

Overview, Design Concepts and Details of the Model

Overview of the model

Purpose

The purpose of the model is to investigate the underlying factors contributing to the high concentration observed in China's e-retail markets, with a specific focus on consumer heterogeneity. While previous research has identified various structural and behavioral determinants of market concentration, the potential influence of consumer types has remained underexplored. This study addresses this gap by examining how distinct consumer behaviors—namely, risk-taking and pickiness tendencies—affect market dynamics and concentration levels. By integrating these behavioral dimensions with the dynamics of sales and WoM feedback, the simulation seeks to generate insights that not only advance theoretical understanding but also offer practical implications for e-retailers and regulatory administrators. Specifically, we carry out simulation experiments to analyze different scenarios that are designed by different target mechanisms. In the first scenario, "pure" markets are examined, each of which contains a single extreme type of consumer. The second scenario focuses on various distributions of each dimension and analyzes how they influence market concentration. Finally, the third scenario investigates the evolution of market concentration in the presence of mixed consumer types.

Entities, state variables and scales

There are two types of entities in the model, consumers and sellers. Below, we introduce their characteristics, state variables, and scales. Additional parameter settings used in the model are summarized in Table 1, which shows the main parameters, their definitions, value ranges, and dynamic states regarding their roles within the model.

1. Consumers

Population and Spatial Distribution: The model comprises N_c consumers. N_c is a model parameter to be adjusted by the modeler. Consumers are placed on a two-dimensional continuous space by independently drawing an x and y coordinate from a uniform random distribution. Each consumer is assigned to a coordinate provided that the location is unoccupied.

State Variables: Consumers are characterized by a set of state variables that define their behavior: risk-taking propensity during purchases, which reflects their willingness to buy new products when buying products, and pickiness during online commenting, which indicates their tendency to be critical or positive when leaving reviews. These two variables collectively define consumer types. Additionally, consumers exhibit limited attention, which measures the extent of their online search or exploration before making a purchase decision.

- **Risk-Taking Tendency (R_i):**

- A continuous variable defined in the interval $[0, 1]$.
- $R_i = 0$ indicates a fully conservative consumer who buys only familiar products,

while $R_i = 1$ represents a fully risk-taking consumer who buys only new products. Intermediate values represent varying degrees of risk-taking behavior.

- R_i is initialized using a Beta distribution $R_i \sim \text{Beta}(\alpha, \beta)$.
- **Pickiness Tendency (P_i):**
 - A continuous variable defined in the interval $[0, 1]$.
 - $P_i = 0$ represents a highly tolerant consumer who gives positive reviews, while $P_i = 1$ denotes a highly picky consumer who gives negative reviews. Intermediate values reflect different levels of pickiness.
 - P_i is initialized using a Beta distribution $P_i \sim \text{Beta}(\alpha, \beta)$.
- **Consumer Type Classification:**
 - Extreme consumer types are defined by the combinations of the boundary values:
 - **CT-type:** $R_i = 0$ and $P_i = 0$ (conservative and tolerant)
 - **RT-type:** $R_i = 1$ and $P_i = 0$ (risk-taking and tolerant)
 - **CP-type:** $R_i = 0$ and $P_i = 1$ (conservative and picky)
 - **RP-type:** $R_i = 1$ and $P_i = 1$ (risk-taking and picky)
 - For a broader representation of consumer behaviors, the initial values of R_i and P_i are generated from a Beta distribution, $x \sim \text{Beta}(\alpha, \beta)$. This approach allows for various distribution shapes:
 - Symmetric when $\alpha = \beta$
 - Positively skewed when $\alpha > \beta$
 - Negatively skewed when $\alpha < \beta$
- **Limited Attention (m):** A discrete parameter quantifying the consumer's ability to search for sellers/products. It ranges from 1 to N_s (the total number of sellers), where $m = 1$ implies that a consumer can search only one seller/product and when $m = N_s$ indicates that a consumer can review all available products. This parameter encapsulates bounded rationality and incomplete information processing in consumer decision-making.

2. Sellers

Population and Spatial Distribution: The market consists of N_s sellers, each located within the same two-dimensional space as the consumers.

State Variables:

- **Product Homogeneity.** All sellers produce identical products. The differentiation among sellers arises solely from consumer-based evaluations.
- **Sales.** Sales are defined as the frequency with which consumers purchase products from a seller during a given time period. This variable is normalized to the interval $[0, 1]$ to facilitate comparisons and to capture the normative influence of widespread product adoption.
- **Word-of-Mouth (WoM).** WoM aggregates the subjective evaluations provided by consumers who have purchased from a seller. Given that picky consumers tend to assign lower ratings and tolerant consumers higher ratings, the WoM serves as an informative metric reflecting both positive and negative consumer feedback. Like sales, WoM is normalized to the interval $[0, 1]$. Importantly, information regarding a seller's WoM and sales is only accessible to consumers who can search for it (as determined by their limited attention m).

3. Scales

Spatial Scale. Consumers and sellers are embedded in a shared two-dimensional continuous space, with locations assigned via uniformly distributed coordinates to ensure random spatial distribution. Sellers remain stationary after initialization, while consumers are mobile. Following a purchase, consumers move closer to the seller they bought from. The specific movement rules (e.g., distance, speed, proximity) are detailed in the submodels section.

Temporal Scale. The simulation advances in discrete time steps. At each time step, interactions between consumers and sellers occur, and state variables (e.g., sales and WoM) are updated according to the model dynamics.

Table 1: Parameters used in the model.

Parameter	Description	Values	Dynamic?
Nc	Number of consumers	Integer (model-defined)	No
Ns	Number of sellers	Integer (model-defined)	No
Ri	Risk-taking tendency	[0, 1]	No
Pi	Pickiness tendency	[0, 1]	No
m	Limited attention	Integer from 1 to Ns	No

S	Seller's sales volume	Normalized to [0, 1]	Yes
WoM	Seller's word-of-mouth	Normalized to [0, 1]	Yes
Wm	Weight of the WOM in defining the product attractiveness and (1-Wm) is the weight given to the sales	[0, 1]	No
A	Product attractiveness		Yes
A _{bf}	Attraction level of the best-found product at the current step	[0, 1]	Yes
A _{lp}	Attraction level of the last-purchased product in the last step	[0, 1]	Yes
d	The difference in attraction between A _{bf} and A _{lp}	[-1, 1]	Yes
t	Threshold of tolerance for attractiveness differences, reflecting how decisions are influenced by both objective product attractiveness and subjective risk preferences	[0, 1]	No

α and β	Shape parameters for the Beta distribution used to generate R_i and P_i	$\alpha, \beta \geq 0$	No
----------------------	---	------------------------	----

Process overview and scheduling

The consumer purchasing process in the model unfolds over discrete time steps, with each cycle representing one complete round of interaction between consumers and sellers. The purchasing process consists of three sequential stages: Search, Consumption, and Evaluation, mirroring the typical phases of online shopping identified in previous research. A flowchart summarizing these processes is provided in **Fig. 3** in the main text.

Search Stage. The stage represents the initial phase of the consumer purchasing process, where consumers gather information about sellers and products to identify potential purchase options. This stage is influenced by consumers' limited attention and their historical purchase data. Specifically, at the start of each time step, consumers enter the market with asymmetric information. Initially lacking any purchase history, they conduct random searches for sellers and products. Consumers are motivated to explore the market to obtain information on product attributes, including sales figures and aggregated consumer feedback. However, their exploration is limited by a fixed attention span, quantified by the parameter m . This parameter restricts each consumer to consider at most m sellers during the search phase, reflecting the bounded rationality and information-processing constraints observed in real online shopping behaviors. As the simulation advances over multiple time steps, consumers accumulate a record of past purchases, which in turn informs subsequent search behavior.

Consumption Stage. Consumers decide which product to purchase, the last-purchased product or the new product, based on the information gathered during the search stage. Firstly, consumers identify the best-found product (highest attractiveness) among the m sellers they evaluated. Secondly, they compare the best-found product with their last-purchased product (historical choice) using the following rules: Conservative consumers (with lower R_i) prefer to stick with their last-purchased product unless the best-found product offers a significant advantage; Risk-taking consumers (with higher R_i) prefer to switch to the best-found product unless the last-purchased product is significantly better. Lastly, after the comparison, consumers make their purchase decisions and update their purchase history (last-purchased product).

Evaluation Stage. Consumers provide feedback on the purchased product, influencing the seller's WoM. Consumers evaluate the purchased product based on their pickiness tendency (P_i): With probability P_i , the consumer leaves a negative rating, decreasing the product's rating by 1; With probability $1-P_i$, the consumer leaves a positive rating, increasing the product's rating by 1 (see Eq. 3). This feedback mechanism captures the dual influence of consumer evaluation, where the collective ratings, in conjunction with sales figures, shape the attractiveness of each seller in subsequent time steps.

Scheduling and Model Iteration. Each full cycle of Search-Consumption-Evaluation

constitutes one time step. At the end of each time step, the model updates the market environment. a) Market updates. Sellers' sales and WoM are updated based on the latest purchases and feedback. The market calculates sales rankings and concentration levels (e.g., market share distribution). b) Consumer updates. Consumers update their purchase history (last-purchased product). c) Seller updates. Sellers receive new consumer feedback and update their sales and WoM. With these updates, the system transitions seamlessly into the next time step, and the entire process repeats.

Design concepts

This section outlines the key theoretical and conceptual foundations of the model that capture the micro-to-macro linkages in online retail markets, including emergent phenomena, agent adaptation, decision-making processes, and other design principles. The following points outline the underlying ideas and mechanisms.

Basic principles. The model simulates a market where consumer heterogeneity, specifically in terms of risk-taking and pickiness, drives market dynamics. Sellers are assumed to be homogeneous in terms of the products they offer, with differentiation arising solely from consumer-generated feedback (sales and word-of-mouth ratings). This design isolates the effect of consumer behavior on market concentration, emphasizing the role of individual decision-making processes in shaping aggregate outcomes. The model is grounded in theories of bounded rationality and consumer behavior, particularly in the context of online shopping. It incorporates insights from prior research on:

Emergence. Macro-level phenomena, mainly market concentration and dynamics, emerge from the repeated interactions of individual consumers with sellers. The aggregation of purchase decisions, influenced by consumers' risk-taking and pickiness traits, leads to evolving sales figures and word-of-mouth scores. For example, some sellers may dominate the market due to higher sales and better WoM (Market Concentration), and positive or negative feedback from consumers can amplify or diminish a seller's reputation over time (Feedback Loops). These emergent market patterns reflect the cumulative effect of local decisions and feedback loops, demonstrating how micro-level behaviors can lead to unexpected macro-level properties.

Adaptation. Consumers adapt their purchasing decisions over time based on their past experiences. With each time step, they update their purchase history (last-purchase), which then informs the comparison with newly discovered options (best-found). This adaptive behavior is modulated by each consumer's risk-taking tendency. Conservative consumers tend to adhere to their established purchase patterns unless presented with a significantly superior option, whereas risk-taking consumers are more inclined to switch to new products. Simultaneously, sellers' WoM and sales are updated dynamically based on consumer feedback, influencing their market position. This adaptive decision-making process enables the model to capture the dynamic evolution of consumer preferences and market structure.

Objectives. The primary objective for consumers is to maximize their satisfaction by selecting products that best meet their preferences, as indicated by the attractiveness of products (a function of sales and word-of-mouth). For sellers, while there is no explicit decision-making process, their performance (as measured by sales and WoM) serves as a proxy for market

success.

Learning. Although the model does not incorporate an explicit learning algorithm, consumers indirectly "learn" from their purchase history. Their evolving last-purchase record reflects accumulated market experience, guiding future choices in a manner consistent with bounded rationality. Over time, this historical record influences their search and consumption decisions, thereby mimicking a form of experiential learning.

Interaction. Interactions in the model occur indirectly through market signals rather than direct communication among consumers. Two primary forms of interaction are represented. The first is *consumer-seller interaction*: a) Consumer evaluations (through WoM) and seller performance (through sales) serve as the conduits of interaction. b) Sellers' WoM and sales are influenced by consumer behavior, creating a feedback loop. The second is *consumer-consumer interaction*: Consumers interact indirectly through WoM, as their collective feedback influences the ratings and attractiveness of sellers' products. Such indirect interaction framework allows the model to capture the influence of social feedback mechanisms, such as the spread of word-of-mouth, on market dynamics without modeling explicit social networks.

Stochasticity. Several elements of the model are stochastic. The spatial distribution of consumers is determined by random draws from a uniform distribution. And they randomly select m sellers to evaluate during the initial search stage, introducing variability in their decision-making process. Additionally, the initial distributions of the risk-taking and pickiness tendencies are generated using a Beta distribution, allowing for varied and probabilistic representations of consumer preferences. The evaluation stage introduces further stochasticity, as each consumer leaves a negative rating with probability P_i and a positive rating with probability $1 - P_i$.

Observation. The aggregated effects of individual actions, such as cumulative sales and word-of-mouth, form the basis for the emergent market structure. Sellers are ranked based on these collective measures, and overall market concentration is observed as a macro-level outcome. Although individual consumer decisions are made in isolation, collective behavior is observable through market-level indicators, enabling the assessment of phenomena like market concentration and dynamics. Specifically, the model tracks and outputs the following metrics: a) Sales Concentration: The distribution of sales across sellers. b) WoM Dynamics: Changes in sellers' aggregated ratings over time. c) Market Evolution: Changes in market structure and seller rankings over time.

Details

Initialization

The simulation market is initialized with a predefined number of consumers and sellers, each assigned specific behavioral traits and constraints. The key initialization parameters are as follows.

Population Setup. a) The market consists of **300 consumers** and **30 sellers**, randomly distributed in space. b) Consumers enter the market with no prior purchase history, requiring

them to engage in random searches in the initial steps.

Consumer Attention (m). a) Each consumer has a limited attention span and can only remember up to three sellers per transaction ($m = 3$). b) This constraint reflects bounded rationality in online shopping, as consumers cannot evaluate all available options but must instead rely on a limited consideration set.

Consumer Traits (R_i and P_i). a) The behavioral tendencies of consumers, including their risk-taking propensity (R_i) and pickiness (P_i), are initialized using Beta-distributed random values. b) The specific parameter settings for these distributions are given in Table 2. c) The distribution type determines whether most consumers exhibit risk-taking (left-skewed), symmetric unimodal distribution (centered around 0.5), or less risk-taking (right-skewed) behaviors.

Table 2: Beta distribution parameter settings for two dimensions of consumer type.

Parameter Setting	α	β
Left-skewed	10	1
Unimodal distribution centered around 0.5	10	10
Right-skewed	1	10

Market Experimentation and Replication. a) Each simulation experiment consists of 20 repetitions, and results are averaged across repetitions to ensure statistical reliability. b) One full trade cycle (including search, consumption, and evaluation) constitutes a single time step, with the market iterating over multiple time steps.

Seller Initialization. a) All sellers start with neutral ratings and zero initial sales, ensuring that early market dynamics emerge from consumer decisions rather than preset seller advantages. b) The attractiveness of sellers evolves based on consumer evaluations and cumulative sales performance over time.

Input Data

The model does not use input data to represent time-varying processes.

Submodels

Consumer traits. The heterogeneity of consumer behavior in the artificial market is captured through two key traits: risk-taking (R_i) and pickiness (P_i). These traits are initialized using Beta distributions (as shown in Eq. 1), which allow for flexible modeling of diverse behavioral tendencies among consumers, parameterized as $R_i \sim \text{Beta}(\alpha_1, \beta_1)$ and $P_i \sim \text{Beta}(\alpha_2, \beta_2)$.

The implementation in the model follows three steps: a) At the start of the simulation, each consumer is assigned a unique R_i and P_i value drawn from the appropriate Beta distribution. b) These traits remain constant for each consumer throughout the simulation, ensuring persistent behavioral differences in decision-making and feedback tendencies. c) The selected parameter values allow for experimentation with different market conditions, where populations may be predominantly e.g., risk-averse, risk-neutral, or risk-seeking.

Normalization function. This submodel describes how **Word-of-Mouth (WoM)** and **Sales** are normalized to the range $[0, 1]$ to ensure comparability across sellers. The normalization of WoM is calculated using the formula in (Eq. 4). The same normalization process is applied to Sales according to the formula:

$$S_{norm} = \begin{cases} 0.5 & \text{if } sales_{max} = sales_{min} \\ \frac{sales - sales_{min}}{sales_{max} - sales_{min}} & \text{otherwise} \end{cases}$$

where:

$sales$ is the total number of units sold by a specific seller.

$sales_{max}$ and $sales_{min}$ are the highest and lowest sales figures in the market, respectively.

If all sellers have the same sales ($sales_{max} = sales_{min}$), the normalized sales value is set to 0.5 to avoid division by zero. Otherwise, sales are linearly scaled between 0 and 1, where sellers with higher sales receive a higher normalized score.

Market concentration. To evaluate the structure of the market and detect potential dominance by leading sellers, a concentration ratio (CR) is used. This metric quantifies the degree of market concentration by aggregating the market shares of the top n sellers in the market. In this model, we employ the 1-seller concentration ratio (CR1) as an indicator of market concentration. CR1 specifically captures the proportion of total market sales controlled by the single largest seller at each time step. The CR1 is computed as follows: a) Aggregate Sales. For each time step, the total sales of all firms are summed to determine the overall market size. b) Identify the Largest Seller. Among all firms, the one with the highest sales during the time step is identified. c) Compute CR1. The market share of the largest seller is calculated by dividing its sales by the total market sales:

$$CR1_t = \frac{Sales_{max,t}}{\sum_{i=1}^N Sales_{i,t}}$$

Where: $sales_{max}$ is the sales of the largest seller at time t ; $\sum_{i=1}^N Sales_{i,t}$ is the total sales of all N s sellers at time t , $CR1_t$ ranges from 0 to 1, with higher values indicating stronger market dominance by a single seller.

Product attraction. The attraction of a product determines consumer preferences and purchasing behavior. It is computed based on word-of-mouth and sales, both of which are normalized between 0 and 1. The total attraction of a product is given by (Eq. 2). The purpose of attraction calculation is to provide a unified metric for evaluating products, combining both reputation (WoM) and popularity (Sales), and guide consumer decision-making in the consumption stage, helping consumers identify the most attractive product.

Consumer Movement. After purchasing a product, consumers adjust their position in the market space, moving closer to the seller from whom they made the purchase. This movement is designed solely for visualization purposes, to illustrate market concentration more clearly and vividly, but it does not influence any future interactions or decision-making processes.

The movement mechanism follows these rules: a) Facing the seller. Consumers turn to face the seller they purchased from. b) Movement step. Consumers move 1.5 distance units toward the seller after each purchase. If the consumer is already near the seller, they remain in their current location. c) Location update. The consumer's new position is recorded but does not influence their future decisions.