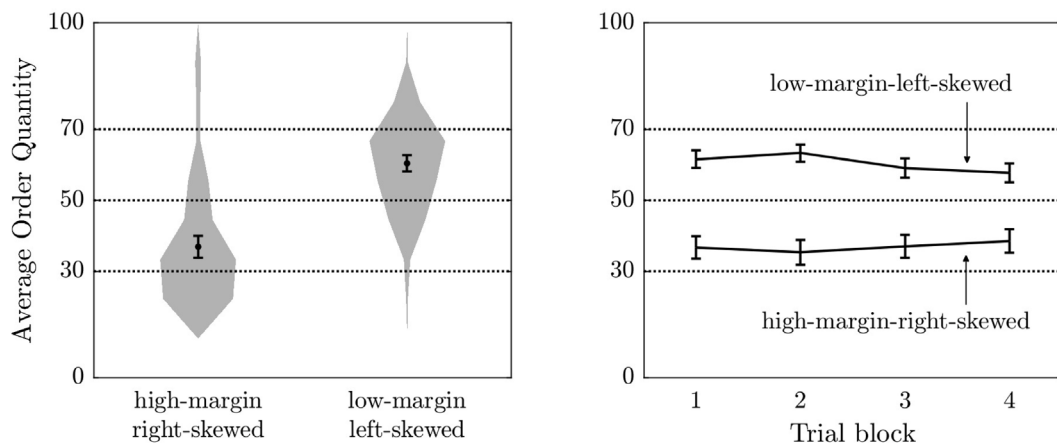


Fig. 11. Average order quantity under DNV paradigm and exponential demand. Error bar: 95% CI; Shaded area: histogram.

This underscores the significant impact of the demand distribution's shape and salience on order behavior. Although the average order quantity remains relatively stable across trials, there is a slight inclination towards optimality in both tasks (Fig. 11b).

In Study 3, we also observe order asymmetry, with the deviation from optimality in low-margin-left-skewed being higher than that in high-margin-right-skewed. This difference (2.69) is statistically significant at the trial level ($p < 0.001$). Additionally, demand chasing, measured by the correlation coefficient between order quantity and past demand, is evident. The subject-level mean correlation coefficient is 0.233, 95% CI = (0.197, 0.269), suggesting a positive correlation between order quantity and previous demand.

8. General discussion and conclusion

This study demonstrates that repetition and information availability fundamentally shape the cognitive mechanisms underlying newsvendor decisions. Comparing a one-shot, payoff-salient task (DFD) with repeated tasks in which payoffs are not explicitly salient (DFE and DNV), we show that choices track event probabilities and respond to recent outcomes, consistent with adaptive belief updating rather than direct maximization of expected utility. Observed asymmetries in adjustment indicate that learning differs across demand realizations, reflecting unequal weighting of experienced outcomes. These patterns are absent in one-shot decisions.

Neural evidence supports this interpretation. Feedback-related activity, including enhanced frontal P2 amplitudes and increased gamma-band responses, reflects greater engagement of memory encoding and learning processes in repeated tasks compared with one-shot decisions. Modulation of beta-band activity during the decision phase further indicates that decision strategies shift with task difficulty. Together, these behavioral and neural findings provide converging evidence that repeated decision environments recruit belief updating and memory processes, which systematically shape subsequent choices.

Below we offer discussions on the theoretical and practical implications of this study, as well as possible future research directions.

8.1. Theoretical implications

The CRAE model provides a unifying framework that connects to existing explanations of newsvendor behavior. First, decision-from-sampling accounts (Stewart et al., 2006) suggest that loss aversion and probability distortion, which are key assumptions in prospect theory, can emerge from sequential sampling and learning. Consistent with this, the CRAE model implies that some aspects of loss aversion and probability distortion observed in repeated decisions may reflect sampling and adaptive belief updating rather than stable preferences. Second, probability matching in CRAE produces systematic distortions in choice distributions: rare events are underweighted relative to their objective probabilities. This aligns with prior accounts of overconfidence and overprecision (Erev et al., 2010; Ren and Croson, 2013). Third, with respect to asymmetry effects, our findings support the existence of differential adjustments across conditions. The direction of asymmetry observed in DNV tasks is consistent with the majority of prior newsvendor studies (Zhang and Siemsen, 2019). Interestingly, DNV and DFE tasks exhibit opposite asymmetry despite identical experimental parameters. We propose that this asymmetry arises from uneven memory encoding, potentially linked to reference-dependent reward prediction errors (Rouhani et al., 2020). By integrating the mechanisms of probability matching, asymmetric learning, and memory-dependent adjustments, the CRAE model offers a comprehensive account of newsvendor decision behavior.

The two mechanisms highlighted in this paper, probability matching and asymmetric learning, are most clearly observed and formally modeled in the DNV paradigm. This paradigm can be generalized to a broader class of repeated economic decisions in which agents receive both probabilistic descriptions and outcome feedback, including inventory management, pricing, capacity planning, and financial allocation. By comparing DNV with DFD and DFE, we show that these mechanisms reflect general cognitive processes that operate in repeated, feedback-based environments, but are absent or attenuated in one-shot descriptive decisions. This interpretation is further supported by the close similarities between DFE and DNV in behavioral patterns, neural signatures, and model performance.

8.2. Practical implications

This research highlights the substantial impact of probability matching on suboptimal newsvendor decisions, particularly when the demand mode and the optimal order are widely separated. Under such conditions, subjects in Study 1 earned 8.6% less profit in DNV than in DFD. Such scenarios are common in practice, for example, high-margin products with right-skewed demand distributions. Probability matching leads decision makers to place orders too close to the demand mode, reducing expected profit. These findings underscore the need for strategies to mitigate probability-matching behavior.

First, probability matching can be exacerbated by a focus on short-term perfection. Prior research shows that subjects tasked with making perfect predictions for consecutive periods increase probability-matching tendency (Gao and Corter, 2015). Inventory contexts, where the supply-demand match is considered essential, intensify the pull-to-center effect relative to general decision tasks (Kremer et al., 2010). To counter this, incentives should emphasize long-term alignment of supply with demand rather than immediate accuracy.

Second, the type of information provided matters. Contrary to the intuition that more information improves performance, providing explicit demand distributions can induce probability matching and aggravate the pull-to-center effect. Our results suggest that emphasizing payoff information over demand information reduces reliance on heuristics and encourages more profit-aligned decisions, consistent with Ben Zion et al. (2010).

Finally, group decision-making may offer another avenue to mitigate probability matching (Gao and Corter, 2015; Schulze and Newell, 2016a). While replenishment decisions are often made individually in organizations, incorporating mechanisms such as internal review or collaborative forecasting could reduce individual biases. Future research should explore the practical feasibility and effectiveness of such interventions.

8.3. Limitations and future research directions

The primary experiment deviates from conventional newsvendor designs by employing a binary set of demand and choice. This simplification was necessary to maximize signal-to-noise ratio in the EEG data. While it does not directly mirror typical multi-alternative inventory settings, it offers theoretical advantages by isolating the core cognitive mechanisms underlying newsvendor bias and allows systematic manipulation of task difficulty. Although the CRAE model was not formally extended to multiple-alternative choices, its key mechanisms, probability matching and asymmetric belief updating, can be generalized with minor adjustments. Prior research suggests that increasing the number of options amplifies reliance on heuristics (Vulkan, 2000), which may enhance the applicability of CRAE in more complex choice environments.

This study assumes risk neutrality as the focus is on the interaction between randomization strategies and learning dynamics rather than on utility formation. As shown in Table 4, prospect-theoretic models incorporating risk aversion or loss aversion do not improve predictive performance. It is also well established that neither risk aversion nor loss aversion can account for the pull-to-center behavior, as decision makers should consistently lean towards the risk-free or loss-free side regardless of profit margin (Schweitzer and Cachon, 2000). EEG evidence further indicates that frontal brain activity associated with outcome evaluation does not differentiate between risk and loss (Figures A.3 and A.4, Online Appendix).

Future research could explore several directions. The behavioral model can be extended to multi-alternative scenarios and incorporate risk or loss preferences. Manipulating prior knowledge of the demand distribution or introducing ordering lead-times could further illuminate how learning and experience shape decisions. Finally, while practice trials and incentives were used to minimize confusion, the possibility of imperfect comprehension of the task instructions cannot be ruled out. Future studies could incorporate worked examples, comprehension checks, or post-experiment debriefings to ensure consistent understanding.

Replication statement

The replication package for this study, including the data, code, and supporting materials necessary to reproduce the reported results, is available at <https://doi.org/10.6084/m9.figshare.12937172>. The authors are responsible for the completeness and accuracy of the replication materials.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Weiwei Han reports financial support was provided by Ministry of Education of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by Humanities and Social Sciences Project, Ministry of Education of China, No. 24YJA630031.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jebo.2026.107776>.

Data availability

The data is available at <https://doi.org/10.6084/m9.figshare.12937172>.

References

- Allais, M., 1953. Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'école Américaine. *Econometrica* 21, 503–546.
- Arrow, K.J., Harris, T., Marschak, J., 1951. Optimal inventory policy. *Econ.: J. Econ. Soc.* 19 (3), 250–272.
- Aydogan, I., 2021. Prior beliefs and ambiguity attitudes in decision from experience. *Manag. Sci.* 67 (11).
- Barron, G., Erev, I., 2003. Small feedback-based decisions and their limited correspondence to description-based decisions. *J. Behav. Decis. Mak.* 16 (3), 215–233.
- Benzion, U., Cohen, Y., Peled, R., Shavit, T., 2008. Decision-making and the newsvendor problem: An experimental study. *J. Oper. Res. Soc.* 59 (9), 1281–1287.
- Benzion, U., Cohen, Y., Shavit, T., 2010. The newsvendor problem with unknown distribution. *J. Oper. Res. Soc.* 61 (6), 1022–1031.
- Bolton, G.E., Katok, E., 2008. Learning-by-doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manuf. Serv. Oper. Manag.* 10 (3), 519–538.
- Bolton, G.E., Ockenfels, A., Thonemann, U.W., 2012. Managers and students as newsvendors. *Manag. Sci.* 58 (12), 2225–2233.
- Bostian, A.A., Holt, C.A., Smith, A.M., 2008. Newsvendor “pull-to-center” effect: Adaptive learning in a laboratory experiment. *Manuf. Serv. Oper. Manag.* 10 (4), 590–608.
- Bush, R.R., Mosteller, F., 1955. *Stochastic Models for Learning*. John Wiley & Sons, Inc.
- Chapman, R.M., McCreary, J.W., Chapman, J.A., 1978. Short-term memory: The “storage” component of human brain responses predicts recall. *Science* 202 (4373), 1211–1214.
- Ciranka, S., Linde-Domingo, J., Padezhki, I., Wicharz, C., Wu, C.M., Spitzer, B., 2022. Asymmetric reinforcement learning facilitates human inference of transitive relations. *Nat. Hum. Behav.* 6 (4), 555–564.
- Cliethero, J.A., 2018. Response times in economics: Looking through the lens of sequential sampling models. *J. Econ. Psychol.* 69, 61–86.
- Daw, N.D., Kakade, S., Dayan, P., 2002. Opponent interactions between serotonin and dopamine. *Neural Netw.* 15 (4–6), 603–616.
- Daw, N.D., Niv, Y., Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neurosci.* 8 (12), 1704–1711.
- Deck, C., Jahedi, S., 2015. The effect of cognitive load on economic decision making: A survey and new experiments. *Eur. Econ. Rev.* 78, 97–119.
- Derks, P.L., Paclisanu, M.I., 1967. Simple strategies in binary prediction by children and adults. *J. Exp. Psychol.* 73 (2), 278.
- D'Urso, D., Di Mauro, C., Chiacchio, F., Compagno, L., 2017. A behavioural analysis of the newsvendor game: Anchoring and adjustment with and without demand information. *Comput. Ind. Eng.* 111, 552–562.
- Erev, I., Ert, E., Roth, A.E., Haruvy, E., Herzog, S.M., Hau, R., Hertwig, R., Stewart, T., West, R., C., L., 2010. A choice prediction competition: Choices from experience and from description. *J. Behav. Decis. Mak.* 23 (1), 15–47.
- Fisher, M., Raman, A., 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* 44 (1), 87–99.
- Gao, J., Corter, J.E., 2015. Striving for perfection and falling short: The influence of goals on probability matching. *Mem. Cogn.* 43, 748–759.
- Gavirneni, S., Isen, A.M., 2010. Anatomy of a newsvendor decision: Observations from a verbal protocol analysis. *Prod. Oper. Manage.* 19 (4), 453–462.
- Gershman, S.J., 2015. Do learning rates adapt to the distribution of rewards? *Psychon. Bull. Rev.* 22 (5), 1320–1327.
- Glöckner, A., Betsch, T., 2011. The empirical content of theories in judgment and decision making: Shortcomings and remedies. *Judgm. Decis. Mak.* 6, 711–721.
- Gu, Y., Chen, A., Kitthamkesorn, S., 2022. Weibit choice models: Properties, mode choice application and graphical illustrations. *J. Choice Model.* 44, 100373.
- Han, W., Wang, X., Petropoulos, F., Wang, J., 2019. Brain imaging and forecasting: Insights from judgmental model selection. *Omega* 87, 1–9.
- Ho, T.H., Lim, N., Cui, T.H., 2010. Reference dependence in multilocation newsvendor models: A structural analysis. *Manag. Sci.* 56 (11), 1891–1910.
- Huxhold, O., Li, S.-C., Schmiedek, F., Lindenberger, U., 2006. Dual-tasking postural control: aging and the effects of cognitive demand in conjunction with focus of attention. *Brain Res. Bull.* 69 (3), 294–305.
- Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47, 263–291.
- Kirschner, H., Fischer, A., Danielmeier, C., Klein, T., Ullsperger, M., 2024. Cortical β power reflects a neural implementation of decision boundary collapse in speeded decisions. *J. Neurosci.* 44 (13).
- Kirshner, S.N., Moritz, B.B., 2020. Comment on “newsvendor demand chasing revisited”. *Manag. Sci.* 66 (12), 6062–6064.
- Kirshner, S.N., Moritz, B.B., 2021. Measuring demand chasing behavior. *Decis. Sci.* 52 (6), 1264–1281.

- Kirshner, S.N., Moritz, B.B., 2023. For the future and from afar: Psychological distance and inventory decision-making. *Prod. Oper. Manage.* 32 (1), 170–188.
- Krajchich, I., Oud, B., Fehr, E., 2014. Benefits of neuroeconomic modeling: New policy interventions and predictors of preference. *Am. Econ. Rev.* 104 (5), 501–506.
- Kremer, M., Minner, S., Van Wassenhove, L.N., 2010. Do random errors explain newsvendor behavior? *Manuf. Serv. Oper. Manage.* 12 (4), 673–681.
- Lau, N., Bearden, J.N., 2013. Newsvendor demand chasing revisited. *Manag. Sci.* 59 (5), 1245–1249.
- Lee, Y.S., Siemsen, E., 2017. Task decomposition and newsvendor decision making. *Manag. Sci.* 63 (10), 3226–3245.
- Lejarraga, T., Gonzalez, C., 2011. Effects of feedback and complexity on repeated decisions from description. *Organ. Behav. Hum. Decis. Process.* 116 (2), 286–295.
- Long, X., Nasiry, J., 2015. Prospect theory explains newsvendor behavior: The role of reference points. *Manag. Sci.* 61 (12), 3009–3012.
- Meyniel, F., Pessiglione, M., 2014. Better get back to work: a role for motor beta desynchronization in incentive motivation. *J. Neurosci.* 34 (1), 1–9.
- Nagarajan, M., Shechter, S., 2014. Prospect theory and the newsvendor problem. *Manag. Sci.* 60 (4), 1057–1062.
- Niranjani, T.T., Ghosalya, N.K., Menon, R.R., Rotaru, K., Gavirneni, S., 2023. Unpacking the cognitive processes of the boundedly rational newsvendor. *Prod. Oper. Manage.* 32 (10), 3138–3157.
- Ockenfels, A., Selten, R., 2014. Impulse balance in the newsvendor game. *Games Econ. Behav.* 86, 237–247.
- O'Connell, R.G., Dockree, P.M., Kelly, S.P., 2012. A supramodal accumulation-to-bound signal that determines perceptual decisions. *Nature Neurosci.* 15 (12), 1729–1738.
- Pochepstova, A., Amir, O., Dhar, R., Baumeister, R.F., 2009. Deciding without resources: Resource depletion and choice in context. *J. Mark. Res.* 46 (3), 344–355.
- Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10 (2), 59–63.
- Polich, J., 2007. Updating P300: An integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118 (10), 2128–2148.
- Pornpattananankul, N., Grogans, S., Yu, R., Nusslock, R., 2019. Single-trial EEG dissociates motivation and conflict processes during decision-making under risk. *NeuroImage* 188, 483–501.
- Regenwetter, M., Davis-Stober, C.P., 2018. The role of independence and stationarity in probabilistic models of binary choice. *J. Behav. Decis. Mak.* 31 (1), 100–114.
- Ren, Y., Croson, R., 2013. Overconfidence in newsvendor orders: An experimental study. *Manag. Sci.* 59 (11), 2502–2517.
- Rouhani, N., Norman, K.A., Niv, Y., Bornstein, A.M., 2020. Reward prediction errors create event boundaries in memory. *Cognition* 203 (1), 104269.
- Roux, F., Wibral, M., Mohr, H.M., Singer, W., Uhlhaas, P.J., 2012. Gamma-band activity in human prefrontal cortex codes for the number of relevant items maintained in working memory. *J. Neurosci.* 32 (36), 12411–12420.
- Ruggeri, K., Alí, S., Berge, M.L., Bertoldo, G., et al., 2020. Replicating patterns of prospect theory for decision under risk. *Nat. Hum. Behav.* 4 (6), 622–633.
- Saldana, C., Claidière, N., Fagot, J., Smith, K., 2022. Probability matching is not the default decision making strategy in human and non-human primates. *Sci. Rep.* 12 (1), 13092.
- Schulze, C., Newell, B.R., 2016a. More heads choose better than one: Group decision making can eliminate probability matching. *Psychon. Bull. Rev.* 23 (3), 907–914.
- Schulze, C., Newell, B.R., 2016b. Taking the easy way out? Increasing implementation effort reduces probability maximizing under cognitive load. *Mem. Cogn.* 44 (5), 806–818.
- Schweitzer, M.E., Cachon, G.P., 2000. Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Manag. Sci.* 46 (3), 404–420.
- Shen, Q., Fu, S., Huang, Y., An, Y., et al., 2022. Asymmetric valuation and belief updating over gain and loss in risky decision making: A behavioral and electrophysiological investigation. *Behav. Brain Res.* 429, 113909.
- Shinde, A., Mehta, P., Amit, R., 2021. An experimental investigation of newsvendor decisions under ambiguity. *Int. J. Prod. Res.* 59 (19), 5960–5971.
- Solway, A., Schneider, I., Lei, Y., 2022. The relationships between subclinical OCD symptoms, beta/gamma-band power, and the rate of evidence integration during perceptual decision making. *NeuroImage: Clin.* 34, 102975.
- Stewart, N., Chater, N., Brown, G.D.A., 2006. Decision by sampling. *Cogn. Psychol.* 53 (1), 1–26.
- Su, X., 2008. Bounded rationality in newsvendor models. *Manuf. Serv. Oper. Manage.* 10 (4), 566–589.
- Surti, C., Celani, A., Gajpal, Y., 2020. The newsvendor problem: The role of prospect theory and feedback. *European J. Oper. Res.* 287 (1), 251–261.
- Sweller, J., Van Merriënboer, J.J., Paas, F.G., 1998. Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10 (3), 251–296.
- Talluri, K.T., Van Ryzin, G.J., 2006. *The Theory and Practice of Revenue Management*. Springer Science & Business Media.
- Twomey, D.M., Murphy, P.R., Kelly, S.P., O'Connell, R.G., 2015. The classic P300 encodes a build-to-threshold decision variable. *Eur. J. Neurosci.* 42 (1), 1636–1643.
- Unturbe, J., Corominas, J., 2007. Probability matching involves rule-generating ability: a neuropsychological mechanism dealing with probabilities. *Neuropsychology* 21 (5), 621.
- Uppari, B.S., Hasija, S., 2019. Modeling newsvendor behavior: A prospect theory approach. *Manuf. Serv. Oper. Manage.* 21 (3), 479–711.
- Van Mieghem, J.A., 2003. Capacity management, investment, and hedging: Review and recent developments. *Manuf. Serv. Oper. Manage.* 5 (4), 269–302.
- Vipin, B., Amit, R.K., 2019. Debating decision bias in the newsvendor problem: A prospect theory model. *Omega* 82, 132–141.
- Vulkan, N., 2000. An economist's perspective on probability matching. *J. Econ. Surv.* 14 (1), 101–118.
- Wang, L., Zheng, J., Huang, S., Sun, H., 2015. P300 and decision making under risk and ambiguity. *Comput. Intell. Neurosci.* 2015, 108417.
- Wilcox, N.T., 2024. Conditional independence in a binary choice experiment. *J. Risk Uncertain.* 68 (3), 205–225.
- Wilson, Jr., W.A., Rollin, A.R., 1959. Two-choice behavior of rhesus monkeys in a noncontingent situation. *J. Exp. Psychol.* 58 (2), 174.
- Wulff, D.U., Mergenthaler-Canseco, M., Hertwig, R., 2018. A meta-analytic review of two modes of learning and the description-experience gap. *Psychol. Bull.* 144, 140–176.
- Wyatt, L., Hewan, P., Hogeveen, J., Spreng, R., Turner, G., 2024. Exploration versus exploitation decisions in the human brain: A systematic review of functional neuroimaging and neuropsychological studies. *Neuropsychologia* 192, 108740.
- Zhang, Y., Siemsen, E., 2019. A meta-analysis of newsvendor experiments: Revisiting the pull-to-center asymmetry. *Prod. Oper. Manage.* 28 (1), 140–156.
- Zhao, Y., Zhao, X., Wang, L., Chen, Y., Zhang, X., 2016. Does elicitation method matter? Behavioral and neuroimaging evidence from capacity allocation game. *Prod. Oper. Manage.* 25 (5), 919–934.