

-RESEARCH ARTICLE-

VISITORS AT AMAZON.COM: A TWO-LEVEL EXPLORATION INTO THEIR BROWSING AND SPENDING BEHAVIORS

Chatpong Tangmanee*

Chulalongkorn Business School, Chulalongkorn University

Email: chatpong@cbs.chula.ac.th

<https://orcid.org/0000-0001-6805-2921>

Chaiyasut Jongtavornvitaya

Chulalongkorn Business School, Chulalongkorn University;

Email: chaiyasut@cbs.chula.ac.th

<https://orcid.org/0000-0003-4492-7046>

—Abstract—

Amazon.com has been investigated extensively. However, research rarely addresses behaviors at the fine level of visit sessions, aggregating them into broad levels of visitors to draw significant conclusions. Our goals are thus (1) to compare the browsing behaviors (a) between the sessions, and (b) between the visitors, with and without purchase experience; and (2) to explore at both levels whether browsing behaviors could explain spending behavior. In order to achieve the research objectives and to test the study hypotheses, data from 1,812,569 usable visit sessions at Amazon.com were converted into 79,696 unique visitors. The t-tests verify that visits are relatively rushed when a purchase is made. Further, the explanatory effect of browsing behaviors on basket value was confirmed using regression analysis that concludes the number of visits significantly helps explaining the basket value in the visit level.

Keywords: Visitors; Amazon.com; Browsing; Spending; Two-Level Exploration.

JEL Classification: L81, M15

1. INTRODUCTION

The growth of electronic business and online transactions has been evident worldwide, especially during the lockdown for the Covid-19 pandemic. It is predicted that total

Citation (APA): Tangmanee, C., Jongtavornvitaya, C. (2022). Visitors at Amazon.Com: A Two-Level Exploration into Their Browsing and Spending Behaviors. *International Journal of eBusiness and eGovernment Studies*, 14 (1), 225-245. doi:10.34111/ijepeg.202214112

online retail sales worldwide will reach 28.47 trillion US dollars in 2022 (Kosower, Maybee, & O'Connell, 2019) with online grocery sales trending in the same direction. In 2022, the total sale number of online transactions will reach 49.58 billion US dollars. This is consistent with the overall picture of online sales in the US. Amazon.com has been listed among the top three most recognized retail websites (Gustafson, 2017; Statista, 2019). In 2017, 11% of online shoppers claimed to have made online transactions once a week using various online platforms. In 2018, Amazon reported that their sales were up 39% in the second quarter, as compared to the same period in the year before (Salinas, 2018). In addition, research predicts that the change in the annual growth of Amazon's retail sales will be 2% higher in 2021 than 2020 (Clement, 2020a). The use of mobiles for online retail transactions has also been increasing at a rapid pace (Clement, 2020b). Omanis had a relatively large number of online purchases in 2020 as compared to the previous years (AL-Hawari, Balasa, & Slimi, 2021).

Given its reputation, Amazon has been researched in various publications (Li, Wang, & Lin, 2018; C Tangmanee, 2019b; Wu, 2021). Using panel data from ComScore, Tangmanee (2017) confirmed that visitors who made a purchase during their visits to retail websites viewed fewer pages than those who made no purchases. The great amount of research has examined browsing and spending behaviors at retail websites in general (Nguyen, Armoogum, & Nguyen Thi, 2021; Zavali, Lacka, & Smedt, 2021), or at Amazon in particular (Panagiotelis, Smith, & Danaher, 2014; Chatpong Tangmanee, 2017). A classification using visit behavior at a UK-based apparel retail website validated that the mobile window shoppers hold the largest segment of the customers with the lowest amount of revenue (Zavali et al., 2021). The "goal-direct" customers identified in Creedy, Sidebotham, Gamble, Pallant, and Fenwick (2017) were also confirmed in Zavali, et al. (2021). Labelled as the visitors with a purpose, these goal-direct customers are highly profitable (Zavali et al., 2021).

Despite this sizable volume of previous research previously cited, there is still opportunity for further research. The critical barrier to investigate online behavior comes from the fact that access to data on such behavior is difficult to obtain. Two explanations are behind this difficulty. First, it is perhaps invalid or unreliable to gather such online behavior using a survey questionnaire as a participant's reply to a request asking how long the participant visits a given website would be subject to skepticism due to the human inability to accurately recall events. Moreover, participants may conceal details. These factors may therefore lead researchers to have different conceptualization of consumers' online behavior, preventing the use of a questionnaire as a means for data collection. Xu, Qi, and Li (2018), for example, adjusted from previous research the questionnaire scales used to measure visit duration. Although acceptable, the scales measured the visitor's perception of the length of their stay in a website, so the resulting data was not the actual duration of the visit.

Second, the findings from the few research projects using programming scripts or clickstream data services to record the browsing or spending behaviors at online retail stores may only cover the consumer's behaviors during a specific visit session. The behaviors of the visitors as individuals may not be correct, thereby limiting the investigation's generalizability. Given a conversion rate as low as 2% (Zhu, Wu, Wang, Cao, & Cao, 2019), it should not be a surprise that researchers may have overlooked the basket value or the visitor's spending behavior; in fact, many researchers have focused on explaining online behavior or using it to identify distinctive clusters of visitors (Creedy et al., 2017; Di Fatta, Patton, & Viglia, 2018; Mallapragada, Chandukala, & Liu, 2016).

For projects that addressed basket value and online behavior using clickstream panel data, their operationalization of the behavior was still unclear. Typically, clickstream data services would record online behavior from one visit session; this session may have only included purchase details if the visitor made a purchase. However, one visitor could have visited multiple times. Hence, the same behavior should be summed up for an examination at the visitor (or individual) level as it is broader than the session level. Nonetheless, this is still ambiguous as previous work has addressed the browsing behavior at the larger visitor level. Panagiotelis, et al. (2014) investigated online sales as a function of pageviews and visit duration; however, their work only provides results from the session level. Other researchers incorporated variables at different levels and shaded new light on visitors' behavior at retail websites (Di Fatta et al., 2018); for example, Sychov and Bakaev (2020) analyzed the extent to which constructs from three levels (i.e., the visitor, the website, and the visit occasion) affected browsing behavior. However, the study is unclear on how the data at the session level were aggregated to the visitor level. Such ambiguous detail might account for the inconclusive findings across websites or might not allow for replication.

Given the serious problem raised when the previous work has attempted to explain the browsing and the spending behaviors online but their data appear to have the quality issues, the current study relied on a clickstream data service to extract the visit-session data at Amazon.com, rolled up the data to the visitor level and performed an empirical analysis that addresses the three following objectives. First, we reported the browsing and spending behaviors at both the session and visitor levels. Second, we compared the behaviors (1) at the session level between the sessions with a purchase and the sessions with no purchase and (2) at the visitor level between those visitors who had made at least one purchase and those who had not made a purchase. Finally, we explored at both levels the extent to which browsing behaviors affect spending behavior.

2. LITERATURE REVIEW

2.1 Online retail business and background of Amazon.com

Online retail has increased significantly worldwide. Its sales volume has reached 2.3 trillion US dollars in 2019 and was projected to be 4.9 trillion in 2021 (Statista, 2017b). The visit sessions at retail websites have also grown substantially (Creedy et al., 2017; Zavali et al., 2021). A survey in Vietnam indicated that the residents prefer to do online shopping for health safety reason (Nguyen et al., 2021). Although 97% of the visit sessions ended with no purchase, these visits still hold business value for their traffic information. According to Huang, et al. (2015), the longer the stay at a unique website is, the more likely the visitors will revisit that specific website. Using machine learning techniques on the data from one German retail website, Raush, et al. (2021) were able to confirm the great number of shopping cart abandonments among new visitors while those with a purchase had high pageviews and had more selected items in their carts.

Amazon.com is currently one of the most well-known retail websites. Established as a bookstore in 1955 before going online in 1994, Amazon has expanded its business to cover other retail items such as cosmetics or electronic appliances, earning high ranks among the world's top leading online retail stores (Statista, 2017a). As such, it went from producing one million US dollars in sales in 2016 to ten billion in 2018, while competitors such as Walmart grew to only five billion dollars in sales during the same period (Salinas, 2018). The majority of Amazon's revenues have been from retail product sales, retail subscription services including Amazon Prime or AWA, and electronic book reading machines (i.e., Amazon Kindle). Its revenue is predicted to reach 356 billion US dollars in 2022 (Statista, 2019). Also, the number of Amazon customers has grown from only 200,000 customers in 2016 to over one million in 2017. Clement (2020a) predicted that global retail electronic commerce sales at Amazon will reach 404.44 billion US dollars in 2020, which is roughly a 64% increase from the same figure in 2017. Obviously, all figures reflect the remarkable success of Amazon.

2.2 The Behaviors of Visitors at Online Retail Websites

Visitors' online behaviors at retail websites including Amazon has been the main interest of several research projects; the behaviors studied primarily consist of the visitors' browsing and spending behaviors. In the current study, the number of pages viewed (or the pageviews) during the visit, the visit duration (in minutes), and the basket value (in US\$), indicating the amount of money visitors spent at Amazon during their visit sessions, are of interest. The first two variables are related to browsing behaviors and the third is relevant to spending behavior. These three variables are addressed at both the session and the visitor level. At the visitor level, an additional variable is also included, the visit frequency or the total number of visits to Amazon one visitor may have. We included it based on a suggestion from Zhu et al. (2019) and Luo, et al. (2021).

2.3 Hypothesis Development

From the previous research focusing on browsing and spending behaviors at retail websites, two streams of empirical work can be identified. The first stream concentrates on examining the correlation among online behaviors and the second stream attempts to classify the visit sessions (or the visitors) into meaningful distinctive clusters. Past research has confirmed correlations among online behaviors at retail websites (Mallapragada et al., 2016; Panagiotelis et al., 2014; Chatpong Tangmanee, 2019a; Xu et al., 2018). These relationships between pageviews, visit duration and basket value are confirmed to be positive. Logically, the longer the session, the higher the number of pageviews (Chatpong Tangmanee, 2017; Wu, 2021). Given Xu, et al. (2018)'s conceptualization of a pageview as a visitor's perception, they discovered its positive connection to a visitor's intention to make a purchase. Also, the positive relationships among pageviews, duration and basket value at the session level are evident in previous work (Aversa, Hernandez, & Doherty, 2021; Mallapragada et al., 2016). With an attempt to model the number of pageviews across websites, Gaspard et al. (2018) used the pageview data at the visit session level and at the aggregated level across websites. However, how he aggregated the data is not clear. Among a relatively small amount of research using the data in the visitor level, Luo, et al. (2021) confirmed that the number of retail website visitors through personal computers had higher impact on sales performance than those accessing the websites through mobiles.

Other research attempts have examined browsing and spending behaviors in conjunction with other variables where each behavior was treated as a dependent variable. Sychov and Bakaev (2020) proved that the number of pageviews at the session level was attributable to a few variables from three levels; it depended upon a visitor's demographic information (i.e., the finest level), the website design (i.e., the intermediate level) and the visit occasion (i.e., the broadest level). Among the others, Sychov and Bakaev (2020) suggested that the duration per pageview be used in addition to the number of pageviews or the visit duration since it might help to look at browsing behavior from a different perspective. Using stickiness as a proxy of the visit duration, Xu et al. (2018) confirmed the significance of visit duration in predicting a visitor's intention to make an online purchase from mainstream media websites.

Alternatively, online behavior was also treated as an independent variable. Using Apple and Amazon websites as the research context from which visitors' browsing behavior were recorded, Panagiotelis et al. (2014) confirmed that online sales depended on pageviews and visit duration. Yet, the degree of dependency relied on what the products were. Based on a control experiment, Buchanan et al. (2018) confirmed significant links between advertising recall, its recognition and the length of stay on retail websites. Given the nature of the controlled experiment, all conditions in their study were fictional, thereby affecting the study's validity.

The second stream of empirical research has focused on using visitors' browsing and spending behaviors as well as other variables to group visit sessions or visitors into distinctive categories. Such findings have helped online practitioners to be able to satisfy the different needs of various groups. [Moe \(2003\)](#) developed and tested her typology of online store visits using clickstream panel data. Based on visit behavior and purchase horizon, the sessions were classified into four groups: direct buying, hedonic browsing, knowledge building, and search or liberation. The first group, for example, could be an online business's primary target group since these customers make a purchase directly. [Later, Wu et al. \(2021\)](#) developed a visitor-level model using clickstream panel data from one car sales website. Their model is insightful since it captures a considerable portion of the variation at the session level and the individual level, and changes over time as visitors gain experience from the visit. Nonetheless, how they aggregated the data from the session level up to the visitor level was imprecise. [Pallant, et al. \(2017\)](#) replicated [Moe's \(2003\)](#) study; not only did they reconfirm the four clusters previously found by [Moe \(2003\)](#), but they also discovered a fifth group and labelled it as the cart-only group. Although members in the fifth group may not make a purchase during their visit sessions, they still use the cart to arrange their shopping items. Recently, [Zavali, et al. \(2021\)](#) used the clickstream data from one UK-based online apparel retailer to segment customers into six groups. Their finding of "visitors with a purpose" are similar to [Pallant, et al. \(2017\)](#)'s the "goal-direct." In addition, those visitors with the purpose generate the highest revenue for the website ([Zavali et al., 2021](#)).

Although the two research streams have considerably advanced knowledge on browsing and spending behaviors during visits to retail websites, there is still room for empirical research. The opportunity for further studies may be from the fact that the gathering of (or accessing to) information from one's actual visit to a retail website is a great endeavor. To maneuver around this expensive effort, researchers have had to reconceptualize visit behavior. A sizable amount of previous work has used stickiness as a proxy for browsing behavior ([Huang, Jia, & Song, 2015](#)). Stickiness could be measured using questionnaire items, the quality of which is adequate; however, it is not the actual visit behavior. Moreover, a visitor's recall of such granular behavior as visit duration may not be accurate. Similarly, a visitor's disclosure of his or her basket value may be invalid because people are uncomfortable with sharing their own financial information.

Among previous studies that have used clickstream panel data on visit behavior at retail websites including Amazon.com, some of the research may not have addressed browsing and spending behavior while others may not have made it clear whether the data used in their studies were from the session or visitor levels. [Moe \(2003\)](#) treated spending behavior as a dichotomous variable of whether visitors did or did not make a purchase during their visit sessions. Recently, [Li et al. \(2018\)](#) addressed both browsing and spending behaviors; nonetheless, their scope on real estate websites may slightly limit the findings' generalization.

Given that one person may have multiple visits to Amazon, little research has clearly addressed both the session and the visitor levels. Panagiotelis, et al. (2014) offer econometrical insight into online sales as a function of visit behavior, yet their work was based on data from the session level. According Di Fatta et al. (2018)'s findings in which they used in-store clickstream data to examine buying, browsing and searching behaviors at one apparel website, it is still unclear how they handled the aggregation of data in the session up to that at the visitor level. Even in the recent study (Zavali, et al., 2021), how the data were rolled up from the fine session level to the broad visitor level is still ambiguous.

It can thus be reasonable to claim the serious flaw in the past research that (1) has examined the browsing and the spending behaviors at Amazon, using the data which seem to have the quality problems, or (2) has ambiguous findings of whether they were based on the fine level of visit sessions or on the broad level of visitors. Therefore, based on the prior literature the following hypothesis have been developed:

H1a: The number of pages viewed in one visit session in the sessions with purchase are statistically different from the number of pages viewed in one visit session in the sessions with no purchase.

H1b: The length of one visit session in minutes in the sessions with purchase are statistically different from the length of one visit session in minutes in the sessions with no purchase.

H1c: The amount purchased in US\$ in one visit session in the sessions with purchase are statistically different from the amount purchased in US\$ in one visit session in the sessions with no purchase.

H2: The amount purchased in US\$ in one visit session is significantly explained by the number of pages viewed in one visit session and the length of one visit session in minutes.

H3a: The average number of pages one visitor viewed per visit among the visitors with purchase experience are significantly different than the average number of pages one visitor viewed per visit among the visitors with no purchase experience.

H3b: The average length of time spent each visit per visitor (in minutes) among the visitors with purchase experience are significantly different than the average length of time spent each visit per visitor (in minutes) among the visitors with no purchase experience.

H3c: The average length of stay each visitor spent on each page among the visitors with purchase experience are significantly different than the average length of stay each visitor spent on each page among the visitors with no purchase experience.

H3d: The total number of visits to Amazon.com one visitor had among the visitors with purchase experience are significantly different than the total number of visits to Amazon.com one visitor had among the visitors with no purchase experience.

H4: The average amount of purchase in US\$ each visitor had made per visit is significantly explained by the average number of pages one visitor viewed per visit, the average length of time spent each visit per visitor (in minutes), and the total number of visits to Amazon.com one visitor had.

Hence, our research objectives were (1) to report and compare browsing and spending behaviors at Amazon.com (a) between sessions with purchase and those without purchase (H1a - H1c), and (b) between the visitors who made a purchase and those who made no purchases (H3a – H3d); and (2) to explore the extent to which browsing behaviors may explain spending behavior (H2 and H4).

3. RESEARCH METHODOLOGY

3.1 Research Approach and Data Manipulation

Our research approach is quantitative and the two units of analysis are the visit-session level and the visitor level. The former is nested in the latter. Using the secondary source of data, we obtained the visit-session-level household panel detail from the ComScore web behavior project, which is subscribed to by Chulalongkorn University's Business School. This allowed us to download the actual visit behavior at Amazon for our research project.

We extracted the data from January 1 to December 31, 2020, resulting in a sufficiently large set of data with a total of 2,959,745 visit sessions at Amazon.com. We further removed the sessions with a visit duration of less than one minute or with one or less pageview because such a short visit may signify an accidental encounter. This yielded a dataset of 1,812,569 visit sessions for further analyses. 13% of these sessions included a purchase; the others were just visit sessions. As one visitor may have multiple visit sessions, we thus combined all sessions by one visitor to create a second dataset at the broader layer termed the "visitor level" in the current study. The dataset at this level includes a total of 79,696 records of the visitors, 40% of whom had no children and the remaining who had at least one child and 53% who attended college or had a bachelor's degree. The largest proportion (35%) had a household size of two members. Geographically, 19% of the visitors in our dataset lived in the Northeast, or in the Central region, 38% were in the South and the rest lived in the West. Given the data from the two layers, session and visitor, we included four variables from the fine layer and five from the broad layer. In other words, the data from the former are nested within the latter. Following [Buchanan, Kelly, Yeatman, and Kariippanon \(2018\)](#)'s and [Luo, et al \(2021\)](#)'s suggestions, we added to our examination the duration per page at both levels.

Given the data aggregation, we were able to include the visit frequency at the visitor level. Hence, the datasets were deemed ready for further analyses.

3.2 Data Analysis and Hypothesis Testing

From the study's objectives, we first reported descriptive statistics for all major variables listed in Table 1. Second, we used an independent t-test to compare the pageviews (PV_SESS), the visit duration (VD_SESS), and the duration per page (DP_SESS) between the sessions with purchase and those with no purchases. Similarly, we used the same test to compare the variables at the visitor level (i.e., PVV_VSTR, VDV_VSTR, DPV_VSTR, and NOV_VSTR) between the visitors who had made no purchases and those who had made at least one purchase at Amazon.com. Finally, we explored data at the session level to see whether the basket value significantly depends on the number of pageviews and the visit duration. Also, we performed the same analysis on the visitor data to explore whether a visitor's basket value per a visit significantly depends on the number of pageviews, the visit duration and the number of visits. To accomplish this, we used regression analysis.

Table 1: Major Variables, Their Meaning and Their Names

Variables	Meaning	Name
Session Level		
Pageview	The number of pages viewed in one visit session	PV_SESS
Visit duration	The length of one visit session in minutes	VD_SESS
Basket value	The amount purchased in US\$ in one visit session	BV_SESS
Duration per page	The average length of stay for each page in one session	DP_SESS
Visitor level		
Pageview per visit	The average number of pages one visitor viewed per visit	PVV_VSTR
Visit duration per visit	The average length of time spent each visit per visitor (in minutes)	VDV_VSTR
Basket value per visit	The average amount of purchase in US\$ each visitor had made per visit	BVV_VSTR
Duration per page per visit	The average length of stay each visitor spent on each page	DPV_VSTR
Number of visits	The total number of visits to Amazon.com one visitor had	NOV_VSTR

Table 2: Hypothesis Statements

Level of data	Scope of data	Hypothesis statements	Research objective to which the hypothesis responds
Visit session	All sessions	H1a: PV_SESS in the sessions with purchase are statistically different from PV_SESS in the sessions with no purchase.	Objective 1(a)
		H1b: VD_SESS in the sessions with purchase are statistically different from VD_SESS in the sessions with no purchase.	Objective 1(a)
		H1c: BV_SESS in the sessions with purchase are statistically different from BV_SESS in the sessions with no purchase.	Objective 1(a)
	Only the sessions with the purchase	H2: BV_SESS is significantly explained by PV_SESS and VD_SESS.	Objective 2
Visitor	All visitors	H3a: PVV_VSTR among the visitors with purchase experience are significantly different than PVV_VSTR among the visitors with no purchase experience.	Objective 1(b)
		H3b: VDV_VSTR among the visitors with purchase experience are significantly different than VDV_VSTR among the visitors with no purchase experience.	Objective 1(b)
		H3c: DPV_VSTR among the visitors with purchase experience are significantly different than DPV_VSTR among the visitors with no purchase experience.	Objective 1(b)
		H3d: NOV_VSTR among the visitors with purchase experience are significantly different than NOV_VSTR among the visitors with no purchase experience.	Objective 1(b)
	Only the visitors with the purchase	H4: BVV_VSTR is significantly explained by PVV_VSTR, VDV_VSTR, and NOV_VSTR.	Objective 2

Given the mean comparison and the dependency validation, all hypotheses are included in [Table 2](#). Based on the session-level data, we tested if the pageview, the visit duration, and the duration per pageview from the sessions with purchase and those without purchase were significantly different (i.e., H1a to H1c). Also, we used the data from the visitor level to test if the browsing behaviors per visit between the visitors who had purchase experience and the visitors who had no experience were significantly different (i.e., H3a to H3d). Given the exploratory nature, we wanted to assess, using the data from each level, whether the basket value could be explained by browsing behaviors (i.e., H2 and H4).

4. RESULTS

Using the ComScore services, the record of 79,696 visitors to Amazon.com yielded 1,812,569 visit sessions. Descriptive statistics of the browsing and spending behaviors are shown in [Table 3](#). Briefly, a visitor in one session viewed an average of 19.60 pages and stayed on Amazon for roughly 22.46 minutes. Approximately 1.62 minutes were spent on each page during a visit. Moreover, 13.1% of these 1,812,569 sessions included transactions, which had an average basket value of US\$ 72.86 dollars. An observation of the skewness and the kurtosis statistics of PV_SESS, VD_SESS, BV_SESS, and DP_SESS in [Table 3](#) confirmed these variables are not normally distributed for absolute values greater than 1 ([Chan, 2003](#); [Ringle, Sarstedt, Mitchell, & Gudergan, 2020](#)). Hence, all data must be transformed. A natural logarithmic function was used to transform all key variables in the session level, after which their distribution appeared normal and the subsequent parametric analyses were performed.

Table 3: Descriptive Statistics

Variables	Units	Mean	Standard deviation	Skewness	Kurtosis
Session level (n = 1,812,569)					
PV_SESS	Pages	19.60	34.889	28.87	3,354.93
VD_SESS	Minutes	22.46	47.913	15.58	360.73
BV_SESS	US dollars	72.86	737.492	457.96	218,284.73
DP_SESS	Minutes per page	1.62	2.009	30.780	6,358.88
Visitor level (n = 79,696)					
PVV_VSTR	Pages	9.11	10.051	6.28	162.99
VDV_VSTR	Minutes	9.62	11.294	9.90	297.12
BVV_VSTR	US dollars	74.82	754.82	151.68	23,541.19
DPV_VSTR	Minutes per page	1.15	0.830	6.72	143.68
NOV_VSTR	Times	37.14	90.036	12.88	613.91

From the broader view of the visitors, they spent approximately 9.11 minutes visiting 6.92 pages on average. A visitor spent an average of 1.15 minutes per pageview. Moreover, a visitor had approximately 37.14 visits to Amazon.com. 30.9% of the 79,696 visitors had at least one purchase transaction and the rest (69.1%) just looked around the store and left without making any purchases. The average amount a visitor spent per visit was US\$ 6.04. Given the absolute values of the skewness and the kurtosis statistics of PVV_VSTR, VDV_VSTR, BVV_VSTR, DPV_VSTR, and NOV_VSTR larger than one (see Table 3), these variables do not have normal distribution and data transformation was needed. All five variables were transformed using the natural logarithmic function, after which the distribution appeared normal and parametric techniques were used for further analyses.

The results of comparing the browsing and spending behaviors between the sessions with and without the purchase are shown in Table 4. The outcomes of the independent t-test analysis confirm that pageviews, visit duration, and duration per page between the two types of the sessions are significantly different. This means H1a to H1c are supported. Interestingly, the statistics in Table 4 show that sessions where a purchase occurred were greater in terms of length of visit (39.47 minutes) with more pageviews (46.01 pages) than sessions without purchase. Alternatively, the former group had a shorter stay per page (1.09 minutes per page viewed) than the latter (1.70). Further discussion will be included in the conclusion. The comparison using the data from the visitor level also yields similar findings. As can be seen in Table 5, visitors to Amazon who had made at least one purchase possessed distinctive browsing behavior (e.g., the pageview, the length of visit per pageview or the number of visits) compared to those who had made no purchase at all. In fact, visitors with previous purchasing experience viewed more pages, stayed longer and had more frequent visits than those with no experience. However, the former appears to stay for a shorter visit per page than the latter. Additional discussion will be included in the conclusion.

Table 4: Browsing Behaviors at The Session Level, Classified by Whether the Sessions Had, Or Had No, Purchase (236,909 Sessions Having the Purchase and 1,575,660 Sessions Having No Purchase)

Variables	Whether the sessions had, or had no, purchase	Mean	Standard deviation	Corresponding hypothesis (p-value)
PV_SESS	Yes	46.01	47.62	H1a (.000)
	No	15.62	30.64	
VD_SESS	Yes	39.47	37.22	H1b (.000)
	No	19.90	48.81	
DP_SESS	Yes	1.09	0.81	H1c (.000)
	No	1.70	2.12	

Table 5: Browsing Behaviors at The Visitor Level and The Visit Frequency, Classified by Whether the Visitors Had Made At Least One Purchase or Had Made No Purchase (24,667 Visitors Had Purchasing Experience And 55,029 Visitors Had No Experience)

Variables	Whether the visitors had made purchase	Mean	Standard deviation	Corresponding hypothesis (p-value)
PVV_VSTR	Yes	14.16	11.87	H3a (.000)
	No	6.85	8.16	
VDV_VSTR	Yes	14.35	11.92	H3b (.000)
	No	7.50	10.32	
DPV_VSTR	Yes	1.10	0.49	H3c (.019)
	No	1.18	0.93	
NOV_VSTR	Yes	85.28	135.42	H3d (.000)
	No	15.56	44.89	

Based on the mean comparison, the relationships between the purchases in Amazon and browsing behaviors, such as pageviews or visit duration, is confirmed. Our final objective is to explore the extent to which the purchase amount (or the basket value in the current study) is a function of online behavior in both the session and the visitor levels.

Using only the data from the sessions where a purchase was made, we explored the extent to which the number of pageviews (PV_SESS) and the visit duration (VD_SESS) could explain the variation of the basket value (BV_SESS). The analysis results are illustrated in [Tables 6](#) and [7](#); and four findings have emerged. First, the significant correlations among the PV_SESS, VD_SESS, and BV_SESS support the subsequent regression analysis. Second, the F statistics of 4,104.07 with p-value of .000 indicates that at least one of the independent variables contribute significantly to the basket value in a visit session at Amazon. Third, all statistics in [Table 6](#) and [Table 7](#) confirmed the significance of pageviews (PV_SESS) and visit duration (VD_SESS) in explaining the basket value (BV_SESS). Finally, the tolerance and the VIF statistics implies bearable concern regarding multicollinearity. However, the small value from the Durbin-Watson test (0.728) may point to a slight problem of residual autocorrelation. Furthermore, the tiny adjusted R^2 of 0.041 suggests the marginal variation of the basket value is accounted for by the two browsing behaviors. Given the exploratory approach, however, the regression analysis findings are still acceptable. Nevertheless, further use of these findings must be made with great caution.

Table 6: Correlation Matrix Based on The Session-Level Data

Variables	PV_SESS	VD_SESS
BV_SESS	0.191*** (.000)	0.187*** (.000)
PV_SESS		0.761*** (.000)

In parentheses are the p-value, *** indicates a significance level of 0.05

Table 7: Regression Analysis Results Based on The Session-Level Data Where BV_SESS is the Dependent Variable

Variables	Regression Coefficients (b)	Beta	t- statistics	p- value	Tolerance	VIF
Constant	2.779		249.08	.000		
PV_SESS	0.165	0.114	33.26	.000	.420	2.38
VD_SESS	0.128	0.101	29.32	.000	.420	2.38

Adjusted R^2 is 0.041 with a Durbin-Watson value of 0.728.

Based on the data from visitors who had made at least one purchase at Amazon.com, we explored the extent to which the basket value (BVV_VSTR) is attributable to pageviews (PVV_VSTR), visit duration (VDV_VSTR) and the number of site visits (NOV_VSTR) using regression analysis. Its output is shown in [Tables 8](#) and [9](#); and four findings were uncovered. First, the correlation matrix confirms the feasibility of subsequent analysis. Second, the F-statistic of 263.316 proved the plausibility of the regression modal. Third, the t-statistics and their p-value in [Table 8](#) and [Table 9](#) verified the significant effects of the number of pageviews (PVV_VSTR), the visit duration (VDV_VSTR) and the number of site visits (NOV_VSTR) on the basket value (BVV_VSTR) visitors spent on Amazon. Finally, the Durbin-Watson statistic of 2.030 shows no serious concern regarding autocorrelation in the model residuals. Based on the small amount of adjusted R^2 (0.032) and the possible issue of multicollinearity as suggested by the tolerance and the VIF statistics; however, the regression findings in the visitor level should be used with care.

Table 8: Correlation Matrix Based on The Visitor-Level Data

Variables	PVV_VSTR	VDV_VSTR	NOV_VSTR
BVV_VSTR	0.013*** (.037)	0.014*** (.024)	0.006 (.350)
PVV_VSTR		0.687*** (.000)	0.151*** (.000)
VDV_VSTR			0.162*** (.000)

In parentheses are the p-value, *** indicates a significance level of 0.05

Table 9: Regression Analysis Results Based on The Visitor-Level Data Where BVV_VSTR is the Dependent Variable

Variables	Regression Coefficients (b)	Beta	t-statistics	p-value	Tolerance	VIF
Constant	2.839		76.10	.000		
PVV_VSTR	0.116	0.067	6.37	.000	.372	2.691
VDV_VSTR	0.207	0.121	11.51	.000	.374	2.677
NOV_VSTR	0.034	0.040	6.14	.000	.991	1.009

Adjusted R² is 0.032 with a Durbin-Watson value of 2.030.

5. CONCLUSION AND DISCUSSION

In an attempt to examine the browsing and spending behaviors of customers at Amazon.com using the data from the visit-session and the visitor levels, we were able to study 1,812,569 sessions by 79,696 visitors during 2020. The demographics (e.g., the household size, or the resident location) accurately tap the profile of Amazon visitors (Similarweb, 2020). The visit sessions in the current study are different from those previously reported (Chatpong Tangmanee, 2017). Such difference may indicate a volatile change in behavior by Amazon customers.

13.0% of these sessions involved a purchase with an average basket value of US\$ 72.86 dollars. Each session consisted of approximately 19.60 pageviews and a visit duration of 22.46 minutes. At the broader level, one Amazon visitor visited 9.11 pages with an average session duration of 9.62 minutes. 30.9% of the 79,696 visitors in the current study made at least one purchase with an average basket value of US\$ 74.82 dollars per session.

Compared to C Tangmanee (2019b), the browsing and the spending behaviors at Amazon.com in the session and the visitor levels in the current study are of higher magnitude than previously reported. For instance, one visit session in Tangmanee (2019b) had approximately 9.67 pageviews with a visit duration of 11.38 minutes. But those in 2020 (the current study) were about twice as high as the figures in Tangmanee (2019b). Moreover, the basket value in Tangmanee (2019b) was US\$ 8.33 dollars but in 2020 it was US\$ 72.86 dollars, which is 8.75 times higher. We speculate that the relatively larger quantities in 2020 could be a result of the COVID-19 pandemic when many people around the world had to work from home and/or practice social distancing.

The comparisons between sessions with and without purchase using pageviews, visit duration, and duration per pageview as a basis confirmed the findings of previous studies (Rausch, Derra, & Wolf, 2022; Chatpong Tangmanee, 2017). Our findings have ascertained that the number of pageviews and the duration of the sessions where a purchase took place are both larger than those in the sessions where no purchase was made. Hence, an Amazon visitor's decision to purchase during the visit session may

influence them to view more pages and stay longer than sessions when they just browsed the store and a purchase was not made. Nonetheless, a look at the duration per page between the sessions with and without the purchase presents an intriguing insight. The visitors in the sessions with a purchase appeared to be quicker than those in sessions without a purchase for a shorter duration per pageview. This may signify the goal-direct purpose of the visit, which was found in the major cluster of the visit sessions by [Pallant, et al., \(2017\)](#) and [Zavali, et al. \(2021\)](#).

Similarly, a comparison of the pageviews, visit duration and duration per pageview between the visitors who had purchase experience with those who had not exhibits findings similar to those at the session level. That is, those who had made at least one purchase had a longer duration and higher pageviews than those who had made no purchase at Amazon. However, the former may have rushed to visit the store compared to the latter. This is because those with a purchase history had a shorter duration per pageview per visit than those with no purchase history. These findings at the visitor level are consistent with those at the session level. The discovery that sessions in which a purchase was made were shorter than those where a purchase was not made is perhaps our unique contribution. However, a recent study shows that the large group of visitors with the long length of stay may enjoy only browsing many pages but have little contribution to the store's revenue ([Zavali et al., 2021](#)).

In addition, visitors with a previous purchase history had more frequent visits to Amazon than those with no previous history. Taken together with browsing behaviors and purchase history, visitors with a purchase history tend to have a longer length of visit with larger pageviews and more frequent visits than those with no purchase history. Referring to the findings on goal-directed visitors in previous work that examined Amazon and other retail websites ([Creedy et al., 2017](#)), an Amazon visitor with a previous purchase history in the current study could be an example of this type of visitor.

The exploration into the extent to which browsing behavior could explain the basket value of the session and the visitor levels extend three issues. First, the basket value in the session was positively and significantly dependent on both the number of pageviews and the visit duration (see [Tables 6 and 7](#)). In other words, a longer session with more pageviews will likely increase the basket value. This is in line with previous studies ([Köster, Matt, & Hess, 2021](#); [Zavali et al., 2021](#)) Based on visitors' perception, the intention to purchase was positively related to visit duration (or stickiness in their own terms) ([Xu et al., 2018](#)) The channels to visit online retail stores could moderate the relationship between the visit duration and the pageviews. Should visitors get in the retail websites through the social advertisement, they tend to have higher pageviews and longer visit duration than those who had arrived the websites through social referrals ([Köster et al., 2021](#)).

In addition, [Wu et al. \(2021\)](#) suggest that pageviews at the session level are more indicative of a purchase than visit duration since the former is not affected by a visitor's browsing pace or personal habits. Given the small amount of adjusted R^2 (0.041) in the current study, the basket value is accounted for by other influential factors. The use of explorative analysis at the session level must be made with care.

Second, the basket value at the visitor level was significantly attributable to the pageviews, the visit duration, and the number of visits (see [Tables 8 and 9](#)). The regression results confirm that the higher pageviews, the longer visit duration and the larger number of visits lead to the higher basket value. The significance of all three explaining variables is feasible. Should a visitor view several pages, have long length of stay or have frequent visits, he or she is likely to make an expensive purchase. It may further imply that the visitors are searching for information in order to make purchase decision ([Zavali et al., 2021](#)). This finding is in line with [Mallapragada, et al. \(2016\)](#)'s work which discovered that the pageview and the visit duration were positively correlated to whether a purchase was made. Similar to the basket value, the sale performance was examined in [Luo, et al. \(2021\)](#) to see if it could be explained by the set of the visitor-level variables. One of their findings is that the sale performance was different between the two channels of the visits (i.e., personal computers vs. mobiles). The effect of the former was larger than the latter.

Given the limited research into the connection between the number of visits to retail websites and the number of transactions, we must rely on the work of [Luo, et al. \(2021\)](#). Using the time-series analysis, the number of prior visits was one of the significant predictors of the online sale performance in one Chinese retail website. Given the marginal yet significant amount of R^2 (0.032) in the current study, we encourage fellow researchers to examine the determinants of the online retail store's basket value using the data in the visitor level.

Finally, the juxtaposition of our regression analysis demonstrates the conceptual resemblance between the findings on the session and on the visitor levels. First, the browsing and spending behaviors at both levels between when the purchase was made and when it was not are similar. When at least one transaction was made, the pageview and the length of visit were higher than when it was not. However, the duration per pageview was smaller, implying a quick visit to the Amazon site, especially when the visit ended with the purchase. Such behaviors correspond to those reported in past research projects in which the visits that ended with a transaction were determined to be goal-directed. It is labelled as the direct purchase in [Moe \(2003\)](#), the goal-direct in [Pallant, et al. \(2017\)](#), and recently the visitors with a purpose in [Zavali, et al. \(2021\)](#). Moreover, our findings have validated the marginal yet significant explanatory effects of the pageviews and the visit duration on the basket value in the session and the visitor levels. In addition, the number of visits in the latter is also significant in explaining the basket value. This is consistent with [Wu, et al. \(2021\)](#), and [Luo, et al. \(2021\)](#).

6. RESEARCH IMPLICATIONS

Our findings offer both theoretical and practical contributions. The theoretical implication comes from insight into the analysis of browsing and spending behaviors from the session and visitor levels at Amazon.com. It is our goal to shed new light on online behavior at one of the world's leading retail websites. Our findings add to the body of empirical work through which two assertions can be made. First, an online purchase can be triggered by similar sets of browsing behavior during visit sessions or across visitors. Similarly, between the sessions and the visitors, one may have a longer duration of visit with higher pageviews when a purchase is made as compared to when it is not. Yet, the visit with a purchase appears to be relatively hurried as compared to that with no purchase. Second, the number of pageviews is prominent in explaining the substantial variation in the basket value. This is also common between the visit sessions or among the visitors. Creating a visit journey through a long series of pageviews at Amazon.com could result in a high basket value.

The practical contribution offers two recommendations for Amazon specially; however, other online retailers could benefit from them as well. First, online retailers should be attentive to browsing behavior in addition to basket value. Many websites offer neither products nor services, their major revenue instead being mainly from the traffic. A pornography website would be one example (C Tangmanee, 2019b). The recommended browsing behaviors including the pageviews or visit duration could help explain spending behavior. Second, online retailers may consider tracking a visitor's number of visits since it helps characterize visitors who may become direct users as, following the works of Pallant, et al. (2017) and Luo, Ngai, Li, and Tian (2021), visit frequency indicates who could potentially become direct buyers. Such information would be helpful in targeting the right market segment for online retail business.

7. LIMITATIONS, AND FUTURE RESEARCH DIRECTIONS

Similar to other research, our project does include limitations. Relying on browsing and the spending behaviors during visit sessions at Amazon.com, our conclusion is inevitably bound by this set of data. Although our study is valid for this set of data, we are unable to offer any discussion beyond our scope. We strongly expect that future work should be extended to cover other websites over a broader timeframe. Given the limitation and the contributions, we suggest two directions for future research. First, scholars may want to examine the other retail websites using the clickstream visit data so our finding's generalizability can be validated. Second, the investigation into what drives the basket value appears promising. We suggest this direction for the marginal contribution of the browsing behaviors in explaining the basket value in the current study. As a result, researchers may want to include the other predictors so the finding can be more valid than ours.

REFERENCES

- AL-Hawari, A. R. R. S., Balasa, A. P., & Slimi, Z. (2021). COVID-19 impact on online purchasing behaviour in Oman and the future of online groceries. *European Journal of Business and Management Research*, 6(4), 74-83. doi:<https://doi.org/10.24018/ejbmr.2021.6.4.923>
- Aversa, J., Hernandez, T., & Doherty, S. (2021). Incorporating big data within retail organizations: A case study approach. *Journal of Retailing and Consumer Services*, 60, 102447. doi:<https://doi.org/10.1016/j.jretconser.2021.102447>
- Buchanan, L., Kelly, B., Yeatman, H., & Kariippanon, K. (2018). The Effects of Digital Marketing of Unhealthy Commodities on Young People: A Systematic Review. *Nutrients*, 10(2), 148. doi:<https://doi.org/10.3390/nu10020148>
- Chan, Y. (2003). Biostatistics 101: data presentation. *Singapore medical journal*, 44(6), 280-285. Retrieved from https://medicine.nus.edu.sg/rsu/wp-content/uploads/sites/15/2020/02/biostat101_resources3.pdf
- Clement, J. (2020a). Total retail sales worldwide Retrieved from <https://www.statista.com/statistics/443522/global-retail-sales>
- Clement, J. (2020b). Global Amazon retail e-commerce sales 2017-2021. Retrieved from <https://www.statista.com/statistics/1103390/amazon-retail-ecommerce-sales-global/>
- Creedy, D. K., Sidebotham, M., Gamble, J., Pallant, J., & Fenwick, J. (2017). Prevalence of burnout, depression, anxiety and stress in Australian midwives: a cross-sectional survey. *BMC Pregnancy and Childbirth*, 17(1), 13. doi:<https://doi.org/10.1186/s12884-016-1212-5>
- Di Fatta, D., Patton, D., & Viglia, G. (2018). The determinants of conversion rates in SME e-commerce websites. *Journal of Retailing and Consumer Services*, 41, 161-168. doi:<https://doi.org/10.1016/j.jretconser.2017.12.008>
- Gaspard, H., Wigfield, A., Jiang, Y., Nagengast, B., Trautwein, U., & Marsh, H. W. (2018). Dimensional comparisons: How academic track students' achievements are related to their expectancy and value beliefs across multiple domains. *Contemporary Educational Psychology*, 52, 1-14. doi:<https://doi.org/10.1016/j.cedpsych.2017.10.003>
- Gustafson, J. L. (2017). *The end of error: Unum computing*: Chapman and Hall/CRC.
- Huang, L., Jia, L., & Song, J. (2015). Antecedents of user stickiness and loyalty and their effects on users' group-buying repurchase intention. *Conference proceedings Twenty First Americas Conference on Information Systems, Puerto Rico*, , 1-17. Retrieved from <https://www.researchgate.net/profile/Jiahe-Song/publication/280051952>
- Kosower, D. A., Maybee, B., & O'Connell, D. (2019). Amplitudes, observables, and classical scattering. *Journal of High Energy Physics*, 2019(2), 137. doi:[https://doi.org/10.1007/JHEP02\(2019\)137](https://doi.org/10.1007/JHEP02(2019)137)

- Köster, A., Matt, C., & Hess, T. (2021). Do All Roads Lead to Rome? Exploring the Relationship Between Social Referrals, Referral Propensity and Stickiness to Video-on-Demand Websites. *Business & Information Systems Engineering*, 63(4), 349-366. doi:<https://doi.org/10.1007/s12599-020-00660-1>
- Li, Q., Wang, Q., & Lin, Z. (2018). Effects of consumer visit to online community and product channel on local sales of large consumer goods: Evidence from real estate industry. *The Journal of Strategic Information Systems*, 27(2), 191-204. doi:<https://doi.org/10.1016/j.jsis.2017.11.001>
- Luo, P., Ngai, E. W. T., Li, Y., & Tian, X. (2021). Effects of visit behavior on online store sales performance: personal computer (PC) versus mobile channels. *Internet Research, ahead-of-print(ahead-of-print)*. doi:<https://doi.org/10.1108/INTR-11-2020-0655>
- Mallapragada, G., Chandukala, S. R., & Liu, Q. (2016). Exploring the Effects of “What” (Product) and “Where” (Website) Characteristics on Online Shopping Behavior. *Journal of Marketing*, 80(2), 21-38. doi:<https://doi.org/10.1509%2Fjm.15.0138>
- Moe, W. W. (2003). Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream. *Journal of Consumer Psychology*, 13(1), 29-39. doi:https://doi.org/10.1207/S15327663JCP13-1&2_03
- Nguyen, M. H., Armoogum, J., & Nguyen Thi, B. (2021). Factors Affecting the Growth of E-Shopping over the COVID-19 Era in Hanoi, Vietnam. *Sustainability*, 13(16), 9205. doi:<https://doi.org/10.3390/su13169205>
- Panagiotelis, A., Smith, M. S., & Danaher, P. J. (2014). From Amazon to Apple: Modeling Online Retail Sales, Purchase Incidence, and Visit Behavior. *Journal of Business & Economic Statistics*, 32(1), 14-29. doi:<https://doi.org/10.1080/07350015.2013.835729>
- Rausch, T. M., Derra, N. D., & Wolf, L. (2022). Predicting online shopping cart abandonment with machine learning approaches. *International Journal of Market Research*, 64(1), 89-112. doi:<https://doi.org/10.1177%2F1470785320972526>
- Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2020). Partial least squares structural equation modeling in HRM research. *The International Journal of Human Resource Management*, 31(12), 1617-1643. doi:<https://doi.org/10.1080/09585192.2017.1416655>
- Salinas, S. (2018). Amazon sales grew by \$15 billion year over year and its peers don't even come close to that level of growth. *grew-by-15-billion-year-over-year-blowing-away-peers.html* Retrieved from <https://www.cnbc.com/2018/07/27/amazon-sales>
- Similarweb. (2020). Amazon Analytics Retrieved from <https://www.similarweb.com/website/amazon.com#overview>
- Statista. (2017a). E-commerce worldwide - Statistics & Facts. Retrieved from <https://www.statista.com/topics/871/online-shopping>

- Statista. (2017b). Retail e-commerce sales worldwide from 2014 to 2021(in billion U.S. dollars). Retrieved from <https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales>
- Statista. (2019). Projected global revenue of Amazon from 2002 to 2022 (in billion U.S. dollars). Retrieved from <https://www.statista.com/statistics/745091/annual-net-revenue-of-amazon>
- Sychov, V., & Bakaev, M. (2020). *Stay some more and buy? Modeling the effects of visit time on online shopping purchases*. Paper presented at the International Workshop on Data Mining and Knowledge Engineering.
- Tangmanee, C. (2017). Comparisons of Website Visit Behavior between Purchase Outcomes and Product Categories. *International Journal of Research in Business and Social Science* (2147- 4478), 6(4), 1-10. doi:<https://doi.org/10.20525/ijrbs.v6i4.760>
- Tangmanee, C. (2019a). An Empirical Analysis of the Pageview and Visit Duration of Pornography Websites. *International Journal of Research in Business and Social Science* (2147- 4478), 8(3), 72-82. doi:<https://doi.org/10.20525/ijrbs.v8i3.250>
- Tangmanee, C. (2019b). Visit Behavior at Amazon.Com: An Analysis of Viewers' Demographics in the proceedings of the 2019 International Conference on Business Management Research
- Wu, J., Liu, L., & Cui, T. (2021). What drives consumer website stickiness intention? The role of website service quality and website involvement. *International Journal of Services Technology and Management*, 27(3), 189-208. doi:<https://doi.org/10.1504/ijstm.2021.115163>
- Xu, F., Qi, Y., & Li, X. (2018). What affects the user stickiness of the mainstream media websites in China? *Electronic Commerce Research and Applications*, 29, 124-132. doi:<https://doi.org/10.1016/j.elerap.2018.05.001>
- Zavali, M., Lacka, E., & Smedt, J. d. (2021). Shopping Hard or Hardly Shopping: Revealing Consumer Segments Using Clickstream Data. *IEEE Transactions on Engineering Management*, 1-12. doi:<https://doi.org/10.1109/TEM.2021.3070069>
- Zhu, G., Wu, Z., Wang, Y., Cao, S., & Cao, J. (2019). Online purchase decisions for tourism e-commerce. *Electronic Commerce Research and Applications*, 38, 100887. doi:<https://doi.org/10.1016/j.elerap.2019.100887>