Associate Professor Dan-Andrei SITAR-TĂUT, PhD
Email: dan.sitar@econ.ubbcluj.ro
Department of Business Information Systems
Babeş-Bolyai University, Cluj-Napoca
Assistant Professor Daniel MICAN, PhD (Corresponding author)
Email: daniel.mican@econ.ubbcluj.ro
Department of Business Information Systems
Babeş-Bolyai University, Cluj-Napoca
Associate Professor Codruţa MARE, PhD
Email: codruta.mare@econ.ubbcluj.ro
Department of Statistics-Forecasts-Mathematics
Babeş-Bolyai University, Cluj-Napoca

CUSTOMER BEHAVIOR IN THE PRIOR PURCHASE STAGE – INFORMATION SEARCH VERSUS RECOMMENDER SYSTEMS

Abstract. The online shopping chain consists of a transaction, starting from the intention to buy and up to the purchase decision, regardless of the adopted strategy. For the buyer, the transaction cost is not only made up by the price of the commodity, but also by the cost of all other resources involved in the entire process. If simple information search (IS) is used, then the additional cost is given by the ineffective time amount spent. If recommender systems (RS) are used, then the privacy or data disclosure risk may be the price. If such a risk is accepted by the online community and managed diligently by third parties, this cost may turn into a real benefit. Time is very important for us, that's why IS habits may not be effective. RS facilitate the content to look for us, without explicit searches. This paper outlines the behavior on the purchase-pull approach – here IS – and the purchase-push adoption – here RS – in a 140-user sample through Structural Equation Modeling. It assesses the impact of personal and contextual features upon IS, respectively its relationships with the attitude towards the perceived usefulness, acceptance, and intention of different RS use.
Keywords: Recommender Systems, Structural Equation Modeling, Personalization, Information Searching, Perceived Usefulness.

JEL Classification: L81, M31

1. Introduction

Searching for information on the web (IS) and evaluate buying alternatives regarding the commodities is a common habit. Even if information access is cheap

and fast, the impartial character of the information is debatable. Thus, site owners can manipulate people, and web content may have a dualist nature (Roscoe et al. 2016). To face these negative effects, organizations have started to invest in developing recommender systems (RS) that help users discover items that might represent interest for them. Content customization is a widespread practice and its purpose is to optimize the time spent before completing the purchase. RS are used by companies wishing to cut processing costs, helping customers reduce their range of possible options (Roscoe et al. 2016). Therefore, sets of RS (Pathak et al. 2010) has been developed, derived from different philosophies. No matter the philosophy behind, the RS make the content to look for the user in real-time, and maybe in the near future, the products will be ordered and delivered right in our homes and refrigerators. But to do this, any company must first listen to and aggregate user's preferences and then provide products specially created to meet the needs and these preferences.

IS and RS are very complex and interrelated domains. The goal of our study is to capture feedback on the interfaces of the RS, respectively on the data these systems collect from users. Therefore, in our opinion, a study regarding the utility, perceived efficiency, and acceptance by users of the main types of RS existing on the market, is required. The openness regarding data input contributing to personal profile building to support better recommendations is also evaluated. These main and complex RS-related aspects will be confronted with the prior-purchase information search behavior to discover valuable relationships and insights.

Our research is focused on IS and its related factors. Personal and contextual features of the online buyers, such as awareness, use, utility, and acceptance of the RS versus IS are assessed and several working hypotheses are tested. Results were obtained on a valid sample of 140 respondents.

The remainder of the paper is structured as follows: Section 2 indicates related work on information search and RS, underlying their importance in online shopping. Section 3 presents the problem statement, working hypotheses, and research method. Section 4 descriptively presents the sample, provides the results of the Structural Equation Modeling (SEM) analysis, assesses, and interprets the working hypotheses. The final section highlights the paper contributions and offers an outlook on future research perspectives.

2. Theoretical foundation and literature review
2.1. Information searching

In most of the cases, the purchase decision is not a random or an unintentional process. People look first for the related information over the Internet or inform themselves from additional sources before a commodity is finally purchased. We believe that IS is performed differently in online shopping according to the main personal characteristics, like gender (Sramova and Pavelka 2019), age (Wan, Nakayama, and Sutcliffe 2012; Rowlands et al. 2008), recurrent

online buying intensity (Hjorthol 2009), and knowledge level regarding alternative awareness and assessments in the online environment, like RS (Roscoe et al. 2016; Wan, Nakayama, and Sutcliffe 2012; Winoto, Pinata, Tang 2014), or various ads present on websites (Sigal 2012; Rezola, Gutierrez, and Linaza 2016).

No matter the commerce type, demographics like gender and age are important attributes in marketing segmentation (Velaudham and Baskar 2015; Blanca 2011), alongside with sociocultural orientation or other features, which are out of this paper's topic. In terms of IS for online buying purposes in respect to gender, the literature is not very clear in respect to the existence of a certain behavior for men or women. For example, (Szymkowiak and Garczarek-Bąk 2018) consider both genders having similar habits, while (Velaudham and Baskar 2015) state that women spend more time searching for information, but men look more for alternatives. In respect to age, the "Google Generation" (Blanca 2011) is, on one hand, inclined towards IS and, on the other, they are direct descendants of a population already used with IS habits from the brick-and-mortar commerce (Kluge, Dina 2012).

However, information search is a very expensive resource waster. This is the reason why specialists look for solutions to reduce the time spent on online shopping but to make user satisfaction higher. Here, the RS and the ads (Rezola, Gutierrez, and Linaza 2016) put on all types of websites and social networks (SN) may be a solution. The results are remarkable, the RS use considerably are substituting IS (Roscoe et al. 2016; Wan, Nakayama, and Sutcliffe 2012; Murali, Vishnu, and Victor 2019).

2.2. Recommender systems

The RS are extensively used on e-commerce sites to make the customer experience more pleasant and effective and to reduce the time spent. Recommendations impact positively on products and services' sales (Pathak et al. 2010) the IS cost and time are significantly reduced. RS are changing the e-commerce universe and suggest the most important and useful products out of a predefined list of products (Bobadilla et al. 2013). Only a few papers have researched issues related to users' subjective opinions and try to develop additional criteria for evaluating RS. Therefore in (Pu, Chen, and Hu 2011) a unified evaluation framework, called ResQue, has been proposed, which aims to measure the quality of RS taking into account the suggested elements, the interface and the quality of the interaction, the users' satisfaction and their influence on the behavioral intentions of the users.

In line with was previously stated, we consider as necessary to take into account the users' opinion to evaluate not only a system but the most popular RS that they met and became aware of within the sites, respectively the applications they use each day. The usefulness and effectiveness of an RS from an e-commerce site should increase the website's usability (Bobadilla et al. 2013). RS perceived usefulness facilitates the buying impulse, like the unplanned purchases, or trips

Dan-Andrei Sitar-Tăut, Daniel Mican, Codruța Mare

_____

promoted by RS (Lu et al. 2015). To improve the shopping experience, an RS may consider many features and inputs from users, but in compliance with privacy, security, and trust aspects (Bobadilla et al. 2013). Perceived ease of use and perceived usefulness noticeably affect the impulse to use such systems (Pu, Chen, and Hu 2011). Literature mentions a series of RS types (Pathak et al. 2010; Lu et al. 2015) etc., as the online shoppers perceive them. Also, the industry – Amazon, Netflix, eBay, Booking, YouTube, Facebook, LinkedIn, etc. – applies very complex RS relying on many of the features exposed in Table 1.

Table 1. Indicator variables for construct UUARS – user opinion about the perceived usefulness/acceptance of different RS types

Variable	Item / RS type
CPVS	Similarity with the products viewed (products/movies)
CPH	Inspired from your browsing history (time spent on product/movie page)
CPSL	Similarity to the products you liked (you liked)
CPSB	Similarity to the products you bought
CPS	Similar products (based on similar characteristics)
CPSO	Similar products (based on user preferences with similar tastes)
CPC	Accessories or complementary products
CPVL	Products seen/liked frequently by users
CPCS	The products bought together by the customers
CPIF	Products viewed, liked or purchased together by groups I follow on social media
CPV	Other visitors have an interest in them and have saved them to favorites
CPBN	What users are buying now
CPBT	Top Products/Brands
CPP	Popular/best-selling products
CPT	Trend products
CPO	Products on offer
CPOCF	Products on offer taking into account the product category

Generic statistics/metrics used in online shopping and user demographic information on a side, and the social media behavior and awareness of a recommender system on the other side, are both aspects that cannot be ignored. The users are mainly open to rely on and also to share some relevant aspects, to have, and to offer for other buyers a better shopping experience. The online shopping community agrees to explicitly or implicitly share with the e-commerce sites certain types of personal data or information (Bobadilla et al. 2013) such as those listed in Table 2.

_____

Table 2. Indicator variables for construct CRS – user agree/believe that RS should collect data about

Variable	Item / User input data
UL	What user liked
UDL	What user disliked
UPR	Product ratings (star ratings)
URC	Reviews and comments from customers who have already used the product
UPI	Different personal preferences and interests selected from a list
UPIN	General personal information (gender, age, location)
UPSIN	Sensitive personal information (name, job, address)
UPV	Products viewed
UPTV	How much time the user spent and how many times the user viewed a product
UPP	Purchased products
UOF	Products bought/owned by friends from the user's SN
UAF	The products appreciated/desired by friends from the user's SN
UACF	Products appreciated/desired/owned by close friends/with whom the user constantly interact from his/her SN
UACFR	The products appreciated/desired/owned by close friends/with whom the user interact from time to time from his/her SN
UFOL	The products appreciated/wanted/owned by the people followed in his/her SN
UPAI	As many elements possible presented above
UPNI	No information from users.

3. Research Method
3.1. Questionnaire design and data collection

All information presented above regarding information search and RS are met and depicted widely in the online shopping field relevant literature. We consider they worth to be also empirically examined through the SEM modeling multivariate approach. To meet this goal, a survey was designed. The variables involved in the analysis are presented in Table 3.

Starting from the variables we study, the survey items were prepared in a clear, clean, and descriptive way, where it was necessary. The questionnaire addressed to an online shopping user sample contains 8 items, 2 of them being broken down in 17 sub-items each. The first one is IS before buying and the second indicates the user's RS awareness level. The following two are compound items and represent the core questions of the questionnaire (see Table 1 and Table 2 from the preceding subsections). Together with two demographic inquiries (age and gender), the other two information types – usually met in online shopping surveys – were included. They are the buying frequency and ads' utility on websites.

Dan-Andrei Sitar-Tăut, Daniel Mican, Codruţa Mare

Table 3. Variables used in the analysis

Variable	Description
IS	The respondent searches for information, advice, and recommendations on the Internet before buying a product
URS	Awareness utility level of the recommender system
UUARS	An index created by the authors that comprises information related to the perceived usefulness, acceptance, and intention to use different types of recommender systems presented in Table 1. Indicator variables for construct UUARS - user's opinion about the perceived usefulness/acceptance of different RS types.
CRS	An index created by the authors for the aspects that should be considered by the RS to provide the most suitable information for the client. It is made up of the aspects presented in Table 2. Indicator variables for construct CRS – user agrees/believes that RS should collect data about.
AGE	Age of the respondent
GEN	Gender of the respondent
OSF	How often a respondent buys products or services online.
UA	Level of the utility of the ads on the site

Except for the age and gender, all our model determinants rely on a five-point Likert scale. Here, "totally inadequate/effective", "strongly disagree", "totally unimportant" are item answers valued with 1, while "totally adequate/ effective", "strongly agree", "totally important" are valued with 5. The data was gathered in late 2019 through an online survey based on the Google Forms platform. It was filled in anonymously and voluntarily by young online buyers. After running all the quality checking procedures (e.g. removing missing data and straight-line answers) a valid sample of 140 respondents remained as grounding for further analyses.

3.2. Hypotheses formulation

In this paper, we study the extent to which the acceptance and use of RS influence the information searches that underlie the purchase decision. We also investigate whether certain demographic characteristics are correlated/influence the search for information and the frequency of online shopping. The variables used in our research are presented in Table 3.

Many persons search for information before buying a product or service. Starting from this idea, we have developed several research hypotheses that evaluate the impact of personal and contextual features upon information search, on one hand. On the other, we hypothesized features of the relationships between searching for information and the attitude towards the intention to use, the awareness, the utility, usefulness, acceptance, and agreed input types in RS. The following hypotheses were constructed and tested:

H1: Information searching is NOT dependent on age (AGE −/→ IS).

We assume that younger individuals are very homogeneous in terms of Internet use for online shopping. Consequently, in such a group, there should be no differences between people of different ages. Search for information is a natural

characteristic for the older generations, inherited from the brick and mortar traditional market environment (Wan, Nakayama, and Sutcliffe 2012). However, the new generation uses instinctively and extensively the search engines (Rowlands et al. 2008) more than elderly people starting from education, health problems, and least but not last to online shopping.

H2: There are significant differences between men and women in terms of information searching (GEN → IS).

In terms of online search behavior, we assume that there are significant differences between the two genders, each group having a different mindset in shopping habits, too. Women are more impulsive, instinctive, and inflexible in the purchasing behavior than men since the latter category is more flexible, being predisposed to look for alternatives (Velaudham and Baskar 2015). Looking for choices involves, by default, more searching.

H3: People that shop online more frequently are more likely to search for information (OSF → IS).

To search for information on the Internet is a normal habit. It is very difficult to quantify how much and how often are we searching exactly for a certain purchase goal (Hjorthol 2009) from all the searches that precede the actual buy. We constructed this assumption based on the idea that people that buy online more frequently are probably searching more for information.

H4: Age positively influences online shopping frequency (AGE → OSF).

It is well-known that income and age are positively correlated. That is why we assume that older people in the sample are more likely to shop online, as they have the required financial resources for this. However, a higher income does not guarantee to spend more money exclusively for online purchases. We have launched this hypothesis since the literature does not sustain unanimously the existence of a direct or reverse relationship (Kluge, Dina 2012).

H5: The RS awareness does not significantly affect the information search habit (CRS ─/→ IS).

IS and RS are widely used in online shopping (Wan, Nakayama, and Sutcliffe 2012) and other various fields (e.g. in academics, to search domain-relevant scientific sources (Murali, Vishnu, and Victor 2019)). In the last years, RS overtook the information search because time and effort are significantly saved (Roscoe et al. 2016; Wan, Nakayama, and Sutcliffe 2012; Murali, Vishnu, and Victor 2019). If the users are aware of RS they should limit straight IS and use recommendations where it is possible. This would be a rational operation way. However, we sustain the contrary due to the well-rooted habit in search engines use for the young generation (Rowlands et al. 2008) and possible minor inherited influence in traditional information search (Wan, Nakayama, and Sutcliffe 2012).

H6: The awareness level of RS utility is homogenously perceived in both genders (GEN ─/→ URS).

Gender is an important factor considered for marketing segmentation purposes (Blanca 2011) both in traditional and online shopping. RS must take into

65

account gender preferences in their mechanism (Winoto, Pinata, Tang 2014) for common goods and services, but especially for the sensitive ones (Winoto, Pinata, Tang 2014) from the ethical perspectives. Even if the RS outputs are heterogeneous according to gender, their perceived utility level is homogenously treated in both genders (Gordillo, A., Barra, E., Quemada 2017).

H7: The online shopping frequency is NOT gender reliant on (GEN −/→ OSF).

Young people tagged as the "Google Generation" (Blanca 2011) or "Generation Z" (Sramova and Pavelka 2019) shop online more than the elders. Globally, in this age cluster, there is not a significant difference in habits between men and women (Sramova and Pavelka 2019), even if in respect to information search or product/service type they may differ.

H8: The level of ads utility awareness influences information search behavior (UA → IS).

The ads present on a website designed for online shopping purposes should influence the information search behavior. The ads reduce the time spent in prior-purchase stages if the ads are well-tailored according to the user's profile (Rezola, Gutierrez, and Linaza 2016), being more intense for women (Sigal 2012).

3.3. Research methodology

The variables were, first, descriptively analyzed. Frequencies and percentages were computed for nominal and ordinal data. For quantitative data, the main descriptive statistics were computed (mean, median, standard deviation, etc.). To evaluate the proposed assertions, we used the Structural Equation Modeling (SEM) methodology. This is suitable for multiple interdependencies among the variables, in comparison with the traditional tools such as regression modeling, that are relied on a one-way approach. We believe this method of analysis is appropriate in our setup, due to its rapid increase in the use in many subjects, but still almost a newness in RS (Chen, Chia-Chen, Tsai 2019). Moreover, it is also suitable for small samples. Additionally, the normality assessment is not required for each variable (Hair et al. 2017).

Our model was obtained in the composite form, mode B. In the present study we employ a model with two composite indicators. These are linear combinations of the variables considered for each group. Assuming that we have the two groups of variables, each composite indicator can be written as a weighted linear variables combination in the group (Hair et al. 2017):

$$CI = \sum_{i=1}^{K} w_i \times X_i \,,$$

Where CI = the composite indicator,
w_i = the weight of each variable in the group
X_i = variable in the group
K = the total number of variables in the group.

_____

The SEM (Jöreskog 1973) is made up of two main parts, out of which the first one is always the structural part, which is described by the following equation:
$$\eta = B\eta + \Gamma\xi + \zeta \,,$$
Where η is the vector of endogenous variables (usually latent),

B is the regression coefficients matrix that relates the endogenous variables among them,

ξ is the vector of the predictors (exogenous variables),

Γ is the regression coefficients matrix relating the endogenous variables to the exogenous ones,

ζ is the error vector.

The B and Γ comprise the path coefficients of the model, which represent the main interest in the analysis, as they describe the interrelationships that appear in the final model. The weights are necessary for the estimation of the model in the sense that they give information about the relative importance of each variable in a certain construct.

The final model scheme is presented in Figure 1 and was validated using standard procedures. First, the reliability and consistency of the constructs were evaluated using the classical Cronbach's alpha (α) and two additional reliability measures: Dijkstra-Henseler's ρA and Joreskog's ρC. The classical threshold of 0.7 was considered for validity. The goodness of fit (Henseler, Hubona, and Ray 2016) for both saturated an estimated model was evaluated based on three parameters: Standardized Root Mean Residual - SRMR, Unweighted Least Squares Discrepancy - dULS and Geodesic Discrepancy - dG. A valid model (well fitted) has all these three values lower than both the 95% and the 99% confidence limits (Dijkstra and Henseler 2015). Post-estimation validation and assessment of the significance of the relationships were conducted, including bootstrapping (for significance analysis inference) (Chin 2010). Among them, the discriminant validity of the model was assessed both based on the Fornell-Larcker criterion and the Heterotrait-monotrait ratio of correlations (HTMT). Values lower than 0.9 indicate discriminant validity between constructs. Analyses were run in Power BI (data visualization), SPSS 24 (data visualization and analysis) and Adanco 2.1 (structural equation modeling).

4. Results and discussions
4.1. Descriptive data analysis
The sample is made up of 140 respondents, most of them being females (60.7% - 85 subjects). Female respondents are, on average, significantly younger (21.47 versus 24.69 for men, p-value < 0.001). The same type of relationship is also to be found for the median age (20 for women and 22 for men). And, even though the range is higher for women, the standard deviation shows a higher variation for men in terms of age.

Men and women have the same type of behavior in terms of online buying frequency. Consequently, H7 is accepted. Most of them use the online environment

for purchasing processes 5-6 times a year (women 31.43%, men 19.29%), followed by the group that does it once a month (13.57% in both genders). Very few persons use the Internet weekly for purchasing reasons (women 2.86%, men 3.57%).

In terms of devices used, most of the sample uses a desktop or laptop for buying online. Only 11% of the men use the mobile or the tablet for this, while almost 32% of women do it.

4.2. Data modeling with SEM

The central aspect of the analysis is if people search for information, suggestions, and recommendations on the Internet before buying a product. A significant majority of approximately 94% of the sample responded that they do. This proves that people want to be informed and want to know exactly what they buy when purchasing online. Taking this as a central action in the online purchasing process, we have developed a structural equation model to test the working hypotheses.

The two constructs in the final model were obtained in the composite form, mode B, which ensures the consistency of the weights (Dijkstra and Henseler 2015). As in the case of the Principal Component Analysis, for example, the weights are a standardized form to show the impact of each variable on the total variance. Consequently, variables that form the construct can be ranked in terms of importance based on the values of the weights (see Figure 1). According to these weights, we can see that in UUARS, the most important elements are: CPOCF, CPO, CPVL, CPSB, and CPSL. While in the CRS construction the main positions are occupied by the following variables: UPV, UPI, UL, UAF, and UDL.

The meaning of the variables mentioned here is detailed in Table 1 and Table 2. As stated in the methodological part, the most suitable/efficient/valid model was chosen based on quality parameters related to SEM and bootstrap procedures that were run. The final model is well fitted in both forms, saturated and estimated. For each of them, all three goodness-of-fit statistics are below 95% and 99% boundaries (see Table 4).

Table 4. Goodness-of-fit for the estimated SEM model.

	Saturated model			Estimated model		
	Value	HI95%	HI99%	Value	HI95%	HI99%
SRMR	0.0671	0.0718	0.0940	0.0681	0.0723	0.0907
d_{ULS}	3.6873	4.2302	7.2449	3.8007	4.2885	6.7452
d_G	1.1086	1.2915	1.5364	1.1147	1.3178	1.5505

To also evaluate the issues of utility, use, and acceptance of the RS, two major constructs have been built in the analysis (see Table 3 and related tables, Table 1 and Table 2). Both constructs passed the reliability validation. For the construct related to the aspects that should be considered by the RS to provide the most suitable information for the client – CRS – the classical Cronbach's alpha had

a value of 0.875, while Joreskog's rho one of 0.82. The other construct, related to the utility and acceptance of different types of RS – UUARS – had the values of 0.9 and 0.85 respectively. Both constructs had a value of 1 for the third reliability indicator, Dijkstra-Henseler's rho. Discriminant validity appears in the final model, as all HTMT values are much lower than the boundary of 0.9 (Table 5).
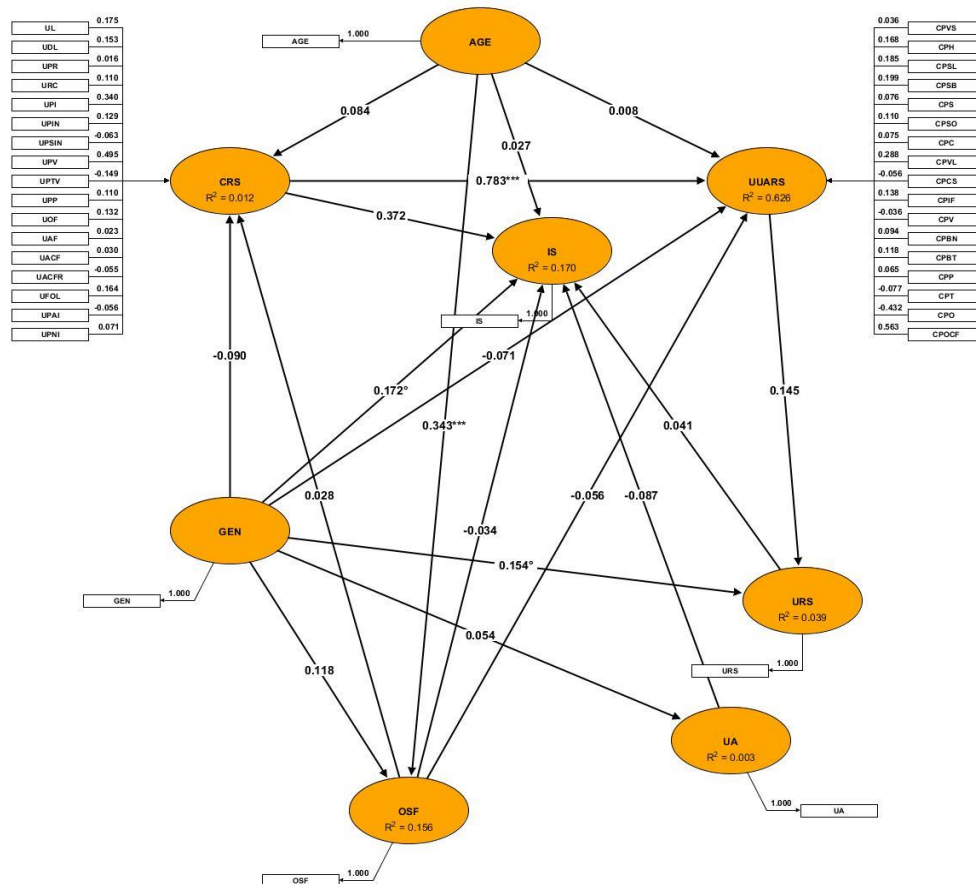

Figure 1. The model of the SEM analysis

Table 5. Discriminant validity assessment – HTMT results for the final model.

Construct	IS	URS	UA	GEN	AGE
IS					
URS	0.2306				
UA	0.2063	0.2851			
GEN	0.2367	0.2776	0.1963		
AGE	0.1399	0.1717	0.1982	0.4528	
OSF	0.2309	0.2278	0.2586	0.3507	0.4949

69

The effect overview and the results of the bootstrapping procedure are presented in Table 6. Effect size measures the strength of the relationship between two numeric-scale variables. Table 6 emphasizes the strength of the effect based on Cohen's f^2 coefficient. The higher the value, the stronger the relationship. We may note there is a strong effect ($f^2 \geq 0.35$ strong effect) between our constructs, a moderate one ($0.15 \leq f^2 < 0.35$) between CRS and IS, UUARS and URS, GEN and IS, GEN and URS, AGE and OSF, while an inexistent or unsubstantial effect ($f^2 < 0.02$) between the rest of the variables.

Table 6. Effect overview, total effects inference, and their significance – bootstrapping results for the final model

Effect	Beta	Indirect effects	Total effect	p-value	Cohen's f^2	Mean value	Std, err,	t-value
CRS → UUARS	0.783		0.783	0.000	16.204	0.647	0.146	5.354
CRS → IS (**H5**)	0.372	0.005	0.377	0.144	0.1642	0.329	0.258	1.461
CRS → URS		0.114	0.114	0.420		0.086	0.141	0.807
UUARS → IS		0.006	0.006	0.794		0.002	0.023	0.261
UUARS → URS	0.146		0.146	0.500	0.0217	0.121	0.216	0.675
URS → IS	0.042		0.042	0.630	0.0020	0.030	0.086	0.482
UA → IS (**H8**)	-0.087		-0.087	0.272	0.0088	-0.078	0.079	-1.099
GEN → CRS	-0.090	0.003	-0.087	0.690	0.0074	-0.070	0.217	-0.399
GEN → UUARS	-0.071	-0.075	-0.146	0.513	0.0121	-0.109	0.223	-0.655
GEN → IS (**H2**)	0.172	-0.035	0.136	0.043	0.0312	0.137	0.067	2.031
GEN → URS (**H6**)	0.154	-0.021	0.133	0.133	0.0243	0.143	0.088	1.504
GEN → UA	0.054		0.054	0.532	0.0029	0.053	0.086	0.626
GEN → OSF (**H7**)	0.118		0.118	0.151	0.0150	0.114	0.082	1.438
AGE → CRS	0.084	0.010	0.094	0.686	0.0058	0.073	0.232	0.404
AGE → UUARS	0.009	0.054	0.063	0.832	0.0002	0.052	0.296	0.212
AGE → IS (**H1**)	0.027	0.024	0.050	0.308	0.0007	0.044	0.049	1.021
AGE → URS		0.009	0.009	0.887		-0.018	0.064	0.142
AGE → OSF (**H4**)	0.343		0.343	0.000	0.1274	0.346	0.076	4.545
OSF → CRS	0.028		0.028	0.878	0.0007	0.039	0.181	0.154
OSF → UUARS	-0.056	0.022	-0.034	0.860	0.0071	0.021	0.195	-0.176
OSF → IS (**H3**)	-0.034	0.010	-0.024	0.858	0.0011	-0.024	0.132	-0.178
OSF → URS		-0.005	-0.005	0.914		0.004	0.046	-0.108

H1. The first working hypothesis is related to age versus IS. The literature says that the older generation (Wan, Nakayama, and Sutcliffe 2012), but also the "Google Generation" (Rowlands et al. 2008) search more for info. Results obtained in this analysis accept H1 (AGE $-/\rightarrow$ IS, with p-value > 0.05 for AGE $\rightarrow$ IS), as can be seen in Table 6), which means there is no significant difference in information search behavior in our sample. This is justified by the fact that, as stated in the hypothesis, young people have a homogeneous behavior in terms of Internet-related issues.

H2. Gender, similar to age, is an important variable in marketing segmentation (Blanca 2011) and it is proved that men and women have different behavior in various stages of online shopping (Velaudham and Baskar 2015). Analysis has validated that men and women have different behavior in terms of IS (GEN $\rightarrow$ IS, p-value < 0.05, H2 confirmed). The total effect of gender upon IS is positive, which indicates that men have a higher tendency to look for information before buying a certain product or service online. We may conclude that women are more impulsive in terms of online shopping. Since men are more analytic, their buying decisions require a deeper prior-purchase IS process (Kumaravel 2017). Some studies do not point out any behavioral difference between men and women in this respect (Szymkowiak and Garczarek-Bąk 2018), while others concluded the opposite – that women are searching more for information in the online buying process (Velaudham and Baskar 2015). But the latter study sustains that in online shopping, usually, men are looking more for alternatives than women.

H3. Prior-purchase IS is natural because the users must have all the details related to the object of the purchase. However, it is very difficult to distinguish exactly the amount of time strictly designed for online shopping and also to state that these searches will be converted actually into a buy (Hjorthol 2009). No relationship was found between the frequency of online shopping and IS, which means H3 (OSF $\rightarrow$ IS) is rejected. Both the direct and total effects of the frequency are extremely low, almost 0. Consequently, the probability resulted is higher than the critical limits even for the 10% one. However, both the total and direct effects are negative. So, in the event of a significant result, they would have been indicative of a reverse relationship between the two variables – people that buy more frequently, do not search for information anymore, or not to such an extent as those with a lower buying frequency. This can be explained that once with frequent online purchases the buyer acquires experience in these tasks and finally the search for information becomes unnecessary.

H4. The literature does not reveal very clear if the age is influencing the online shopping behavior or frequency. This hypothesis can be treated more likely as inconsistent, some saying the elder people buy online more than the younger ones (Kluge, Dina 2012). Every hypothesis regarding this is relative and depends dramatically on the methodological issues and the structure of the sample, considering overall statistics or age groups breakdowns, region influence, or product/service category. Our first working hypothesis has proven that young

71

people have similar online IS behavior. However, the current hypothesis (H4, AGE → OSF) is accepted (p-value<0.01) and demonstrates that age significantly influences their shopping frequency. The relationship between the two variables is positive, meaning that older people (coefficient=0.343) from our sample buy more. This can be explained by the fact that very young people from our sample are usually students, with very low incomes and dependent on others (like parents) for financial issues. On the contrary, older people in the sample are already employed, which means, first of all, that they are financially independent of their family.

H5. To say that IS is influenced by CRS cannot be sustained in our study. Consequently, our hypothesis – CRS −/→ IS – is accepted.

Even if expensive, IS is still an old-fashion habit, since RS act like search mediators and optimizers. If the users are conscious of the RS existence and their types, then a rational user behavior would be to limit the information search and to make the shift to RS (Murali, Vishnu, and Victor 2019). However, the results we obtained do not have much in common with rational reasoning and effectiveness in online shopping goal seeking. Paradoxically, old-fashion style, traditionalism, and inherited habits beat efficiency. The RS awareness does not influence the information search behavior.

H6. Gender is an important variable in many aspects of online shopping, including e-commerce RS (Winoto, Pinata, Tang 2014). The users perceive the RS utility equally, this perception not being influenced by the gender (Gordillo, A., Barra, E., Quemada 2017), as also our hypothesis – GEN −/→ URS – states for influence.

H7. Our study indicates that gender does not affect significantly the shopping frequency in an online environment for young users (median age 20 for women and 22 for men) for influence. In such circumstances, H7 (GEN −/→ OSF) is accepted. This conclusion is also met in (Sramova and Pavelka 2019). Thus, men and women have significant similar behavior according to online shopping in terms of frequency, but there may be severe discrepancies in the item categories they buy. Men buy more frequently gadgets, while women buy clothes or cosmetics.

H8. The ads present onto online shopping websites should lower the IS time especially when the users are aware of their utility (Rezola, Gutierrez, and Linaza 2016). In our study, this hypothesis – UA → IS – is rejected. The theoretical scenario works only on accurate ads customization (Rezola, Gutierrez, and Linaza 2016), since in other cases, including ours, the ads may be irrelevant, annoying, blocked, aggressive, or simply ignored by users due to the good sake of the old-fashion IS habit. Thus, even if the users know the ads' possible utility, their development, or the users' deliberate or accidental behavior annihilate the potential benefits.

Concluding, from the set of 8 hypotheses, 6 of them – H1, H2, H4, H5, H6, H7 – were accepted, while H3 and H8 were rejected. Even though medium value coefficients were obtained between each of the two constructs and IS, in none of the cases these coefficients were significant. We estimated different models, with

different paths to link IS to the two constructs, but in all cases, the effects turned out not to be significant based on the bootstrap simulations or the quality of the models was much lower than the present's one, things that led to their invalidation.

Our focus was on IS and how various factors may influence it. However, the model depicted in Figure 1 reveals an interesting insight. To be useful and accepted, a recommender system must take into account what clients consider important to be presented (CRS $\rightarrow$ UUARS). There is a strong relationship between the two constructs – direct and total effect of 0.783, with a p-value < 0.001. Thus, this result is statistically confirmed.

It means that if users perceive the usefulness, accept and have the intention to use different types of RS, they are aware of what kind of information is used by RS to improve their experience, and also to share some personal implicit or explicit information with the rest of users community. This attitude is fair for the entire community, all users being directly involved in the same online shopping process and all of them contributing individually for the welfare of the entire ecosystem, with the final goal to raise their satisfaction. The perceived ease of use and usefulness on applications empowered by RS strongly affect the intention to use such applications (Chen, Chia-Chen, Tsai 2019). Trust in the recommendation system is an important aspect and influences the acceptance of recommendations. This may embrace the form of a system providing good quality explanations or good quality reviews. Emphasizing the importance of information quality (Chen, Chia-Chen, Tsai 2019) may substantially contribute to the improvement of the user's experience, and a quick acceptance, use, and trust in such systems.

In order to control for the stability of these results, we also included in the analysis several aspects related to the importance of different other RS in comparison to other sources of information, like commercials, information external to the site, the searching system, or even the situation in which the client does not rely on such systems and knows exactly what she/he wants. All of them had very low coefficients, meaning almost no effect on the model. We kept in the analysis the ones with the highest values among them, namely awareness level of utility perceived for the recommender system and the one for the ads on the Internet page. But none of them has a significant impact on the search for information before buying online. An important aspect is that the utility, use, and acceptance of a recommender system is not significantly linked to the perceived utility of the system in respect to other information sources.

5. Conclusions and future work

The present research addresses the issues of IS and RS adoption in e-commerce. The topics became more and more interesting with the development of online buying practices. Still, a review of the most significant studies in the field proves that there are many aspects not enough assessed or debated. Our study comes to compensate for a share of this lack by linking together RS and behavioral aspects related to prior purchase IS.

_____

Six hypotheses that we offered have been confirmed, while two were rejected. However, the most important conclusion of our study is that there is a strong relationship between the most important sources of information used and the perceived usefulness, acceptance, and intention to use a recommender system. But none of them influences the search for information. IS is deep-rooted in young users, this habit being also enriched with the heritage from their parents, old-fashion information seekers in classical commerce. We proved that IS is not dependent on any RS-specific issues, purchase frequency, ads, and age, but is more related to personal client features, and specifically based on gender. Women rely on intuition or impulse in shopping decision, while men are more analytical and IS-oriented. Besides, young people behave similarly in terms of IS and have homogeneous opinions concerning the use and utility of such systems. Moreover, the most relevant aspects that should be considered by the RS are semblable to all individuals in the sample.

But, contrary to what was discussed above, age does significantly influence the frequency of online shopping, even among young people. We assume this is due to a confounding effect. Very young people usually do not have a stable income (as they study) or are involved in very low paid jobs (again, related to the fact that they are studying, so their jobs are either very low-skilled or are part-time). Consequently, even though they search for information as much as the older ones, they lack the financial resources to buy. After finishing their study period, they get more high-skilled jobs, which are better paid, so they start having the financial basis required for purchasing. Such results open a wide discussion about people's preferences and their behavior that goes beyond Business and enters the field of Psychology or beyond, of Philosophy.

IS is strongly related to RS. Therefore, in the future, we aim to examine the relationships to reveal the key factors affecting purchase decision by examining the influence of RS types and various IS sources. Moreover, purchase decision analysis, trust, and privacy topics open the venues to continue research in this complex and challenging field.

REFERENCES

[1] **Blanca** (**2011**),*Age, Gender and Income: Do They Really Moderate Online Shopping Behaviour?* Edited by Julio. *Online Information Review* 35 (1). Emerald: 113–33;

[2] **Bobadilla, Ortega, Hernando, and Gutiérrez** (**2013**), *Recommender Systems Survey. Knowledge-Based Systems* 46 (July). Elsevier Inc.: 109–32;

[3] **Chen, Chia-Chen, Tsai** (**2019**), *Determinants of Behavioral Intention to Use the Personalized Location-Based Mobile Tourism Application: An Empirical Study by Integrating TAM with ISSM. Future Generation Computer Systems-The International Journal of Escience* 96: 628–38;

[4] **Chin** (**2010**), *Bootstrap Cross-Validation Indices for PLS Path Model*

Assessment. In *Handbook of Partial Least Squares: Concepts, Methods and Applications*, edited by *Springer,* 83–97. Berlin;

[5] **Dijkstra, and Henseler** (**2015**), *Consistent and Asymptotically Normal PLS Estimators for Linear Structural Equations.* Computational Statistics & Data Analysis 81: 10–23;

[6] **Gordillo, A., Barra, E., Quemada** (**2017**), *A Hybrid Recommendation Model for Learning Object Repositories. IEEE Latin America Transactions* 15 (3): 462–73;

[7] **Hair, Hult, Ringle, and Sarstedt** (**2017**), *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM).* Sage;

[8] **Henseler, Hubona, and Ray** (**2016**), *Using PLS Path Modeling in New Technology Research: Updated Guidelines. Industrial Management & Data Systems* 116 (1): 2–20;

[9] **Hjorthol** (**2009**), *Information Searching and Buying on the Internet: Travel-Related Activities?*. *Environ Plann B* 36 (January): 229–44;

[10] **Jöreskog** (**1973**), *A General Method for Estimating a Linear Structural Equation System.* University of Uppsala, Department of Statistics, Selected Publications. Uppsala;

[11] **Kluge, Dina** (**2012**), *Online Consumer Behavior and Its Relationship with Socio-Demographics, Shopping Orientations, Need for Emotion, and Fashion Leadership. Journal of Business and Media Psychology*;

[12] **Kumaravel** (**2017**), *Consumer Shopping Behaviour and the Role of Women in Shopping-A Literature Review. Research Journal of Social Science and Management* 7 (December): 50;

[13] **Lu, Wu, Mao, Wang and Zhang** (**2015**), *Recommender System Application Developments: A Survey. Decision Support Systems* 74 (June). North-Holland: 12–32;

[14] **Murali, Vishnu and Victor** (**2019**), *A Collaborative Filtering Based Recommender System for Suggesting New Trends in Any Domain of Research.* In *Proceedings of ICACCS*, 550–53;

[15] **Pathak, Garfinkel, Gopal, Venkatesan and Yin** (**2010**), *Empirical Analysis of the Impact of Recommender Systems on Sales. Journal of Management Information Systems* 27 (2). Routledge: 159–88;

[16] **Pu, Chen, and Hu**. (**2011**), *A User-Centric Evaluation Framework for Recommender Systems.* In *Proceedings of the Fifth ACM Conference on Recommender Systems - RecSys '11*, 157. ACM Press;

[17] **Rezola, Gutierrez, and Linaza** (**2016**), *Automatic Persistent Personalization of Ads in Tourism Websites.* In *Information and Communication Technologies in Tourism 2016*, edited by Inversini and Schegg, 17–30. Cham: *Springer*;

[18] **Roscoe, Grebitus, O'Brian, Johnson, and Kula** (**2016**), *Online Information Search and Decision Making: Effects of Web Search Stance. Computers in Human Behavior* 56 (March). Elsevier Ltd: 103–18;

[19] **Rowlands, Nicholas, Williams, Huntington, Fieldhouse, Gunter, Withey,**

75

Jamali, Dobrowolski, and Tenopir (**2008**), *The Google Generation: The Information Behaviour of the Researcher of the Future.* Aslib *Proceedings* 60 (4): 290–310;

[20] **Sigal** (**2012**), *Gender Differences in Brand Commitment, Impulse Buying, and Hedonic Consumption.* Edited by Ram. *Journal of Product and Brand Management* 21 (3). Emerald Group Publishing Limited: 176–82;

[21] **Sramova, and Pavelka** (**2019**), *Gender Differences and Wellbeing Values in Adolescent Online Shopping.* International Journal of Retail & Distribution Management 47 (6): 623–42;

[22] **Szymkowiak, and Garczarek-Bąk** (**2018**), *Gender Differences in E-Commerce*. 4 (January): 250–61;

[23] **Velaudham, and Baskar**. (**2015**), *Influence of Gender in Consumer Buying Behavior towards Air Conditioner in Chennai City.* Annamalai Journal of Management*, April, 3–9;

[24] **Wan, Nakayama, and Sutcliffe** (**2012**), *The Impact of Age and Shopping Experiences on the Classification of Search, Experience, and Credence Goods in Online Shopping.* Inf. Syst. E-Business Management 10 (March): 135–48;

[25] **Winoto, Pinata, Tang** (**2014**), *The Ethics of a Recommendation System.* In *15th International Conference on Web-Age Information Management (WAIM)*, 287–98.