ARTICLE IN PRESS

International Journal of Research in Marketing xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Predicting livestream shopping success

Zekun Liu [a], Weiqing Zhang [b], Xiao Liu [c], Eitan Muller [c,*], Feiyu Xiong [d]

[a] Indiana University, USA
[b] Rice University, USA
[c] New York University, USA
[d] Alibaba Group, China

ARTICLE INFO

Keywords:
Livestream shopping
Survival
E-commerce
Entertainment
Social network

ABSTRACT

Despite the popularity of the livestream shopping industry—projected to surpass $830 billion in China and $68 billion in the U.S. by 2026—research on identifying the factors that drive success in this industry remains scarce, primarily because of the complexity of this environment. In this study, we use data from Taobao Live, the largest livestream shopping platform, to predict the success of sellers hosting livestream sessions. We develop a scoring model using a comprehensive set of livestream strategies and metrics, including session strategies, product assortment strategies, livestream content metrics, and engagement metrics. We find that incorporating session and assortment strategies and livestream content and engagement metrics improves the accuracy of sales performance predictions by 25.6%. Additionally, we identify the key strategies and metrics that are most predictive of sales performance. The proposed predictive framework not only helps in identifying promising sellers but also enhances the understanding of the complex dynamics of the livestream shopping ecosystem, providing valuable insights for stakeholders such as investors, lenders, suppliers, and platform operators.

1. Introduction

Livestream shopping, a relatively new format of e-commerce, uses video streaming to demonstrate products in real time as a strategy for engaging audiences and generating sales. This format exists alongside the longstanding options of brick-and-mortar stores and traditional internet shopping. The livestream shopping industry has grown rapidly. In its largest market, China, livestream shopping reached 600 million people and is projected to reach $830 billion in revenue by 2026.[1] In the U.S., online platform giants such as Amazon and Facebook have incorporated livestream shopping into major sales events such as Black Friday, and the value of the livestream shopping industry is expected to exceed $68 billion in 2026.[2]

However, the increased popularity of livestream shopping does not guarantee seller success, and many sellers phase out within a year.[3] Hence, identifying the key predictors of success of businesses (e.g., Audretsch et al., 2000; Dekimpe & Morrison, 1991; Zhang & Luo, 2023) is of great importance to both investors and sellers. Because of the complex landscape of the livestream shopping industry,

* Corresponding author.
E-mail addresses: zl73@indiana.edu (Z. Liu), wz60@rice.edu (W. Zhang), xl23@stern.nyu.edu (X. Liu), emuller@stern.nyu.edu, em101@stern.nyu.edu (E. Muller), feiyu.xfy@alibaba-inc.com (F. Xiong).

1 Source: https://uxspot.com/live-stream-ecommerce-evolution-in-china
2 Source: https://tinyurl.com/dft7mksw
3 Source: https://tinyurl.com/2mj6a9pu

https://doi.org/10.1016/j.ijresmar.2025.12.002
Received 26 June 2022
Available online 28 December 2025

where sellers must navigate both e-commerce factors and the unique dynamics of live interaction, it is crucial for investors, sellers, and retail platforms to understand the predictors of success.

In this paper, we build a scoring model to predict the success of sellers using livestream shopping and explore the key indicators of sales performance, going beyond the traditional metrics associated with an online store's characteristics. Specifically, our analysis seeks to uncover which aspects of livestreaming (such as session strategies, product assortment strategies, livestream content, and audience engagement) are most predictive of seller sales performance. By analyzing these factors, we aim to provide valuable insights that help sellers monitor key metrics over time, to flag potential issues or opportunities in the competitive world of livestream shopping, while also guiding investors in identifying standout sellers.

We address these research questions by collaborating with Taobao Live, the largest livestream shopping platform in the world, with a 55% share of China's livestream e-commerce market in 2020.[4] We selected a representative sample of 4061 sellers who launched livestreaming channels between August 2019 and February 2020. These sellers operate within Taobao Live's primary product categories: Fashion, Food, and Jewelry. We tracked their sales performance and livestream activities for five months following their first livestream. Additionally, we collected data on each seller's online store characteristics. Our analysis focuses on livestream strategies, including session strategies (such as length and frequency), assortment strategies during livestreams (such as product price and tenure), livestream content metrics, and social engagement metrics. We examined the predictive power of these factors on sales performance. To further assess the incremental value of these factors, we employed the XGBoost algorithm, comparing their predictive strength with traditional metrics associated with online store characteristics. We then applied SHapley Additive exPlanations (SHAP) to decompose the model's predictions, which enabled us to interpret the importance of each feature at the seller level. This approach clarifies how various livestream features contribute to each seller's predicted performance.

We obtained the following results. Adding session strategies, assortment strategies, livestream content metrics, and social engagement metrics substantially improves the prediction of seller success beyond traditional online store factors, particularly for jewelry sellers. The importance of these factors varies with seller tenure: livestream content and assortment strategies matter more for newer sellers, and livestream content metrics are the strongest predictors for more experienced sellers. Sessions featuring higher-priced, moderately popular products, and a diverse mix of offerings by popularity, tenure, and presentation rate are linked to higher sales. Additionally, a higher seller service scores is associated with better sales performance. When considering social engagement, only meaningful interactions (such as comments, watch time, and unique visitors who clicked "like") emerge as strong predictors of sales.

The main contributions of our paper are twofold. First, our paper contributes to the growing body of literature on livestream shopping. Previous studies have empirically explored various aspects of influencer-hosted livestreams, including characteristics such as influencer popularity and type (Gu et al., 2024; Sun & Tang, 2024; Wang et al., 2022), script style and sales pitch (Li & Han, 2024), relationship capacities (Wongkitrungrueng et al., 2020; Zhang et al., 2024a), emotional appeal (Lin et al., 2021), facial expressions (Bharadwaj et al., 2022), and engagement (Yang et al., 2025); and the dynamics of brand-influencer negotiations (Lin et al., 2024). While much of the existing literature centers on influencer-hosted livestreams, our paper focuses on seller-hosted livestreams, an increasingly popular format adopted by sellers. Notably, the livestream shopping industry has witnessed a significant shift toward seller-led broadcasts. For instance, on the Taobao platform, approximately 70% of livestream sessions in 2020 were operated by sellers rather than influencers, and this format is rapidly becoming the norm. Despite the growing prevalence of seller-hosted livestreaming, it remains relatively underexplored in academic research. The distinction between seller-hosted livestreams and influencer-hosted livestreams allows us to move beyond traditional metrics such as influencer characteristics, and to identify novel livestream strategies and metrics that are most predictive of sales success. Zhang et al. (2024b) examines a similar context of seller-hosted livestreams and investigates how hosting livestream shopping affects seller revenue. Our study contributes to this stream of literature by introducing a predictive framework that encompasses comprehensive seller strategies, including session strategies, assortment strategies, and livestream content metrics. This framework provides additional managerial insights to various stakeholders. In Table 1 we summarize the livestream shopping literature and highlight our contribution.

Second, our paper also contributes to the stream of literature exploring the role of engagement in e-commerce. For instance, audience reach and engagement activities have been treated as important metrics for evaluating the effectiveness of mobile apps and videos (Hilvert-Bruce et al., 2018; Van Heerde et al., 2019; Wongkitrungrueng & Assarut, 2020; Wymer, 2019), and they are linked with purchase intent and conversion in promotional campaigns (John et al., 2017; Tucker, 2015; Voorveld et al., 2018; Yang et al., 2025). Our paper examines the role of engagement metrics within the livestream shopping industry, revealing that only meaningful interactions between viewers and sellers have a positive impact on sales performance. These findings offer valuable benchmarks for stakeholders and researchers, helping to deepen their understanding of engagement metrics in this rapidly expanding industry.

Our findings can inform the decision-making processes of various stakeholders. First, the scoring model provides significant benefits to investors, suppliers, lenders, and platforms. For investors, the ability to predict a firm's success can help evaluate investment opportunities, enabling them to optimize their portfolios by choosing sellers with higher chances of success or divesting from riskier ones. Suppliers, who often rely on other firms for parts, services, or strategic partnerships, can use these success predictions to evaluate the reliability of potential partners, ensuring long-term collaboration and minimizing disruptions due to partner failure. Lenders, focused on firms' repayment capacity, can use these success indicators to better assess credit risk and adjust loan terms, helping them avoid defaults and identify firms with lower financial risk. In addition to traditional lenders such as banks, digital platforms often play an important role in providing financial services to sellers. For instance, Alibaba, the owner of Taobao, actively supports small- and medium-sized sellers through interest-free financing programs. The success score developed in this study can

[4] Source: https://www.huxiu.com/article/396903.html

Table 1
Summary of livestream shopping studies.

Author(s)	Research Question	Key Findings	Livestreamer Type
Wongkitrungrueng et al. (2020)	How do utilitarian, hedonic, and symbolic values impact customer trust and engagement in livestream shopping?	Customer engagement is driven by trust in products and sellers. Utilitarian and hedonic values influence engagement indirectly, whereas symbolic value has a direct positive effect.	Not specified
Bharadwaj et al. (2022)	How do facial expressions and face presence affect sales performance in livestream retail?	Emotional displays do not directly increase sales. A balanced facial presence helps, with a straight-faced approach outperforming excessive emotion.	Salespeople
Gu et al. (2024)	What are the sales effects of using big vs. small influencers, and how do these effects interact with price discounts and co-promotion?	Big influencers significantly boost sales. Discounts are more effective with big influencers; mixing big and small influencers can dilute effects.	Influencers
Li and Han (2024)	What is the sales lift from livestreaming sales pitches?	Sales pitches increase sales by about 150% without cannibalization or higher returns.	Influencers
Yang et al. (2025)	What visual and textual elements drive user engagement in video ads?	High-energy, fast-paced, and novel videos attract more engagement.	Influencers
Wang et al. (2022)	How do verbal and nonverbal livestreamer cues influence sales?	Streamer traits and nonverbal cues matter. Sales pitch types (product, promotion, CRM) also shape sales outcomes.	Influencers
Zhang et al. (2024a)	How do influencer characteristics and relationship capabilities affect livestream commerce sales?	Influencer traits matter more for mega-influencers; relationship capabilities are key for smaller influencers. Interaction capability is especially impactful.	Influencers
Sun and Tang (2024)	How do form and behavioral realism of AI streamers affect purchase intentions?	Both realism types increase purchase intention, with form realism enhancing the effect of behavioral realism.	AI-enabled virtual streamers
Zhang et al. (2024b)	What is the causal impact of online store sellers hosting livestream sessions?	Livestream sessions lead to higher overall revenue for sellers and generate spillover effects that indirectly boost online store revenue.	Online store owners
This paper	What livestream strategies and social engagement metrics predict seller sales performance?	Livestream strategies and social engagement significantly improve seller sales prediction beyond traditional store factors, aiding stakeholders such as Alibaba Loans.	Online store owners

serve as a decision-support tool for such financing initiatives, helping identify which sellers are more likely to benefit from and repay loans, thereby guiding more effective capital allocation.[5] Additionally, livestream shopping platforms can use our scoring model as input for the recommender system or search ranking, potentially improving the overall user experience. Therefore, our paper can help platforms identify high-potential small businesses, thereby promoting a more dynamic commercial environment and reducing operational costs associated with credit default risks.

In addition to the scoring model itself, many stakeholders seek insight into the specific indicators that are most predictive of firm success. This additional layer of analysis is valuable to analytics service providers, investors, and sellers alike. Analytics service providers can feature these leading indicators prominently within dashboards. Investors can determine whether these key indicators align with their risk tolerance or strategic objectives. Sellers can use these indicators for multiple purposes: identify areas for deeper causal investigation, monitor key metrics over time to flag potential issues or opportunities, and benchmark their performance against platform averages or top performers to drive strategic improvement.

The rest of the paper is organized as follows: In the next section, we describe the livestream shopping industry. In Section 3, we introduce our data and key variables. In Section 4, we describe our prediction model. In Section 5, we apply SHAP to interpret the model's predictions and examine the top contributing features. We conclude the paper in Section 6.

2. Background on livestream shopping

Livestream shopping relies heavily on e-commerce platforms for seamless integration with online retail, with revenue driven by the interactions between networks of consumers, livestream hosts, and business owners. The relatively low cost of launching a livestreaming channel has encouraged many small business owners to enter the livestreaming space as hosts. The streaming format is more interactive than that of traditional pre-recorded videos (such as QVC) and enables sellers to build real-time social connections with consumers and explain product features in detail. Table 2 provides an overview of the main livestream shopping platforms.

[5] Source: https://www.scmp.com/business/banking-finance/article/1557338/alibaba-teams-mainland-banks-offer-loans-smes

Table 2
Livestream shopping platforms.

	Taobao	Kuaishou	TikTok	Amazon	Wayfair	Facebook	Google
Start	2016	2018	2018	2019	2019	2018	2020
Source	Taobao	Taobao	Taobao	Amazon	Wayfair	Multiple	Multiple
DAU	30 million	100 million	100 million	~	~	~	~
Product	Apparel Cosmetics Jewelry	Deals	Cosmetics	Deals Fashion Beauty	Furniture	Used Goods	Cosmetics

Note: DAU stands for daily active user.

There are two main types of livestreams: those hosted by influencers and celebrities, and those hosted by sellers. While viewers can engage and interact with the hosts in both formats, the structures differ. In livestreams hosted by influencers and celebrities, brands compensate the hosts for promoting products. Compensation plans can vary significantly; for example, during the Singles' Day promotion in 2019, a fashion brand paid a top influencer a lump sum of ¥150,000 (approximately $22,000) along with a 20% profit commission for promoting the brand's specialized slippers in his livestreams (Jiang, 2020). In such arrangements, brands typically do not have full control over session or product assortment strategies, because brands rarely purchase entire livestream sessions for exclusive product promotions.

Recently, the livestream shopping industry has seen a shift toward more seller-hosted livestreams. Seller-hosted sessions now account for 70% of all livestreams on Taobao Live.[6] When sellers host livestreams, they exclusively promote their own brands, allowing them full control over their livestream strategies.

All livestreams in our sample were hosted by sellers on Taobao, enabling us to analyze a comprehensive range of livestream strategies and engagement metrics. During a seller-hosted livestream, the seller's Taobao store name and page link are prominently displayed in the corner of the screen, making it easy for viewers to subscribe or visit the seller's online store.

3. Data

We collect data on sellers who adopted the livestream shopping channel on Taobao Live, the largest livestream shopping platform in the world, launched by the Alibaba Group in 2016. In China, nearly 68% of consumers used the platform's services in 2020. The Alibaba e-commerce ecosystem facilitates the seamless integration of online stores and livestream shopping for sellers. We focus on sellers who adopted the livestream shopping channel within Taobao Live's top product categories: Fashion, Food, and Jewelry. The sellers conducted their first livestream sessions between August 2019 and February 2020.

We randomly select 4061 eligible sellers for analysis. For each seller, we track their sales and livestream activity. For each livestream, we recorded the session's length, product assortment, and various engagement metrics, including likes, shares, comments, new subscribers, and watch time. Additionally, we collected data on seller characteristics, such as ratings and subscriber counts. We aggregated the data to a monthly level and analyzed the first five months after a seller's adoption of the livestream shopping channel. All sellers in our dataset are small-scale businesses exclusively selling their own brands, none of which are nationally recognized. Table 3 shows the summary statistics at the seller level for the first month in which the sellers adopted the livestream shopping channel.

In the following sections, we provide a detailed explanation of the variables used to predict seller sales performance, which is measured as the logged monthly sales. The variables are defined in Table 4.

3.1. Session strategies

For each of the months in our dataset, we analyze sellers' approaches to conducting livestream sessions, focusing on their session strategies. First, we determine whether the seller conducted any livestreams during that month. If yes, we assess both the total duration of all livestream sessions and the number of days the seller hosted livestreams within that month. These factors provide key insights into how sellers allocate their time and manage the frequency of audience engagement during this period.

In addition to analyzing the livestream activity for the current month, we also track the cumulative total session length and the cumulative number of livestream days since the seller began conducting livestream sessions. This long-term perspective allows us to evaluate whether the seller's strategy evolved over time, such as by increasing audience engagement through longer or more frequent sessions. By examining both recent and historical livestreaming behavior, we gain a comprehensive understanding of how sellers develop and refine their session strategies to balance immediate audience engagement with sustained long-term growth.

[6] Source: http://www.nbd.com.cn/articles/2021-04-29/1727040.html

Table 3
Seller-level summary statistics.

Variable	Mean	SD	Min.	Median	Max.
Seller Tenure (in months)	40	42	0	24	177
Seller Credit Rating	68,242	1,384,782	0	351	63,817,954
Seller Positive Rating	0.876	0.3	0	0.996	1
Seller Subscriber Count	33,215	403,129	0	418	14,803,841
Seller Categories					
Fashion	64.6%				
Jewelry	12.4%				
Food	23.0%				
Number of Obs.	4,061				

Note: This table presents seller-level summary statistics based on their characteristics during the first month of adopting the livestream shopping channel.

Table 4
Definition of independent variables.

Variable	Definition
Panel A: Session Strategies	
If Livestream$_{it}$	Whether any livestream sessions were held
Total Livestream Session Length$_{it}$	Total duration (in hours) of all livestream sessions
Average Livestream Session Length$_{it}$	Average duration (in hours) of livestream sessions
Livestream Frequency$_{it}$	Total number of days holding livestream sessions
Panel B: Assortment Strategies	
Price$_{it}$	Average price of products featured within the same session
Variation of Price$_{it}$	Coefficient of variation in product prices within the same session
Product Tenure$_{it}$	Average tenure of products featured within the same session
Variation of Product Tenure$_{it}$	Coefficient of variation in product tenure within the same session
Product Save Popularity$_{it}$	Average save popularity of products featured within the same session
Variation of Product Save Popularity$_{it}$	Coefficient of variation in save popularity of products within the same session
Product Category Count$_{it}$	Number of distinct product categories featured within the session
Concentration of Product Category$_{it}$	Distribution concentration of product categories within the session
Product Count$_{it}$	Average number of products introduced per hour during the session
Panel C: Social Engagement Metrics	
Page Views Count$_{it}$	Average number of livestream page views
Likes Count$_{it}$	Average number of likes
Comments Count$_{it}$	Average number of comments
Shares Count$_{it}$	Average number of shares
Subscription Count$_{it}$	Average number of subscriptions
UV$_{it}$	Average number of unique viewers that watched the session
Likes UV$_{it}$	Average number of unique viewers that liked
Comments UV$_{it}$	Average number of unique viewers that commented
Shares UV$_{it}$	Average number of unique viewers that shared
Watch Time$_{it}$	Average watch time
Panel D: Livestream Content Metrics	
Content Value Score$_{it}$	The content quality metric including factors such as video quality and content relevance
Fans Value Score$_{it}$	The value of a fan base based on engagement, purchases, and loyalty
Area Value Score$_{it}$	The performance metric for a seller's livestream sessions relative to those of other sellers within the same category (area)
Seller Service Score$_{it}$	The seller service metric including factors such as responding to comments in livestream sessions
Account Health Score$_{it}$	The performance metric based on the trajectories of livestream activities and consumer purchases
Panel E: Controls	
Seller Livestream Tenure$_{i,t}$	The number of months since the seller first began livestreaming
Seller Tenure$_{i,t}$	The time length in months since the seller opened the online store
Seller Credit Rating$_{i,t-1}$	The number of successful transactions with positive feedback
Seller Positive Rating$_{i,t-1}$	The percentage of positive feedback among all successful transactions
Seller Subscriber Count$_{i,t-1}$	The total subscriber count of the seller
Seller Category$_i$	Includes three category dummies: Fashion, Jewelry, and Food
Month$_t$	Includes all month dummies starting from August 2019 to May 2020

Note: This table summarizes the definitions of the independent variables used.

3.2. Assortment strategies

In addition to session strategies, we also evaluate the sellers' assortment strategies during livestreams. A key aspect of this analysis is examining the range of products presented by sellers, focusing on several crucial characteristics: price, product tenure, product save popularity, and product category. Product tenure reflects how long a product has been available in the store, providing insight into whether sellers are showcasing newer items or relying on established, longstanding products. Product save popularity measures how many users have saved a particular product, serving as a proxy for consumer interest and potential future demand.

To quantify the seller's assortment decisions, we compute both the mean and the variance (using the coefficient of variance) of key product attributes (price, tenure, and save popularity) across all products introduced during a session. This method captures not only the average characteristics of the products but also the diversity and variation within the assortment. For instance, a higher variance might indicate a broader mix of product prices or tenures, suggesting a diverse assortment, whereas a lower variance points to a more uniform product offering.

Additionally, we assess the variety of product categories presented during each session. By calculating the unique number of categories, we measure the breadth of the assortment and employ the Herfindahl-Hirschman index (HHI) to evaluate the concentration of products within specific categories. The HHI is crucial for determining whether the seller is concentrating on a narrow range of categories or offering a more diversified portfolio within a session.

Moreover, we calculate the product count per hour for each livestream session, which is determined by dividing the total number of products introduced during the session by its duration in hours. This metric provides insight into the pace at which products are showcased and may indicate the seller's strategy to maintain viewer engagement with a continuous flow of new offerings. Finally, we average these assortment decisions across all sessions within the month to establish the seller's overall product assortment strategy for that period.

3.3. Social engagement metrics

During livestream sessions, viewers can interact with sellers through multiple forms of engagement, such as simply watching, pressing the "like" button, sharing the livestream, commenting to communicate directly with the seller, and subscribing to the seller's store. These interactions not only enhance the real-time connection between the seller and their audience but also provide valuable feedback on the effectiveness of the livestream.

For each livestream session, we track several key social engagement metrics to quantify audience participation. These include the total number of page views, likes, comments, shares, and subscriptions generated during the session. We also capture the number of unique viewers who engaged with the session by watching, liking, commenting, or sharing, offering a clearer picture of the breadth of audience interaction beyond simple view counts.

In addition to raw engagement numbers, we monitor the average watch time per viewer, which reflects the level of sustained interest from viewers. A higher average watch time suggests deeper viewer engagement, whereas lower watch times may indicate that viewers are less captivated or engaged. This metric can help us understand the quality of interaction, because it provides insight into how long viewers remain engaged with the livestream content. Finally, we average these metrics across all sessions held within the month to obtain the social engagement metrics for the livestreams during that month.

3.4. Livestream content metrics

The platform generates a suite of metrics based on sellers' livestream activities, capturing dimensions such as content quality (content value score), fan purchasing behavior and loyalty (fans value score), comparative performance within a category (area value score), responsiveness to viewer interactions (seller service score), and the evolution of livestream and consumer activity over time (account health score). These scores are produced by the platform's automated algorithms and are updated weekly. For instance, the content value score is derived from automated analysis of livestream videos, incorporating factors such as video resolution, scripting, and other quality indicators. Similarly, the account health score integrates livestream content features with the historical trajectories of sellers' activity and consumer purchases. The documented input domains indicate that these metrics fundamentally center on livestream videos and evaluate their quality from different dimensions. In our empirical specification, we use these scores as recorded at the beginning of each observation month to ensure that no post-outcome information enters the predictive model.

Appendix A.4 presents a robustness check using lagged $(t-1)$ versions of all livestream content metrics. While this specification naturally relies on less timely information and therefore exhibits slightly lower predictive accuracy, it produces consistent qualitative results. Even if information leakage were a concern, the lagged specification would effectively mitigate it while maintaining strong predictive performance.

Table 5 gives the summary statistics of the session strategies, assortment strategies, social engagement metrics, and livestream content metrics at the seller-month panel level.

3.5. Seller online store characteristics

We monitor various characteristics of the sellers' online stores, which we use as important control variables in our analysis. These controls help us to account for variations in seller experience, reputation, and scale, enabling a more accurate evaluation of how livestream-related variables influence seller sales performance.

One key control is seller tenure, which is calculated based on when the seller first opened their online store. Additionally, we track seller livestream tenure, or the time since the seller began using the livestream shopping channel by holding their first livestream session. These variables help us account for the seller's experience both in the broader online marketplace and specifically in livestreaming, which may influence the seller's ability to engage audiences and drive sales. We also use seller credit rating as a control for the seller's historical transaction volume. This rating serves as a proxy for the seller's total number of completed transactions and reflects their accumulated experience on the platform. Higher credit ratings typically indicate more established sellers who may have developed stronger operational capabilities and customer relationships over time.

Table 5
Seller-month level summary statistics.

Variable	Mean	SD	Min.	Median	Max.
If Livestream	0.67	0.47	0	1	1
Total Livestream Session Length	89.23	162.61	0	14.01	1201.31
Average Livestream Session Length	4.096	5.322	0	2.346	24
Total Livestream Session Length (cumulative)	255.12	459.72	0.08	76.11	4640.22
Livestream Frequency	9.50	10.92	0	4	31
Livestream Frequency (cumulative)	30.18	31.62	1	19	152
Price	938.15	4061.76	0	54.50	84150.78
Variation of Price	0.52	0.64	0	0.36	6.86
Product Tenure	53.07	94.89	0	17.34	1440.45
Variation of Product Tenure	0.45	0.58	0	0.28	9.84
Product Save Popularity	2328.02	20301.46	0	0.25	1,319,040
Variation of Product Save Popularity	1.31	1.67	0	0.85	11.58
Product Category Count	1.17	2.30	0	1	66.63
Concentration of Product Category	0.57	0.44	0	0.78	1
Product Count	96.57	507.68	0	20.97	32365.93
Page View Count	1360.84	9693.49	0	35.88	466928.23
Likes Count	4785.10	41003.94	0	119.09	1980500.74
Comments Count	179.22	921.46	0	3.25	58,819
Shares Count	11.52	64.50	0	0.35	3312.03
Subscription Count	6.59	36.50	0	0.29	1666
UV	448.06	2352.29	0	13.50	106892.37
Likes UV	14.10	101.19	0	1	7007.77
Comments UV	19.22	90.13	0	0.93	3285.80
Shares UV	4.11	25.00	0	0.22	1972.90
Watch Time	2.29	6.90	0	0.95	777.34
Content Value Score	219.40	128.17	0	271.77	671
Fans Value Score	96.24	132.53	0	39.29	942.68
Area Value Score	647.28	261.37	80	794.97	1000
Seller Service Score	34.36	71.17	0	1.81	628.39
Account Health Score	160.90	178.41	0	98.83	993
Logged Sales	7.61	4.674	0	8.351	32.226
Number of Obs.	20,305				

Note: This table presents summary statistics for variables related to livestream strategies, social engagement metrics, and the outcome variable at the seller-month level. UV means unique viewer.

Another important control is seller positive rating, which measures the percentage of positive ratings among all transactions. This variable provides a snapshot of customer satisfaction and service quality, allowing us to adjust for the potential influence of a seller's reputation on their livestream engagement and sales outcomes. Additionally, we control for the number of subscribers each seller has during each time period. Subscribers are a key indicator of the seller's audience size and loyalty. By controlling for subscriber numbers, we can better isolate the effects of livestream strategies from the seller's pre-existing audience reach.

4. Predicting seller sales performance

In this section, we present a predictive model of seller sales performance based on the XGBoost algorithm.[7] XGBoost is a powerful, tree-based machine learning algorithm known for its high predictive accuracy and efficiency in handling large datasets. To ensure robust model evaluation, we randomly split the sellers into two groups: 80% for training and 20% for testing. The model is trained on the training set, and the testing set is reserved solely for assessing the model's out-of-sample predictive performance, ensuring that the evaluation is unbiased. For the training process, we employ 5-fold cross-validation, a commonly used technique for tuning hyperparameters and preventing overfitting. This method divides the training data into five subsets, training the model on four subsets while validating it on the fifth. This process is repeated five times, allowing each subset to serve as the validation set once. By averaging the performance across all five iterations, we can fine-tune the hyperparameters to achieve optimal model performance.

The training dataset consists of data points represented as $(x_{it}, y_{it}), i = 1, \ldots, N_{train}, t = t_i, \ldots, t_i + 4$, where t_i is the first month after seller i started to livestream, and $x_{it} \in R^D$ is a D-dimensional feature vector. This feature vector includes the relevant seller characteristics, session strategies, assortment strategies, and social engagement metrics that influence sales performance. The outcome variable y_{it} represents the logged sales performance of seller i at time t, which is treated as a continuous variable.

By using the XGBoost algorithm with this structure, we aim to capture the complex relationships between the input features and the sales outcomes, enabling a more accurate prediction of seller performance. To further assess temporal stability, in Section 4.3, we conduct an out-of-time validation by training the model on each seller's first four months of data and testing on the subsequent

[7] Appendix A.1 compares XGBoost with Random Forest, Ridge, and Support Vector Regression models (Table A1); XGBoost achieves the highest predictive accuracy.

Table 6
Model accuracy with different feature sets.

	Out-of-sample		
	MSE	MAE	R^2
Baseline	10.103	2.408	0.545
Baseline + Session Strategies	8.746	2.204	0.606
Baseline + Session Strategies + Assortment Strategies	8.220	2.123	0.629
Baseline + Session Strategies + Social Engagement Metrics	8.298	2.139	0.626
Baseline + Session Strategies + Livestream Content Metrics	7.285	1.963	0.672
All Predictors	6.994	1.901	0.685

Note: This table reports the model accuracy across different feature sets based on the test sample. MSE is mean squared error, and MAE is mean absolute error.

month. The model achieves comparable accuracy to the main specification, indicating that the predictive relationships are stable over time.

4.1. Sales performance prediction

To understand the relative importance of the different sets of livestream variables (session strategies, assortment strategies, social engagement metrics, livestream content metrics) in predicting seller sales performance, we compare the predictive power of the following model specifications: (1) the baseline model, which includes all controls related to the sellers' online stores and month dummies; (2) baseline combined with session strategies, which adds session strategies to the baseline; (3) baseline combines with session strategies and assortment strategies, incorporating both session and assortment strategies; (4) baseline combined with session strategies and social engagement metrics, which includes session strategies and social engagement metrics; (5) baseline combined with session strategies and livestream content metrics, which includes session strategies and livestream content metrics; and (6) all variables, which is a specification that comprises controls, session strategies, assortment strategies, social engagement metrics, and livestream content metrics. We do not include models combining the baseline directly with assortment strategies, social engagement metrics, or livestream content metrics because these variables are more meaningful when session strategies are considered. Assortment and engagement metrics are averaged across livestreams and depend on whether a session was held during the period.

For each model specification, we evaluate performance using mean squared error (MSE), mean absolute error (MAE), and R^2. Table 6 compares the predictive power of session strategies, assortment strategies, social engagement metrics, and livestream content metrics on seller sales performance. The baseline model, which includes only controls for store characteristics and macro factors, provides the least accurate predictions, explaining just over half of the variance in sales, with an R^2 of 0.545. When session strategies are added, predictive accuracy improves, and the R^2 increases to 0.606, reflecting an 11.2% improvement in the model's ability to explain sales variance. This highlights the importance of how sellers manage their livestream sessions.

Next, we examine how predictive accuracy improves as we add each set of variables, assortment strategies, social engagement metrics, and livestream content metrics to the baseline model with session strategies, one at a time. Adding assortment strategies increases the model's R^2 to 0.629, a 15.4% improvement over the baseline. This suggests that the selection and variety of products featured during livestreams contribute meaningfully to predicting sales. Adding social engagement metrics instead yields a similar improvement, with an R^2 of 0.626, highlighting the importance of audience interactions, such as likes, comments, and shares, in driving sales performance. Among the three, adding livestream content metrics leads to the greatest improvement: the model achieves the highest predictive accuracy, with an R^2 of 0.672, representing a 23.3% improvement over the baseline.

Finally, the model using all predictors yields the highest R^2 of 0.685, representing a 25.7% improvement over the baseline. These results demonstrate that a combination of these factors is most effective for understanding and predicting seller sales performance.

4.2. Prediction for different types of sellers

We further investigate whether the incremental predictive power of different sets of variables varies across different types of sellers. Our analysis focuses on two key factors: seller categories (Fashion, Jewelry, and Food) and seller tenure.

Table 7 shows that adding more variables improves the predictive power of the model across the three categories. Compared to the baseline, adding session strategies boosts prediction accuracy for the three categories, with the biggest increase in Jewelry (32.5%). We next examine how predictive accuracy improves when we add each set of variables, assortment strategies, social engagement metrics, and livestream content metrics to the baseline model with session strategies, one at a time, for each product category. For Fashion and Food, adding livestream content metrics leads to the highest accuracy improvement, with R^2 increasing by 22.5% for Fashion and 12% for Food. In contrast, adding either assortment strategies or social engagement metrics results in smaller but similar gains in accuracy. However, in Jewelry, social engagement metrics provide the biggest boost in predictive power, increasing R^2 by 47.5%, outperforming both assortment strategies and livestream content metrics. In summary, the Jewelry category sees the most significant improvements, with a 57.3% increase in R^2, whereas Fashion and Food show more modest but steady gains as more variables are added.

Table 7
Comparison of incremental predictive power by seller category.

	Out-of-sample R^2		
	Fashion	Jewelry	Food
Baseline	0.534 -	0.379 -	0.603 -
Baseline + Session Strategies	0.586 (+9.7%)	0.502 (+32.5%)	0.634 (+5.1%)
Baseline + Session Strategies + Assortment Strategies	0.606 (+13.5%)	0.530 (+39.9%)	0.641 (+6.3%)
Baseline + Session Strategies + Social Engagement Metrics	0.608 (+13.9%)	0.559 (+47.5%)	0.644 (+6.8%)
Baseline + Session Strategies + Livestream Content Metrics	0.654 (+22.5%)	0.543 (+43.2%)	0.675 (+12.0%)
All Predictors	0.665 (+24.5%)	0.596 (+57.3%)	0.688 (+14.1%)

Note: This table presents the comparison of incremental predictive power across seller categories, based on out-of-sample R^2.

Table 8
Comparison of incremental predictive power by seller tenure.

	Out-of-sample R^2	
	$\leq$ 3 months	> 3 months
Baseline	0.507 -	0.546 -
Baseline + Session Strategies	0.547 (+7.9%)	0.605 (+10.8%)
Baseline + Session Strategies + Assortment Strategies	0.577 (+13.8%)	0.626 (+14.7%)
Baseline + Session Strategies + Social Engagement Metrics	0.550 (+8.5%)	0.630 (+15.4%)
Baseline + Session Strategies + Livestream Content Metrics	0.584 (+15.2%)	0.676 (+23.8%)
All Variables	0.612 (+20.7%)	0.689 (+26.2%)

Note: This table presents the comparison of incremental predictive power by seller tenure, based on out-of-sample R^2. Sellers are divided into two groups based on whether they had their online shop for at least three months prior to their first livestream session.

Table 8 shows how the model's predictive power improves as more variables are added, with separate analyses for short-tenure sellers ($\leq$ 3 months) and longer-tenure sellers (> 3 months). For short-tenure sellers, the baseline model starts with an R^2 of 0.507. Adding session strategies increases the R^2 by 7.9%. We then examine how predictive accuracy changes when we add each set of variables, assortment strategies, social engagement metrics, and livestream content metrics to the baseline model with session strategies, one at a time. Adding assortment strategies increases accuracy by 13.8%. Livestream content metrics yield the largest improvement, raising the R^2 by 15.2% (to 0.584), suggesting that content quality plays a key role in predicting sales for newer sellers. Social engagement metrics lead to a smaller gain of 8.5%, indicating that audience interaction is less influential for short-tenure sellers. The full model, which includes all variables, achieves the highest R^2 of 0.612, representing a 20.7% improvement over the baseline.

For longer-tenure sellers, the baseline model starts with a higher R^2 of 0.546. Adding session strategies improves the R^2 by 10.8%. Building on this, we assess the added value of each group of variables, assortment strategies, social engagement metrics, and livestream content metrics, when included individually alongside session strategies. Incorporating assortment strategies further raises the R^2 by 14.7%, suggesting that product selection remains important for experienced sellers. Social engagement metrics also contribute meaningfully, with the R^2 rising to 0.630, a 15.4% improvement over the baseline. Livestream content metrics deliver the largest gain for this group, increasing the R^2 by 23.8% (to 0.676), highlighting the critical role of content quality and relevance. The full model, which includes all variables, achieves the highest R^2 of 0.689, representing a 26.2% improvement.

Overall, the findings suggest that for short-tenure sellers, assortment strategies and livestream content metrics are more predictive than social engagement metrics, whereas for longer-tenure sellers, livestream content metrics play a more important role in improving predictive accuracy. Both groups experience significant gains when all variables are included, but the greatest improvement is seen in longer-tenure sellers.

Table 9
Out-of-time validation: model performance on held-out month.

	Out-of-time sample		
	MSE	MAE	R^2
All Predictors	7.339	2.018	0.719

Note: This table reports the model accuracy based on the test sample which shows the out-of-time validation.

4.3. Out-of-Time validation

To assess whether the predictive model remains stable when applied to future periods, we conduct an out-of-time validation. For each seller, we train the model using the first four months of observations and evaluate predictive accuracy on the subsequent (fifth) month. To ensure that the performance is not mechanically driven by calendar-month patterns, we remove calendar-month fixed effects from the predictors. Table 9 reports the performance on the out-of-time test sample.

The out-of-time performance is highly comparable to the original (random) train-test split, indicating that the model retains its predictive power when applied to unseen future periods, and the score is temporally stable rather than overfitting contemporaneous patterns.

5. Enhancing decision support with SHAP

Our goal is to build a scoring model that not only predicts seller sales performance accurately but also offers interpretability into how those predictions are made. We use the XGBoost algorithm because of its strong predictive performance and computational efficiency. However, like many machine learning models, XGBoost operates as a “black box” and does not naturally reveal how individual features influence predictions. To address this limitation, we incorporate SHAP (SHapley Additive exPlanations; Lundberg and Lee 2017), which allows us to interpret feature contributions at the seller level. SHAP enables us to break down the model’s predictions and assign importance scores to each feature, supporting both accuracy and transparency.

SHAP computes Shapley values based on coalitional game theory, treating each feature in a data instance as a “player” in a coalition. These values reflect how the model’s prediction is fairly distributed across features, indicating each feature’s contribution to the final outcome. Features with larger absolute Shapley values exert greater influence on the prediction. This approach enhances the XGBoost model by not only predicting sales performance but also quantifying the role of each input feature. By applying SHAP, we offer stakeholders a clear view of how factors such as session strategies, assortment choices, social engagement, and livestream content contribute to seller performance. This added transparency supports trust in the model’s results and facilitates more informed strategic decisions.

We then identify the top predictors of sales performance in the model that includes all variables by calculating the mean absolute Shapley value of each feature across all sellers. Fig. 1 displays the features with a mean absolute SHAP value of at least 0.05, ranked in descending order within each variable set.

As expected, sellers’ online store characteristics such as credit rating, subscriber count, and tenure emerge as one of the most significant predictors of sales performance. These inherent characteristics are key drivers of sales outcomes, as they are closely aligned with a seller’s overall market presence and credibility.

Additionally, livestream content metrics are key contributors to the model’s predictive performance. Seller service score has the highest impact, followed by fans value score and content value score. Livestream strategies also matter. Total session length, both cumulative and per session, and livestream frequency, are strong predictors. Assortment-related variables such as price, variation in product tenure, product count, and product save popularity are also important. These findings suggest that strategic decisions made during livestreams, such as how products are priced and when sessions are held, are closely linked to sales performance. Social engagement metrics, especially watch time and the number of unique visitors who click “like” or comment, add further predictive value, highlighting the role of audience interaction and engagement in driving sales.

We also examine the individual feature importance plot to better understand how feature values relate to their impact on predictions. Figure 2 combines feature importance with feature effects, showing how different values of each feature influence the model’s output. Consistent with Fig. 1, we include features with a mean absolute SHAP value of at least 0.05. Each point represents the Shapley value of a feature for one seller. The vertical position reflects the feature’s overall importance, and the horizontal position indicates its Shapley value. The color of each point shows the feature value, ranging from low (red) to high (green). To reveal the distribution of Shapley values across sellers, overlapping points are jittered vertically.

Among the livestream content metrics, seller service score stands out as the most influential predictor and is clearly associated with higher predicted sales. In contrast, the effects of fans value score and content value score are less consistent. The SHAP plots show that even at high levels, these scores do not reliably contribute positively to sales predictions. This suggests that simply having a large fan base or producing high-quality content is not sufficient. Their impact likely depends on how well these elements are integrated with other aspects of the seller’s strategy, such as timing, session structure, or audience engagement.

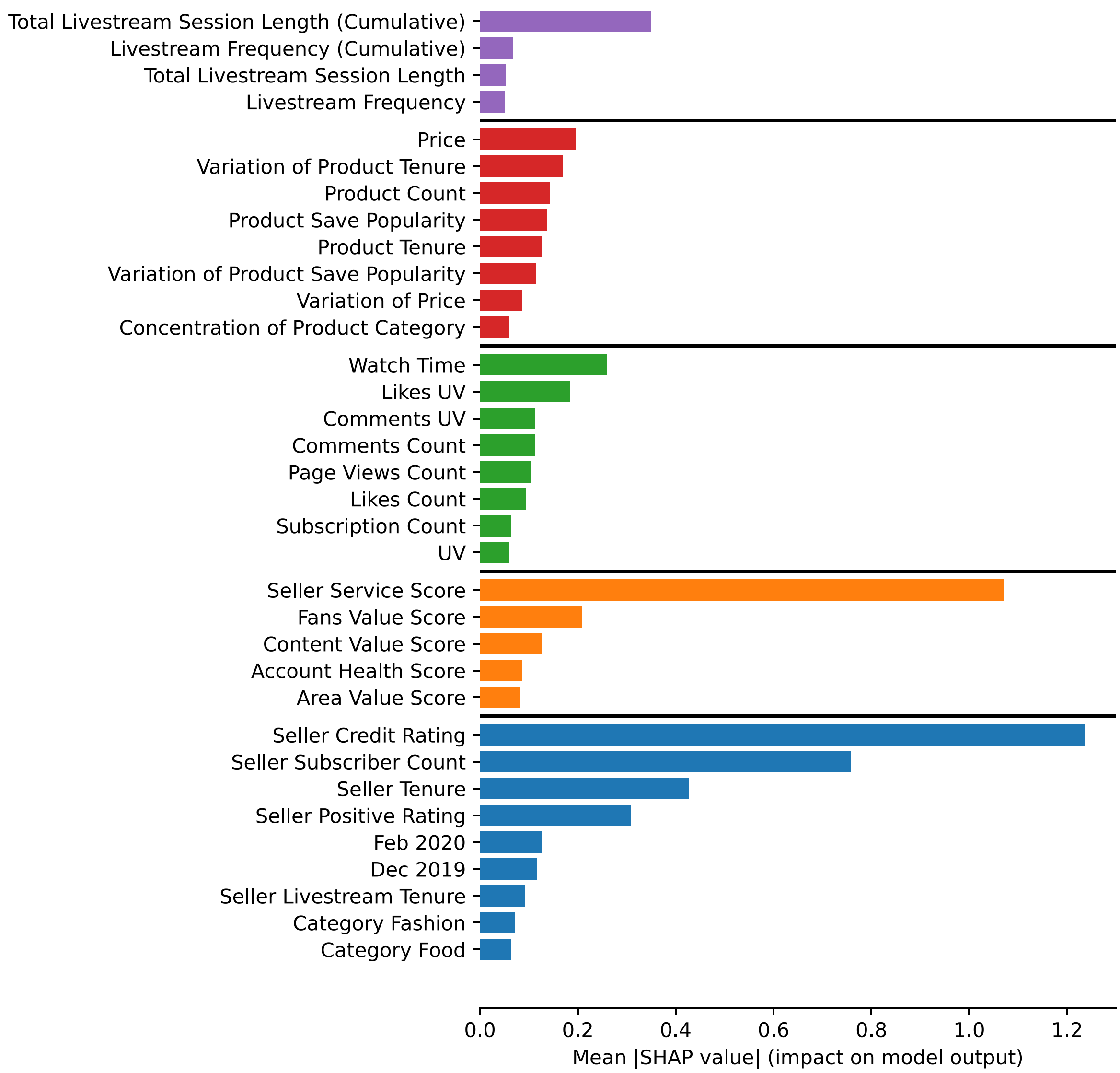

Fig. 1. Features ranked in order of importance for seller sales performance.

In our analysis of livestream strategies, we explore both session strategies and product assortment decisions to assess their impact on sales performance. The results indicate that certain aspects of livestream session management are key drivers of sales. Specifically, the frequency of livestreams and the total length of livestream sessions within a given month are positively associated with higher sales. This suggests that consistent and frequent livestream engagement with the audience can significantly improve sales outcomes. However, we also observe an interesting contrast: the cumulative length of all livestreams conducted since the seller began using the channel shows a negative association with sales. This finding suggests that while regular livestreaming in the short term is beneficial, extending session lengths over a prolonged period may lead to diminishing returns. This could be attributed to factors such as audience fatigue or content oversaturation, where excessive long-term broadcasting reduces its effectiveness in driving sales.

Our analysis of assortment strategies highlights several patterns linked to stronger sales performance. Livestreams featuring higher-priced products tend to perform better, suggesting that viewers respond more positively when the perceived product value is higher. While products with longer tenure are generally linked to lower sales, greater variation in product tenure within a session is positively associated with performance. A mix of newer and older products likely keeps viewers more engaged. Sessions that include products with moderately high save popularity, meaning items that are liked but not overly saturated, also see better outcomes. Livestreams that present more products per hour also perform well, suggesting that a quicker, more varied presentation helps maintain audience attention. In addition, variation in product save popularity contributes positively to sales. Sessions that feature a mix of popular

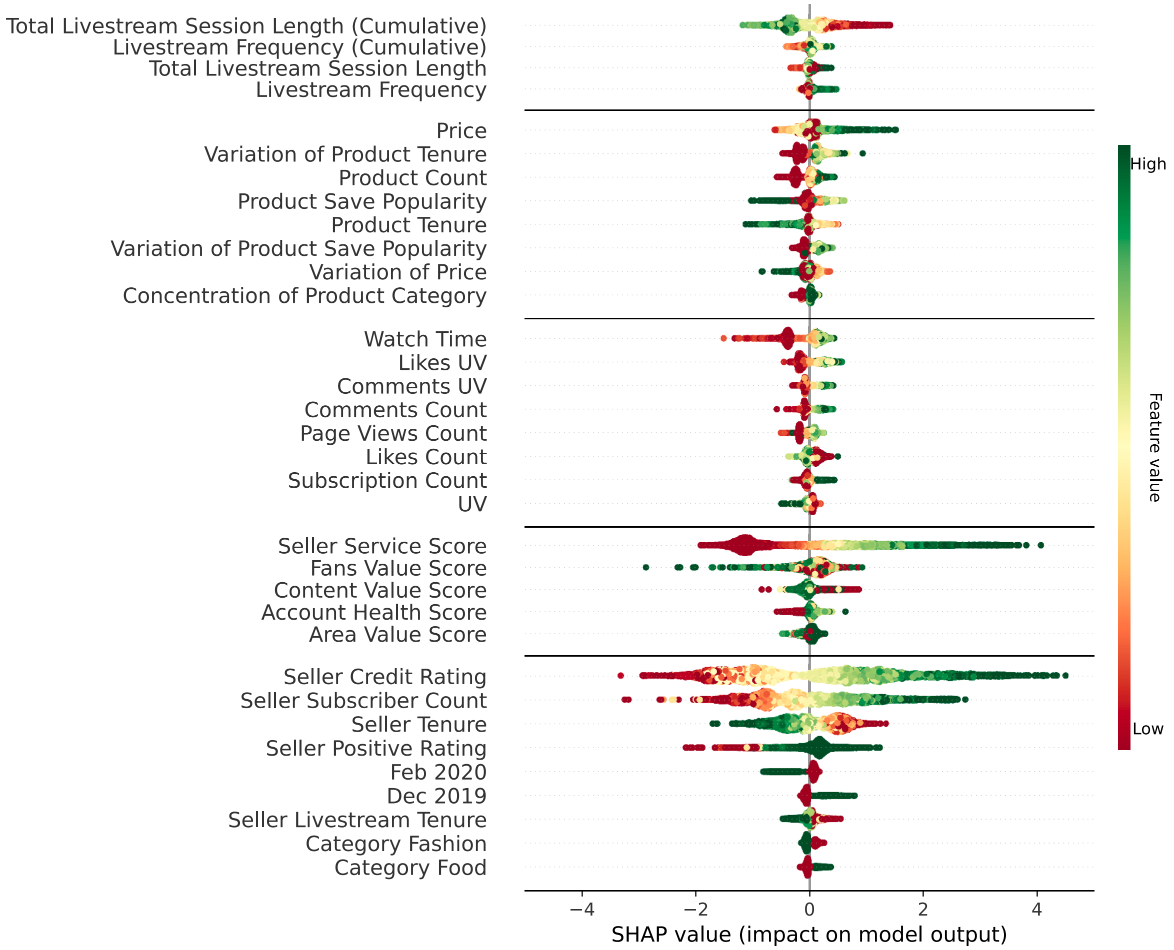

Fig. 2. Features ranked in order of importance for seller sales performance with feature value.

and less popular items tend to attract a broader audience by balancing mass appeal with a sense of discovery. In sum, successful assortment strategies include a range of prices, levels of popularity, and product tenures, delivered in a format that keeps the session lively and engaging.

For social engagement metrics, we find that most variables are positively associated with sales, though their predictive strength varies. Watch time stands out as the most important metric, highlighting the role of sustained viewer attention in driving sales performance. The unique number of viewers who liked the content is the second most influential predictor and proves to be more informative than the total number of likes. Similarly, the unique number of viewers who commented is among the more predictive metrics, emphasizing the importance of individual-level engagement. These results suggest that distinct user interactions, rather than repeated actions from the same user, are more closely tied to sales outcomes. Taken together with the findings on comments, it becomes evident that meaningful engagement is a stronger signal of sales potential. In contrast, the number of shares is the least informative metric, which has a mean absolute SHAP value less than 0.05, indicating that while sharing may increase exposure, it does not directly translate into higher sales.

For the control variables, our analysis shows that some metrics offer more meaningful insight into sales performance than others. Seller credit rating, which serves as a proxy for past performance, and the number of subscribers are the most informative indicators. Both are strongly and positively associated with current sales, suggesting that seller credibility and a sizable audience base are important drivers of success. In contrast, the tenure of the livestream channel and the seller's overall tenure are negatively associated with sales. While this finding may seem counterintuitive, it reflects the nature of the monthly sales data. Sellers or channels with shorter operating histories may have achieved high sales volumes in a more condensed time frame, pointing to recent or rapidly growing success. These sellers may also be more aggressive or benefit from early visibility. On the other hand, those with longer histories may face slower growth or strategic shifts that reduce short-term sales performance.

The SHAP results are robust to alternative model specifications. Using quantity sold or the log percentage change in sales as the outcome yields similar feature rankings and effects (Appendix A.2). Using an alternative dataset with livestream promotion information (Appendix A.3) also produces consistent SHAP patterns, indicating that the main predictors remain stable across specifications and datasets.

6. Conclusion

This paper proposes a framework to predict the sales performance of livestream shopping sellers and examines the key predictors of sales performance, moving beyond traditional online store metrics. We focus on identifying the session strategies, assortment strategies, livestream content metrics, and social engagement metrics that are most strongly associated with sales. Understanding these factors is crucial in the livestream shopping industry, where the combination of e-commerce and live interaction highlights its main character.

We collect data on livestream activities, sales, and basic online store characteristics from Taobao Live. The dataset spans five months of observations following each seller's first livestream session. We use XGBoost to predict sales performance based on sellers' livestream strategies (including session and assortment strategies), livestream content metrics, and social engagement metrics in addition to the online store's characteristics. We then use SHAP to interpret how each feature contributes to predicting sales performance.

We demonstrate that incorporating session strategies, assortment strategies, livestream content metrics, and social engagement metrics into predictive models significantly improves forecasting accuracy, with a 25.6% increase in predictive power compared to traditional online store characteristics alone. This improvement is particularly notable for sellers in the Jewelry category. Our analysis reveals that different livestream strategies, livestream content metrics, and engagement metrics matter depending on seller tenure. Specifically, livestream content metrics and assortment strategies are more critical for newer sellers, whereas livestream content metrics are stronger predictors of success for more experienced sellers. Furthermore, livestream sessions featuring higher-priced and moderately popular products, and diverse product offerings with variations in popularity, tenure, and presentation rate, are associated with better seller sales performance. For livestream content metrics, a higher seller service score is associated with higher sales performance. For social engagement metrics, only meaningful interactions, such as comments, watch time, and unique number of visitors who reacted with a like, emerge as strong predictors of sales.

Our results have implications for various stakeholders, including investors, lenders, suppliers, livestream shopping platforms, and sellers. Investors can use the scoring model to predict firm success, optimizing their portfolios by selecting promising firms and avoiding riskier ones, whereas lenders can assess credit risk and adjust loan terms to minimize defaults. Suppliers can evaluate potential partners' reliability, ensuring long-term collaboration, and livestream shopping platforms can enhance user experience by integrating the model into their recommender systems. Additionally, insight into key success indicators benefits analytics providers, investors, and sellers, enabling better monitoring, benchmarking, and strategic adjustments to improve performance and minimize risk.

Our paper has several limitations that open avenues for future research. First, we focus on predictive analytics rather than causal inferences. While our findings offer valuable insights to guide sellers in optimizing their strategies, they do not establish direct causal relationships. For example, our model highlights which livestream strategies and social engagement metrics are associated with better sales performance. However, this merely indicates correlation, not causation, providing actionable insights by pointing out key areas that correlate with higher sales. In this sense, our predictive model serves as a practical tool for sellers seeking to improve their sales outcomes by focusing on these strategies and metrics. Future research could explore whether any of the leading indicators have a causal impact on seller sales performance. Second, our data come from only one livestream shopping company, Taobao Live, which is connected with the large e-commerce platform Taobao. Future research can test the generalizability of our findings by using data from other livestream shopping platforms, such as Facebook and TikTok, which emphasize the social network and entertainment aspects rather than the e-commerce aspect.

CRediT authorship contribution statement

Zekun Liu: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Weiqing Zhang:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Xiao Liu:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Eitan Muller:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Feiyu Xiong:** Formal analysis, Data curation, Conceptualization.

Data availability

The data that has been used is confidential.

Appendix A. Robustness

A.1. Alternative methods

We estimated alternative models, including Random Forest, Ridge Regression, and Support Vector Regression, to provide benchmarks and assess robustness. Table A1 reports the performance of these models alongside that of the XGBoost model. The results show that XGBoost performs best, with the lowest mean squared error (MSE; 6.994), the lowest mean absolute error (MAE; 1.901), and the highest R^2 (0.685), indicating stronger predictive accuracy in this context.

Table A1
Model Performance With Alternative Prediction Method.

	Out-of-sample		
	MSE	MAE	R^2
XGBoost	6.994	1.901	0.685
Random Forest	8.105	2.133	0.635
Ridge Regression	8.592	2.214	0.613
Support Vector Regression	8.827	2.169	0.602

Note: This table presents the performance of different prediction models based on out-of-sample data. The models compared are XGBoost, Random Forest, Ridge Regression, and Support Vector Regression, and the performance metrics are mean squared error (MSE), mean absolute error (MAE), and R^2.

A.2. Alternative model specification

First, instead of using revenue as the outcome measure, we also examine quantity sold as an alternative outcome. Fig. 2 presents the SHapley Additive exPlanations (SHAP) values for key features predicting quantity sold. The overall results remain consistent with those in Figure 2. Livestream content metrics, session strategies, and seller store characteristics continue to rank among the most important predictors, reinforcing the robustness of our findings. One notable difference is that price is not consistently positively associated with quantity sold. This finding indicates that in some cases, higher sales revenue may be driven primarily by higher prices rather than higher volume. However, because total sales reflect both price and quantity, the decrease in quantity is generally not large enough to offset the effect of higher prices, resulting in increased overall revenue.

Next, we conducted a robustness check using an alternative specification in which the target variable, monthly sales, is expressed as the log percentage change. For the predictors, we applied log percentage transformations where appropriate, such as for livestream session length, frequency, seller credit rating, subscriber count, and seller ratings. The variables for which such transformations were not meaningful were kept in their original scale. As shown in Fig. A1, the results remain consistent. Livestream content metrics, session strategies, and seller store characteristics continue to be among the most influential predictors, reinforcing the robustness of our findings. The log percentage changes in seller rating, subscriber count, and seller service score are all positively associated with changes in sales. Similarly, increases in livestream frequency and total session length are also positively correlated with sales growth.

A.3. Alternative dataset with livestream promotion information

We examine the livestream promotion strategy and how it predicts seller sales performance. Using our transaction-level data (for a subset of sellers and a portion of all transactions), we analyze how sellers run promotions during livestreams. For each livestream, we identify the products featured. For products with sales both through the livestream and the regular online store, we calculate their average promotion level. We then take the average promotion across all featured products we can observe to approximate the livestream's promotion level. Focusing on sellers with at least one such measure, we run a robustness check by adding this variable to our prediction model and calculating its SHAP value.

There are 2741 observations in the training data and 596 in the testing data, when using the full set of features in this analysis. Fig. A2 displays features with a mean absolute SHAP value of at least 0.05, ranked in descending order within each variable set (Fig. A3).

The results show that livestream promotion does not stand out as a strong predictor of seller sales performance. Although higher promotion levels are positively correlated with seller sales at the individual SHAP value level, the overall SHAP values remain quite small. While moderate discounting under favorable conditions can enhance revenue, overly aggressive discounts may compress margins beyond sustainable levels. Furthermore, promotional effectiveness is likely mediated by a set of contextual factors, such as product category, baseline pricing, consumer segment, and livestream host characteristics, that vary widely across sellers. These interacting forces can generate both positive and negative outcomes from livestream promotions, which may cancel out in the aggregate and thereby attenuate the observable predictive power of promotion-related variables in our empirical model.

However, we acknowledge that this analysis is limited to a selected subset of sellers and transactions where we can recover promotion data, and this subset may not be representative of the broader sample. One limitation of this approach is that it allows us to identify only promotions specific to livestream sessions. If sellers offer promotions simultaneously across both their online stores and livestreams, we are unable to distinguish those cases.

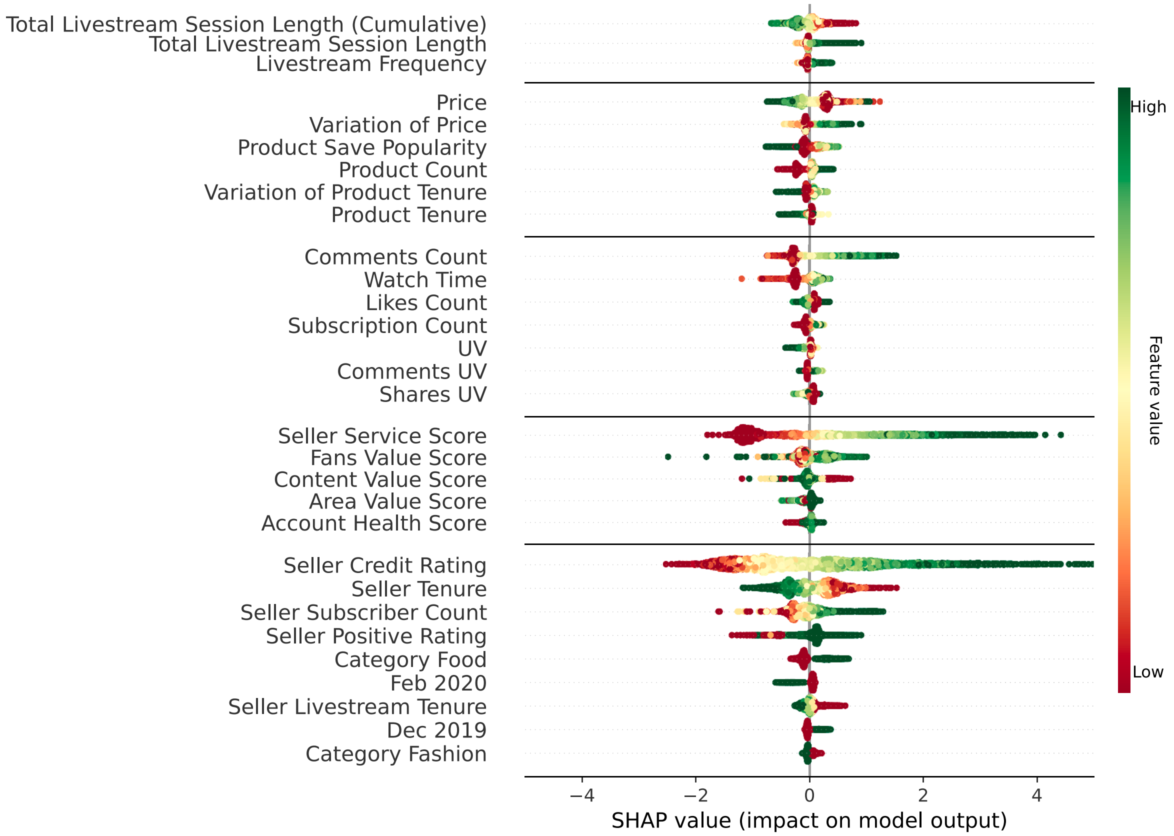

Fig. A1. Features ranked in order of importance for seller sales quantity with feature value.

Table A2
Model Accuracy With Different Feature Sets—Lagged Livestream Content Metrics.

	Out-of-sample		
	MSE	MAE	R^2
Baseline	10.103	2.408	0.545
Baseline + Session Strategies	8.746	2.204	0.606
Baseline + Session Strategies + Assortment Strategies	8.220	2.123	0.629
Baseline + Session Strategies + Social Engagement Metrics	8.298	2.139	0.626
Baseline + Session Strategies + **Lagged Livestream Content Metrics**	7.911	2.063	0.643
All Predictors **With Lagged Livestream Content Metrics**	7.386	1.964	0.667

Note: This table reports the model accuracy across different feature sets based on the test sample. MSE is mean squared error, and MAE is mean absolute error.

A.4. Lagged livestream content metrics (leakage robustness)

To further ensure that no post-outcome information leaks into the training process through platform-generated content variables, we re-estimate the predictive model using lagged ($t-1$) versions of all livestream content metrics. The platform updates these metrics weekly, and in the main specification we have used the values as recorded at the beginning of each observation month; the lagged specification introduces an even stricter temporal buffer. Table A2 reports performance with lagged livestream content metrics. The lagged specification delivers slightly lower predictive power, but the relative feature contributions and substantive conclusions remain consistent.

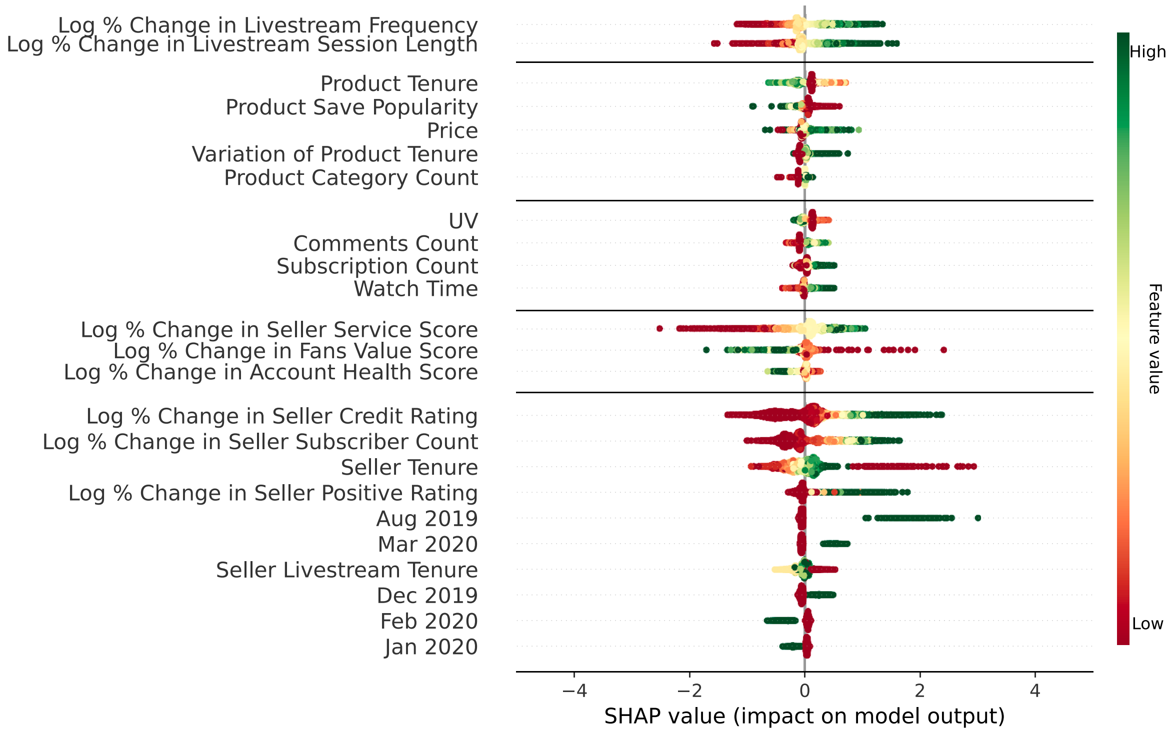

Fig. A2. Features ranked in order of importance for sellers log percentage change in sales with feature value.

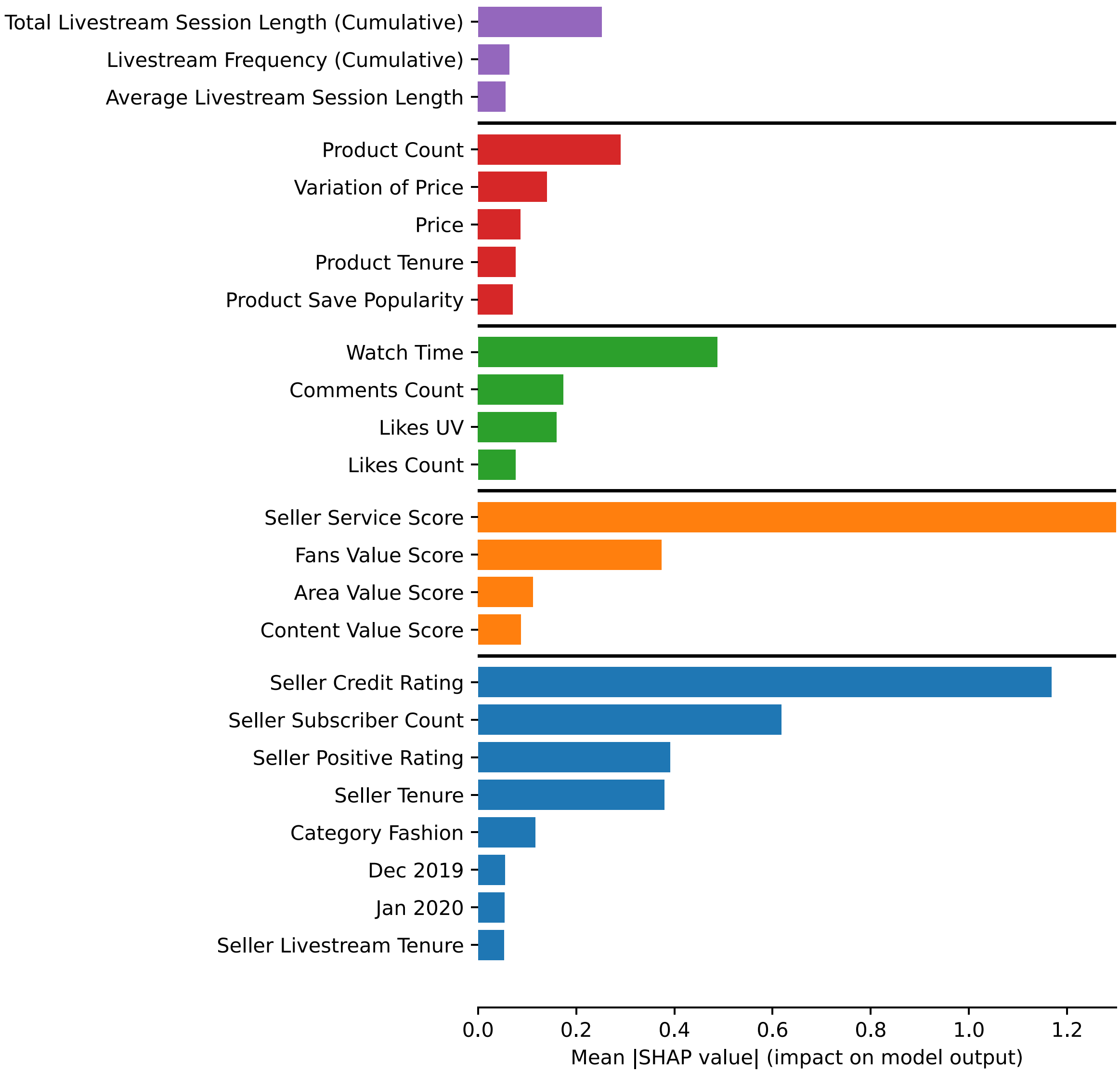

Fig. A3. Features ranked in order of importance for sellers log percentage change in sales with feature value—subset with livestream promotion.

References

Audretsch, D. B., Houweling, P., & Thurik, A. R. (2000). Firm survival in the netherlands. *Review of Industrial Organization*, *16*(1), 1–11.

Bharadwaj, N., Ballings, M., Naik, P. A., Moore, M., & Arat, M. M. (2022). A new livestream retail analytics framework to assess the sales impact of emotional displays. *Journal of Marketing*, *86*(1), 27–47.

Dekimpe, M. G., & Morrison, D. G. (1991). A modeling framework for analyzing retail store durations. *Journal of Retailing*, *67*(1), 68.

Gu, X., Zhang, X., & Kannan, P. K. (2024). Influencer mix strategies in livestream commerce: impact on product sales. *Journal of Marketing*, *88*(4), 64–83.

Hilvert-Bruce, Z., Neill, J. T., Sjöblom, M., & Hamari, J. (2018). Social motivations of live-streaming viewer engagement on twitch. *Computers in Human Behavior*, *84*, 58–67.

Jiang, D. (2020). The end of influencer livestreaming e-commerce? scandal and sales plummet. https://www.163.com/dy/article/FKREBSNI0519H3QD.html.

John, L. K., Emrich, O., Gupta, S., & Norton, M. I. (2017). Does "liking" lead to loving? the impact of joining a brand's social network on marketing outcomes. *Journal of Marketing Research*, *54*(1), 144–155.

Li, X., & Han, R. (2024). Streaming consumers in livestreaming e-commerce: identifying lifts from sales pitches. *working paper*.

Lin, X., Gui, L., & Lu, Y. (2024). Managing sales via livestream commerce: Implications of price negotiation and consumer price search. *Production and Operations Management*.

Lin, Y., Yao, D., & Chen, X. (2021). Happiness begets money: Emotion and engagement in live streaming. *Journal of Marketing Research*, *58*(3), 417–438.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 4768–4777).

Sun, L., & Tang, Y. (2024). Avatar effect of AI-enabled virtual streamers on consumer purchase intention in e-commerce livestreaming. *Journal of Consumer Behaviour*, *23*(6), 2999–3010.

Tucker, C. E. (2015). The reach and persuasiveness of viral video ads. *Marketing Science*, *34*(2), 281–296.

Van Heerde, H. J., Dinner, I. M., & Neslin, S. A. (2019). Engaging the unengaged customer: The value of a retailer mobile app. *International Journal of Research in Marketing*, *36*(3), 420–438.

Voorveld, H. A. M., Araujo, T., Bernritter, S. F., Rietberg, E., & Vliegenthart, R. (2018). How advertising in offline media drives reach of and engagement with brands on facebook. *International Journal of Advertising*, *37*(5), 785–805.

Wang, Z., Wang, Y., Yang, S., & Liu, H. (2022). Sales impacts of live-streaming video ads. *working paper*.

Wongkitrungrueng, A., & Assarut, N. (2020). The role of live streaming in building consumer trust and engagement with social commerce sellers. *Journal of Business Research*, *117*, 543–556.

Wongkitrungrueng, A., Dehouche, N., & Assarut, N. (2020). Live streaming commerce from the sellers' perspective: Implications for online relationship marketing. *Journal of Marketing Management*, *36*(5–6), 488–518.

Wymer, S. (2019). Social media and live streaming: reaching and engaging fans?: the case of Facebook Live and the Queensland Maroons. Ph.D. thesis. Massey University.

Yang, J., Zhang, J., & Zhang, Y. (2025). Engagement that sells: Influencer video advertising on tiktok. *Marketing Science*, *44*(2), 247–267.

Zhang, M., & Luo, L. (2023). Can consumer-posted photos serve as a leading indicator of restaurant survival? evidence from yelp. *Management Science*, *69*(1), 25–50.

Zhang, S. K., Tang, T. Y., & Krallman, A. (2024a). Navigating livestream commerce: A dual-lens framework of influencers' impact on product sales. *Journal of Business Research*, *185*, 114919.

Zhang, W., Liu, Z., Liu, X., & Muller, E. (2024b). The power of livestream shopping: Boosting revenues and catalyzing spillovers. *Available at SSRN 4318978*.