Shifts in consumer behavior towards organic products: theory-driven data analytics

Firouzeh Taghikhah [a*], Alexey Voinov [a,b], Nagesh Shukla [a], Tatiana Filatova [b,a]

[a] *Center on Persuasive Systems for Wise Adaptive Living, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW 2007, Australia,*
[b] *Department of Governance and Technology for Sustainability, University of Twente, Netherlands*
Email: Firouzeh.th@gmail.com, aavoinov@gmail.com, Nagesh.Shukla@uts.edu.au, Tatiana.Filatova@uts.edu.au

* Corresponding Author

Abstract

Consumer behavior is key in shifts towards organic products. A diversity of factors influences consumer preferences, driving planned, impulsive, and unplanned purchasing decisions. We study choices among organic and conventional wine using an extensive survey among Australian consumers (N=1003). We integrate five behavioral theories in the survey design, and use supervised and unsupervised machine learning algorithms for analysis. We quantify a gap between intention and behavior, and emphasize the importance of cognitive factors. Findings go beyond correlation to the causation of behavior when combining predictive prowess with explanatory power. Results reveal that affective factors and normative cues may prompt unplanned and spontaneous purchasing behavior, causing consumers to act against their beliefs.

Keywords:

Organic food; emotion; habit; impulsive purchasing; data mining; explainable artificial intelligence.

1. Introduction

Demand-side policies can significantly contribute to tackling climate change issues and managing environmental resources. The practical implementation of these policies is vital for conserving the ecological-service for the future and controling the exploitation of natural capital assets. The challenge for policymakers is how to change the consumption patterns and increase the demand for environmentally-friendly products, which triggers the market forces to make the adoption of sustainable practices economically attractive to suppliers and producers.

For example, since the introduction of chemicals in the 19th century, viticulture has significantly contributed to a wide range of environmental issues, particularly those related to land and water pollution. By excluding agrochemicals from vineyards, organic agriculture helps preserve biodiversity and the overall quality of agroecosystems (Rugani et al. 2013). Wines produced with organically grown grapes have a higher content of antioxidants (30%) (Vrček et al. 2011) and lower content of orchatoxins (Gentile et al. 2016). Consumer choices and their willingness to pay (WTP) more for organic wines can support farmers in expanding organic vineyards (Taghikhah et al. 2020b). In fact, it can be a game-changing strategy contributing to the economy, ecology, and society.

The demand-side argument is prominent in the ongoing debate about how to increase the organic wine market share. It highlights the need to investigate the characteristics of consumer segments willing to purchase the organic food, identify the factors influencing their decisions, target influential factors in each consumer segment, and develop segment-specific marketing strategies if we wish to nudge behavior

2

towards organic consumption. Prior studies report various factors drivers of consumers' decisions in purchasing organic wine. The key factors include price (Panzone 2014), perceived that health and environmental benefits (Loose & Lockshin 2013), region of origin (Trinh et al. 2019; Yang & Paladino 2015), brand (Ryan & Casidy 2018), superior taste and quality (Kim & Bonn 2015), as well as socio-demographics including age, gender, and income (D'Amico et al. 2016). More recent studies have highlighted the relative importance of occasions like hosting friends and gift-giving (Boncinelli et al. 2019), wine consumption and shopping frequency (Pomarici & Vecchio 2014), and drinking frequency (Pomarici et al. 2016) as predictors of consumers shift from conventional to organic wine.

In the context of pro-environmental behavior, the literature highlights a discrepancy between consumers' stated intentions and their actions, known as the intention-behavior gap. Even though consumers demonstrate WTP for products with sustainability cues, and their intentions are high, these do not necessarily translate into the actual purchasing behavior. With regard to organic wine, the literature focuses on identifying determinants of WTP; yet, this is rarely differentiated from real purchasing behavior. An exception is a study by Schäufele and Hamm (2017), who confirm the inconsistencies between intentions to purchase organic wine and the actual behavior among low-income consumers, identifying prices as the primary purchasing barrier. Poor quality and inferior taste are other reported reasons for avoiding organic purchases (Mann et al. 2012; Stolz & Schmid 2008).

Impulsive and unplanned purchasing behaviors appear to interrupt the intention-behavior relationship. According to the literature on consumer behavior, affective factors as well as cognitive and normative factors, can trigger behavior change

(Russell et al. 2017). The non-cognitive factors, such as emotions, impulse tendencies, and personal goals, may underlie the failure to translate consumers' intentions into actions. Yet, to the best of our knowledge, there have been no quantitative studies to date that have investigated the relative importance of these factors as they relate to organic wine purchasing.

Moreover, quantitative research predicting consumers' intentions and behavior for purchasing organic wine has, to date, been dominated by statistical models. While these models can successfully reveal the relationship between variables, their predictive power and accuracy, as compared to machine learning (ML) algorithms, are low, especially when dealing with a high number of observations and attributes. Indeed, they are powerful tools in identifying unexpected patterns and emergent proprieties of underlying phenomena of interest. Pattern verification is a useful application of ML for confirming whether suggested behavioral theories exist.

Our study aims to explore the determinants of heterogeneity in organic food purchasing intentions and behaviors. To identify the behavioral factors driving purchasing decisions, we consider behavior change theories from psychology and developed a conceptual framework that integrates five relevant theories. We focus on organic wine as a case study and surveyed 1,003 Australian consumers living in the City of Sydney. The collected data enable to quantitatively assess the impact of socio-demographics, shopping and wine consumption patterns, and behavioral factors on consumers' stated intentions and behavior for purchasing organic wine. Our findings reveal factors that cause the intention-behavior gap in pro-environmental food consumption.

This article makes a number of innovative contributions to the literature on consumer behavior.

(i)   It examines the influence of affective factors, including emotions, impulse tendencies, and personal goals, as well as cognitive factors, especially social norms, in the context of wine purchasing. To the best of our knowledge, this is the first study that has fully explored how this set of attributes affects preferences for organic food by integrating the strength of multiple behavioral theories. While many papers in this field focus on studying willingness to pay and intentions for buying organic wine, we focus on the gap between intentions and behavior.

(ii)   It goes beyond the traditional analysis in empirical consumer behavior studies by applying both supervised and unsupervised ML methods. Besides increasing the accuracy of predictions, we explain why AI arrived at a specific decision by identifying the most influential factors. These methodological advances provide new insights into different consumer segments, identify the causality and mechanisms of decisions related to organic products, and, most importantly, verify whether the behavioral patterns will continue to function as expected over time. We apply explainable AI techniques to open the "black box" of ML in consumer behavior area so that the results can be understood by humans.

(iii) It provides empirical insights for industry and policymakers when promoting organic food and can contribute to the facilitation of demand-side solutions in the transition to sustainable agriculture. The demand-side policy is attractive to producers (farmers) and policymakers because the diffusion of organic food consumption is expected to reduce the cost of developing organic farms and

help overcome the trade-offs between economic, health, and environmental goals.

The remainder of this paper is organized as follows. We begin with explaining the proposed theoretical framework used to develop the survey (Section 2) and describe the methodological aspects, data collection, and the analysis process (Section 3). Section 4 presents the results, and Section 5 discusses them in the context of existing literature. We conclude with discussing the implications for practice and outline potential avenues for future research.

2. Theoretical framework

Behavior change theories are widely applied to understand the internal, external, and interpersonal factors driving individual actions. To provide a more holistic perspective on pro-environmental purchasing behavior, we refer to the principles of Stern's buying theory (Stern 1962) that classify decisions as planned, impulsive, and unplanned. Planned purchasing behavior is time-consuming, information-searching, norm-dependent, semi-bounded rational decision making. In contrast, unplanned purchasing behavior refers to decisions driven by atmospheric store-related stimuli (e.g., promotions, posters) or habits (context-dependent stimuli) without any preliminary planning or even actual need. Impulsive purchasing refers to rapid, spontaneous decisions driven by an individual's impulse tendency (i.e., a sudden, irresistible urge). Internal stimuli cause impulsiveness in response to mood swings, excitement, or unpleasant situations. Research shows that the use of sensory cues, such as the addition of scent or music, can influence consumers' emotions and impulse purchasing behavior (Helmefalk & Hultén 2017).

To encompass the complexity of consumer behavior and various stages that lead one to a purchasing decision, we develop a theory-grounded framework (Figure 1). Namely, we combine the strength of relevant theories to understand the influence of cognitive and affective factors behind a variety of purchasing behavior:

1. Theory of Planned Behavior (TPB) (Ajzen 1991) to account for factors driving planned decisions,

2. Theory of Interpersonal Behavior (TIB) (Triandis 1977) to integrate the influence of emotions,

3. Impulsive Buying Theory (IBT) (Stern 1962) to capture factors driving impulsive purchasing,

4. Alphabet Theory (AT) (Zepeda & Deal 2009) to integrate the role of habits,

5. Goal Framing Theory (GFT) (Lindenberg & Steg 2007), to account for a variety of goals.

This framework allows to comprehensively explore purchasing decisions in different situations (e.g., shopping environment), understand the influence of context on the action (e.g., occasions), identify potentials to influence preferences (e.g., social media), and bridge the gap between intention and behavior. Appendix A1 describes these theories and the literature review on consumer behavior for organic food in detail; Taghikhah et al. (2020a) discuss the further rationale for integration.
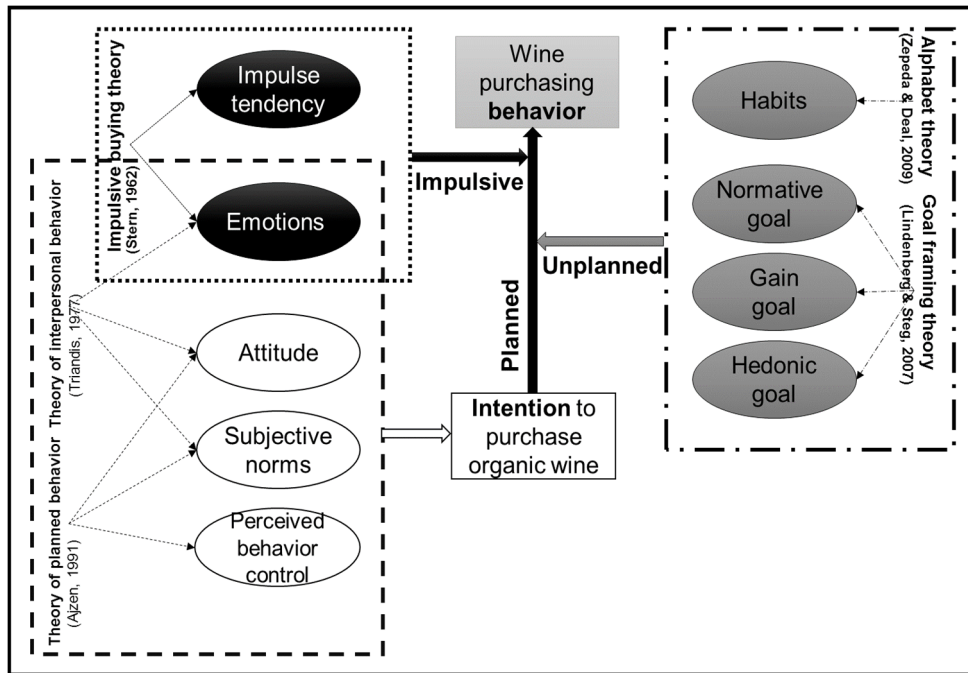
Figure 1. Conceptual model of the determinants of organic wine purchasing behavior.

3. Methodology

3.1. Data collection

Relying on the theoretical framework (Figure 1), we design a questionnaire to elicit data on the corresponding variables (discussed in detail in Appendix A2). The questionnaire includes 7 sections consisting of 35 questions about (i) socio-demographic characteristics (10 questions), (ii) shopping and drinking-related style (7 questions), (iv) habits (1 question), (v) attitudes (3 questions), PBC (2 questions), (vi) social networks (3 questions), (vii) personal goals (4 questions), (viii) emotions, and (ix) impulse tendency (1 question). We align the questions for assessing habits, shopping patterns and emotions and impulsiveness with previous studies (e.g., Verplanken and Orbell (2003), Ogbeide (2013), Watson et al. (1988)). We use a multiple-question approach in assessing each variable to improve the quality of

results. They were measured on a Likert scale ranging from 1–5. An expert panel consisting of two academic researchers and one practitioner reviewed and validated the questions.

In September 2019, the online survey was conducted in 32 suburbs of the City of Sydney through Qualtrics online customer panel (https://www.qualtrics.com). The respondents (18+ years old) were chosen randomly. We ran a one-stage pilot study (N_{pilot}=50) to test the consistency of questions and responses. We check the internal consistency of the survey using Cronbach's alpha. The test result shows a good consistency at 0.766, which confirms the validity of the designed questions for assessing factors.

We acknowledge that self-reported items do not always reflect the actual behavior in stated-preferences studies like surveys. However, the choice of what wine to buy is a regular decision, which stays in the memories of consumers. In this case, consumers were not thinking of a hypothetical decision when filling in our questionnaire; they were explicitly asked about a decision that is learnt and is practiced on an almost weekly basis. Our questionnaire explicitly asked respondents to remember whether they had purchased organic wine and what share of their actual past purchases was organic.

3.2. Methods of analysis

3.2.1. Data pre-processing and correlation analysis

For standardization, the variables containing discrete sequences of values, such as age, shopping frequency, shopping size, family size, etc. are normalized with the min-max normalization method to values between 0 and 1, to scale the differences in the

ranges of the continuous variables. We apply a binary encoding procedure to all categorical variables in our dataset to convert them into binary variables. Our final data set includes 1003 responses and 89 variables. The *descriptive statistics* reveal the characteristics of respondents and the distribution of key variables (Section 4.1). We use *Spearman's rank correlation* to assess the strength and direction of the relationships between the nine latent variables representing behavioral factors (Section 4.2). This allows us to validate the proposed conceptual framework (Figure 1). We consider coefficients greater than +0.4 (and smaller than -0.4) as indicators of a strong relationship, while those between 0.2 and 0.4 (-0.4 and -0.2) as of a moderate correlation. The strength of a correlation depends on the context and sample size. As common in social sciences (Cohen (1992, 2013), coefficients around 0.3 and 0.5 represent moderate and strong correlations, respectively. However, for large sample sizes, a moderate correlation coefficient can be considered as significant as a strong correlation in a small sample, meaning that this relationship is unlikely to occur by chance.

3.2.2. Supervised learning: Classification

To reach beyond correlations towards implying causation of the behavior, we use *classification*, the most commonly applied supervised learning approach, to predict the probability that a consumer prefers organic to conventional wine. We consider 6 classes of intentions and 5 classes of behavior for purchasing organic wine. The consumers with no willingness to pay for organic wine are labelled as class (1) and those with willingness to pay for organic wine up to 10%, 20%, 30%, 40%, and 50% are labeled class (2), (3), (4), (5), and (6), respectively. Similarly, for predicting behavior, labels are assigned to consumers who purchase only conventional wine

(class (1)), organic wine up to 25% (class (2)), organic wine between 25% and 50% (class (3)), organic wine between 50% and 75% (class (4)), and organic wine 75% or more (class (5)). We test both *parametric (logistic regression, LR)* and *nonparametric (support vector machine, SVM)* classification algorithms (Cortes & Vapnik 1995), as well as the *Decision Tree (DT)* (Quinlan 1990) and *Random Forest (RF)* (Ho 1998) algorithms to identify the best performing method for classification of our data. Appendix A3 provides the details of these classification algorithms.

Parametric algorithms assume that a linear combination of variables and coefficients can be fitted to a line, whereas nonparametric algorithms construct the model based on the similarities between patterns in data, without making any assumptions. While the selection of methods depends mainly on the characteristics of the data, higher flexibility and predictive power are generally expected for nonparametric algorithms. However, data requirements and overfitting issues should be carefully controlled when using these algorithms. SVM finds the best prediction model using an optimization process to minimize the error function. DT uses conditional control statements in a flowchart-like structure to predict outcomes. Previous studies have reported better performance of ensemble methods like RF for classification, where multiple predictive models (in this case, trees) vote for the class assigned to a given sample so as to decrease biases and variances in predictions. The partitioning ratio for training and testing for each of these methods is set to 70% vs. 30%, respectively.

3.2.3. Unsupervised learning: Clustering

To identify hidden patterns or distinct groups based on their similarities in our dataset, we use *clustering*, the most common unsupervised learning approach for exploratory data analysis. Density-based clustering algorithms such as DBSCAN automatically

detect the number of clusters and are suitable for cases where the clusters are not compact and well-separated (Ester et al. 1996). In contrast to ad-hoc methods that divide records based on one attribute, this method includes all attributes when computing the cohort outliers. Hence, one may reveal hidden patterns and groups in survey data and study their characteristics along known dimensions to potentially attribute various behavioral factors to consumption patterns. Partitioning methods (e.g., K-means) and hierarchical clustering work by finding spherical-shaped clusters or convex clusters, while DBSCAN identifies arbitrary-shaped clusters under fewer restrictions. However, since our database is highly dimensional and scattered, this algorithm fails to detect clusters of consumers with similar properties. Hence, we utilize its extension – *HDBSCAN* - designed to deal with high-dimensionality. HDBSCAN uses a technique to hierarchically represent every possible cluster generated by DBSCAN and extract a set of flat clusters (Campello et al. 2013). We applied HDBSCAN on the pre-processed dataset with 89 dimensions (Section 3.2.1). As the algorithm fails to extract meaningful clusters when using all 89 dimensions (the noise is 70%), we further use the *principal component analysis (PCA)* method to gradually reduce the dataset dimensions, minimize the clustering noise, and increase the density of resulting clusters. PCA identifies six dimensions where the clustering noise is the lowest, while its density is the highest. Further, relying on the HDBSCAN recommendations for selecting parameters, we use the approach proposed by Rahmah and Sitanggang (2016) to tune its hyper-parameters. Appendix A4 provides details of HDBSCAN and the settings for its hyper-parameters.

4. Results

4.1. Descriptive analysis

  Table 1 compares the socio-demographic characteristics of the City of Sydney population (collected from ABS - 2016 census) with the collected sample (own survey - 2019). The results indicated that, except for the educational level, the sample is representative of the population. Nevertheless, as discussed in Section 4.2 below, education is only moderately correlated with intention and behavior, and indicating that the possible education gap between our sample and the local population should not affect the main conclusions of the study.

Table 1. Socioeconomic distribution in the City of Sydney (LGA) and the survey sample.

Factors	City of Sydney LGA	Survey sample
Total number of households	85,423	1,003
Gender • Female (%) • Male (%)	47% 53%	41% 59%
Median age group	30-40 years old	36-45 years old
Median total income	AU$75,001 to AU$150,000	AU$75,001 to AU$150,001
Average household size	2	2
Education level • Postgraduate Degree levels (%) • Graduate Degree level (%) • School education level (%)	17.9% 39.9% 42.2%	51.1% 36.8% 12.1%

  Table 2 provides summary statistics for the socio-demographic characteristics of the survey respondents. It shows that (1) gender statistics are balanced and can adequately reflect differences, (2) the majority of consumers are highly educated and work full time in the management and engineering occupations, (3) the income level of more than two-thirds of the consumers is higher than the average income, between

AU$ 75 and AU$ 250 thousand, and (4) about half of the respondents are singles and half are couples.

Regarding consumers' patterns of wine purchasing and consumption, the results indicate that the majority of respondents surveyed visit wine shops more than once a week and purchase more than five wine bottles per month. More than 70% of consumers purchase the same brand of wine quite often and report drinking wine 2 to 5 times a week.

Table 2. Socio-demographic characteristics of surveyed consumers.

Socio-demographic items					
Gender		**Household annual income**		**Household size**	
Male	59%	less than 45 thousand AU$	9%	One	15%
Female	41%	45-75 thousand AU$	14%	Two	29%
Age		75-150 thousand AU$	38%	Three	26%
18-25 years	10%	150-250 thousand AU$	26%	Four	0%
26-35 years	25%	More than 250 thousand AU$	13%	Five	13%
36-45	39%	**Occupation**		Six	47%
46-55	16%	Engineering	19%	Seven and more	18%
56-65	6%	Education	12%	**Employment status**	
66 and more	4%	Sales and service	15%	Full-time employed	78%
Education		Management	29%	Part-time employed	10%
Primary	2%	Other	26%	Retired	4%
Secondary	10%			Student	5%
Graduate	39%			Unemployed	3%
Post-Graduate	51%				

Table 3 presents the summary statistics for the behavioral factors related to purchasing behaviors. The results showed that, on average, consumers have positive attitudes towards organic wine and positive emotions during shopping. Most

consumers report high habitual (0.69) and low impulsive purchasing (0.29). Consumers distinguish between organic and conventional wine and like the taste of organic (more than 0.72), whereas the advice of staff, choice of other people at the shop, and social media are not significant predictors of wine choice (less than 0.33). While wine availability is important to our respondents, they indicate no concern for price (comparing 0.56 to 0.37).

Table 3. Importance of behavioral factors among survey respondents.

Behavioral factors	Sub factors (related theory)	Measures	Average	Standard deviation
Cognitive	*Attitude (TPB)*	Trust on organic wine	0.74	(0.19)
		Environmental knowledge of organic wine	0.73	(0.17)
		Health knowledge of organic wine	0.72	(0.16)
	Perceived Behavioral Control (TPB)	Importance of wine price	0.37	(0.35)
		Importance of wine availability	0.56	(0.28)
	Habit (AT)	Automaticity of purchasing	0.69	(0.21)
	Hedonic goals (GFT)	Taste	0.85	(0.29)
		Difference and distinction	0.78	(0.36)
		Likeness	0.72	(0.39)
	Gain goals (GFT)	Change of price at the shop (switch preference)	0.42	(0.35)
		Change of availability at the shop (switch preference)	0.32	(0.37)
Normative	*Social norms (TPB)*	Frequency of socializing about wine	0.57	(0.3)
		Purchasing wine for occasions	0.65	(0.47)
		Advice of family and friends	0.75	(0.15)
	Normative goals (GFT)	Staff and others at shop	0.33	(0.23)
		Social media	0.17	(0.26)
Affective	*Emotions (TIB & IBT)*	Positive emotions	0.75	(0.25)
	Spontaneous urge (IBT)	Impulse tendency	0.29	(0.3)

Our results indicate a significant gap between intention (Figure 2.a) and behavior in organic wine purchasing (Figure 2.b). We consider WTP more for organic wine as an indicator of individual intention and the proportion of purchased organic wine in the shopping basket as an indicator of actual behavior. The respondents are asked to assume that the average price of a wine is $10 per bottle. More than 80% of consumers have a positive intention for purchasing organic wine (Figure 2.a). Interestingly, only 4% of consumers are exclusively organic wine buyers (i.e., purchase organic in 75-100% of cases), with an additional 17% of consumers

indicating they are frequent organic wine buyers (i.e., between 50-75% of their wine shopping basket is organic). Still, 60% of respondents indicate that less than 50% of their wine shopping basket is organic, while 20% had never purchased organic wine before. Our further analysis aims to explore what stands behind various consumption decisions.
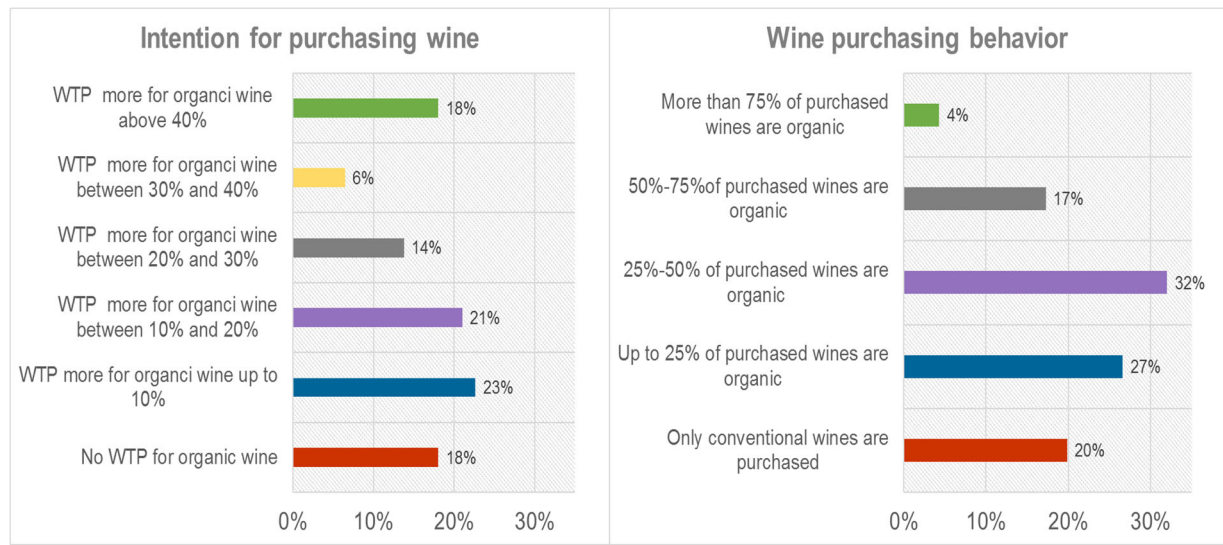
(a) Intention to purchase organic wine   (b) Wine purchasing behavior

Figure 2. Distribution of intention and behavior for purchasing organic wine (in percentage; N=1003).

4.2. Correlation analysis

Table 4 presents the correlation matrix for behavioral factors, perceived behavioral control (PBC), social norms, emotions, habits, impulse tendencies, hedonic, gain, and normative goals. Overall, attitudes and emotions are the most strongly correlated with the other variables, while the weakest correlations are between the gain goals and other variables. We find that habits are strongly positively correlated with hedonic and normative goals. As expected, habits correlate negatively with the impulse tendency (-0.46), meaning those who stick to certain products are less prone to spontaneous shopping. In general, customers with negative attitudes and feelings, and who are against norms and habits, tend to purchase wine more impulsively.

16

Table 4. Triangular matrix of correlations among latent constructs of behavior (bold, underlined values represent strong correlations, and italic values show moderate correlations).

	Variables	Attitude	PBC	Hedonic goal	Gain goal	Habits	Social norms	Normative goal	Emotions	Impulse tendency
Cognitive	Attitude	–								
	PBC	*0.31*	–							
	Hedonic goal	**0.48**	*0.23*	–						
	Gain goal	*-0.24*	-0.1	*-0.22*	–					
	Habits	**0.59**	*0.23*	**0.42**	-0.18	–				
Normative	Social norms	**0.47**	*0.24*	**0.46**	-0.16	**0.49**	–			
	Normative goal	**0.48**	0.18	**0.36**	-0.13	**0.52**	**0.5**	–		
Affective	Emotions	**0.56**	*0.24*	**0.45**	*-0.2*	**0.61**	**0.49**	**0.54**	–	
	Impulse tendency	**-0.42**	-0.18	*-0.24*	0.13	**-0.46**	*-0.35*	**-0.45**	**-0.51**	–
	Variables	Attitude	PBC	Hedonic goal	Gain goal	Habits	Social norms	Normative goal	Emotions	Impulse tendency
		Cognitive					Normative		Affective	

Furthermore, we calculate the correlation matrix for the relationships between wine purchasing intentions and behavior and all the database variables. Table 5 shows that both intention and behavior are strongly and positively correlated with hedonic goals (likeness, taste, distinction), attitudes (health belief, environmental belief, and trust), habits, emotions, social norms (special occasion and socializing), and shopping and drinking-related patterns (wine drinking frequency, purchasing frequency, shopping size, time spent at the wine shop, and the average price paid for wine). At the same time, demographics, including gender, family size, education, and income, are moderately correlated with intention and behavior. Moreover, the relationships between impulse tendency, wine substitution (if the products are unavailable), and organic wine purchasing intention and behavior are negative. Appendix B presents the details of the correlation analysis for all database variables.

Table 5. Correlations between intention and behavior for purchasing organic wine and other variables, where strong correlations are bold and underlined, and moderate correlations are in italics.

			Organic purchasing intention	Organic purchasing behavior
Socio-demographic factors		Gender	*-0.28*	*-0.33*
		Retired	*0.31*	*0.38*
		Household size	*0.2*	*0.32*
		Average household education	*0.37*	*0.34*
		Average household income level	*0.28*	*0.31*
Shopping and drinking-related patterns		Average wine shopping size per month	**0.43**	**0.51**
		Wine drinking frequency	**0.45**	**0.53**
		Wine purchasing frequency	**0.5**	**0.64**
		Time spent in wine shops	**0.42**	**0.45**
		Loyal to certain brand of wine	*0.26*	*0.28*
		Average price paid for wine	**0.54**	**0.6**
		Maximum price willing to pay for wine	*0.26*	*0.22*
Behavioral factors	**Cognitive**	Like organic wine	**0.49**	**0.61**
		Distinction between organic and conventional wine	**0.47**	**0.51**
		Perceive organic wine tastier	**0.48**	**0.56**
		Habitual wine purchasing	**0.45**	**0.53**
		Environment belief for organic wine	**0.57**	**0.5**
		Health belief for organic wine	**0.53**	**0.51**
		Trust in organic wine	**0.59**	**0.56**
		Price importance for purchasing wine	*0.32*	*0.29*
		If price increases, cheaper substitution	*-0.27*	*-0.2*
		If price increases, no substitution	*-0.32*	*-0.19*
		If price increase, loyalty	**0.51**	*0.33*
		If unavailable, no substitution	*0.28*	*0.34*
		If unavailable, cheaper substitution	*-0.2*	*-0.22*
		If unavailable, expensive substitution	-0.15	*-0.2*
	Normative	Influence of family	*0.37*	**0.4**
		Influence of friends	*0.33*	*0.35*
		Influence of other shopper	*0.39*	**0.46**
		Influence of social media	**0.46**	**0.53**
		Frequency of talking about wine when socializing	**0.41**	**0.45**
		Organic wine for special occasion	**0.51**	**0.66**
	Affective	Positive emotions during shopping	**0.46**	**0.63**
		Impulsive/spontaneous shopping	*-0.31*	*-0.29*

4.3. Supervised machine learning: Classification analysis

While correlation analysis for the entire dataset elicits only linear, rough associations, its results cannot be used for describing nonlinear relationships and making predictions. Moreover, correlations analysis fails to describe the causalities. Hence, to imply the causation of organic wine purchasing and select the classification algorithm with the highest accuracy, efficiency, and prediction power, we compare the performance of SVM, LR, DT, and RF in predicting consumers' intentions (4.3.1) and behavior (4.3.2). The comparison helped us to select the best performing algorithm in our survey data, understand the causal factors of organic wine buying behavior, and derive predictive models for consumers' preferences.

4.3.1. Predicting consumers' intentions to purchase organic wine

We test the considered supervised algorithms on the 6 classes of intentions (Section 3.2.2) and also combine classes 2 and 3 as well as 4 and 5 to decrease granularity (Figure 3). The highest accuracy in predicting the likelihood that a consumer will have an intention to buy organic products is achieved if we consider 4 combined classes of intention: "not willing to pay" (a premium), "willing to pay 10% and 20% more", "willing to pay 30% and 40% more", and "willing to pay 50% and higher more." In all cases, RF outperform the other algorithms (DT, SVM, and LR), while LR had the lowest accuracy (Figure 3). Nonparametric algorithms are better able to handle homogeneity amongst classes, resulting in higher accuracy and higher efficiency in processing complex and highly dimensional datasets. Appendix C1 provides the details of the analyses and the decision tree resulting from RF model for predicting 4 classes.
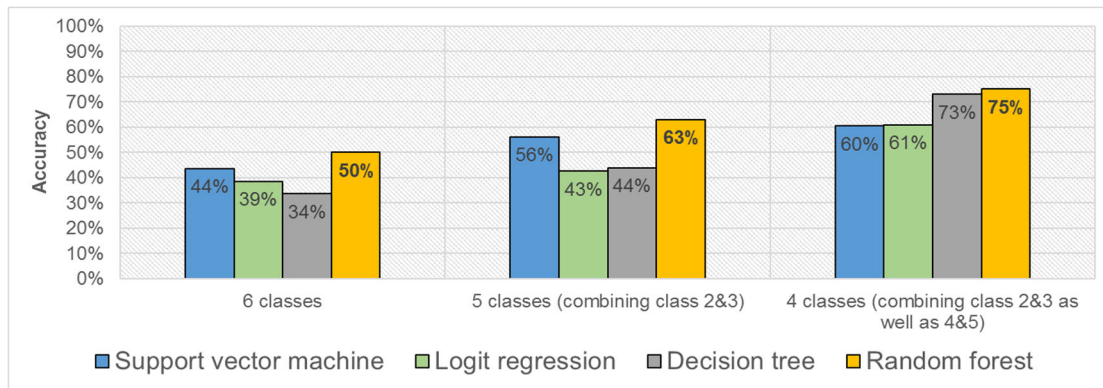
Figure 3. Comparing the performance of the algorithms (i.e., support vector machine (SVM), logit regression (LR), decision tree (DT), random forest (RF)) in predicting consumers' intentions across three models. The original 6 classes of intention range from not willing to pay a premium for organic products (class 1) to willing to pay more than 50% for organic (class 6).

Apart from delivering predictive models, RF provides a deeper understanding and useful information about the relative importance of different variables affecting overall accuracy (Table 6). We find that for organic wine intention, consumers' trust in organic wine has the highest predictive power, followed by environmental belief in organic wine and the average price paid for a bottle of wine (importance weights varied between 0.04 and 0.06 in the three models). On the contrary, factors such as age, loyalty, wine availability, and special occasions are least important (importance weight of 0.02, only in one model). Besides trust in organic farming, environmental belief about organic wine, positive emotions, higher payment for wines, more hedonic motivations, habitual purchasing, and high-frequency wine drinking and purchasing are associated with greater intention to purchase organic wine.

Table 6. The importance of factors in predicting intention according to the Random Forest analysis (variables repeated in the three models are indicated with *; the most important factor and numbers are underlined and bolded). The numbers indicate the weights, where 0.06 has the highest and 0.02 has the lowest influence on the predictions.

Factors		Variables used in RF model	Importance in 6 class model	Importance in 5 class model	Importance in 4 class model
Behavioral factors	Cognitive	**Like** organic wine*	0.02	0.03	0.02
		Perceive organic wine **tastier**	-	0.03	0.03
		Trust in organic wine *	***0.05***	***0.05***	***0.06***
		Environmental belief about organic wine*	**0.05**	**0.05**	0.05
		Health belief about organic wine	-	0.03	0.03
		Habitual wine purchasing*	0.03	0.03	0.03
		Distinction between organic and conventional wine	0.02	-	0.02
		Wine price importance	0.02	-	0.02
		Wine availability importance	-	-	0.02
	Normative	Talking about wine when **socializing** *	0.02	0.03	0.03
		Organic wine for **special occasion**	-	-	0.02
		Family and friend influence	0.02	-	0.02
		Other shoppers influence	0.02	-	0.02
		Wine shop **staff** influence	0.02	0.03	-
		Social media influence on wine choice	-	-	0.02
	Affective	Positive **emotions** *	0.03	0.04	0.05
		Impulsive shopping tendencies*	0.03	0.03	0.02
Shopping and drinking-related patterns		**Average price** paid for wine*	***0.05***	0.04	***0.06***
		Time spent in wine shop *	0.02	0.03	0.02
		Wine **purchasing frequency** *	0.03	0.04	0.04
		Average wine **purchasing size** *	0.02	0.03	0.04
		Wine **drinking frequency** *	0.03	0.03	0.03
		Frequency of comparing different wine **prices**	0.02	-	0.02
		Loyalty to a certain brand	-	-	0.02
Socio-demographic factors		Household average **income**	0.02	-	0.02
		Household highest **education**	-	-	0.03
		Age	0.02	-	0.02
		Household size	0.02	-	0.02
		Gender	-	-	0.02

4.3.2. Predicting consumers' likelihood of purchasing organic wine

We assess the accuracy of the predictive models of the different algorithms for estimating the probability of purchasing organic wine. Similar to intention prediction, RF outperformed the other algorithms, but SVM had the worst performance. Moreover, DT and LR demonstrated comparable performance, except in predicting 3 classes, where DT outperformed (Figure 4). Appendix C2 provides the details of the analyses, and the decision tree resulted from RF model for predicting 3 classes.

Furthermore, we measure the importance of all predictor variables and keep the significant variables in the model. However, there is no full agreement among models about the importance of the variables. For example, the 5-class model indicates that positive emotions and the average price paid for wine had the strongest influence, while the 4-class model indicates that special occasion is the most important factor (for more details, please refer to Appendix C3). Thus, we test the performance of the models when the intention variable is included in our analysis as another predictive factor.
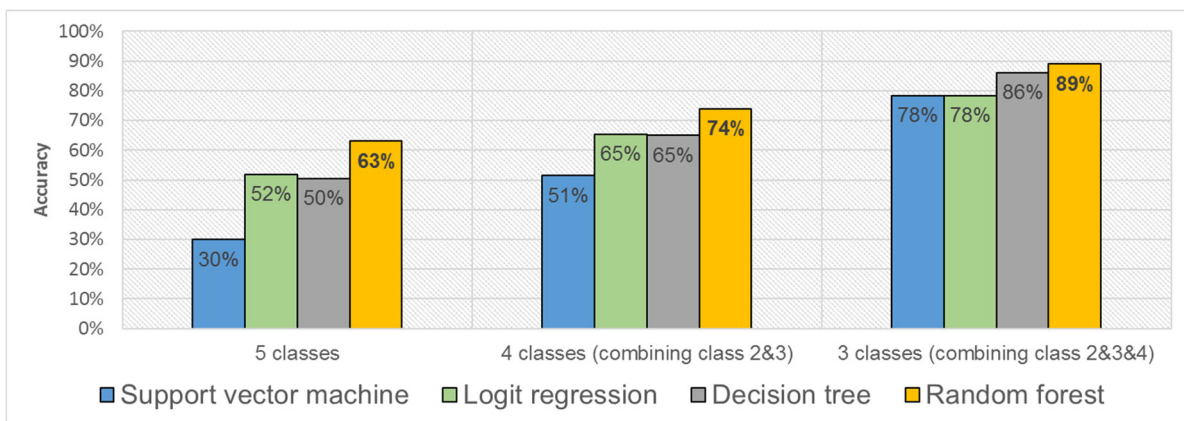
Figure 4. Comparing the performance of the different algorithms (i.e., support vector machine (SVM), logit regression (LR), decision tree (DT), random forest (RF)) in predicting wine purchasing behavior.

Regarding the model accuracy, the inclusion of intention leads to no improvements. However, we find that the average price paid for wine is consistently the most important factor in predicting organic wine behavior, as shown in Table 7 (importance weights between 0.07 and 0.1). Shopping and drinking-related patterns play a similar role in predictor behavior, as observed in relation to intention. Consumers who more frequently purchased more bottles of wine, reported drinking more often, and spend more time at the shops were more likely to purchase organic wine. Behavioral factors, including cognitive (i.e., intention, attitude, habits), normative (i.e., purchase occasions, social media), and affective (only emotions) are other emergent proxies for

22

organic wine purchasing behavior. Finally, socio-demographic factors appeared to be unimportant in predicting purchasing decisions.

Table 7. The importance of factors in organic wine purchasing behavior according to random forest analysis (variables repeated in three models are indicated with * and the most important factor is underlined).

Factors			Variables used in RF model	Importance in 5 class model	Importance in 4 class model	Importance in 3 class model
Behavioral factors	Cognitive		**Intention** for purchasing wine*	0.05	0.05	0.06
			Trust organic wine*	0.03	0.02	0.02
			Health belief about organic wine*	0.03	0.03	0.02
			Environmental belief about organic wine*	0.03	0.02	0.02
			Habitual wine purchasing*	0.03	0.04	0.03
			Like organic wine*	0.03	0.03	0.07
			Distinction between organic and conventional wine	-	-	0.05
	Normative		Influence of **social media***	0.03	0.04	0.04
			Organic wine for **special occasion***	0.04	0.03	0.08
			Influence of **other shoppers**	0.02	-	-
	Affective		Positive **emotions***	0.04	0.04	0.03
Shopping and drinking-related patterns			_**Average price paid for wine***_	_**0.07**_	_**0.10**_	_**0.09**_
			Wine **purchasing frequency***	0.04	0.04	0.06
			Time spent in wine shop*	0.03	0.06	0.04
			Wine **drinking frequency***	0.03	0.03	0.04
			Average wine **purchasing size***	0.03	0.04	0.03
Socio-demographic factors			**Income**	0.02	-	-

4.4. Unsupervised machine learning: Cluster analysis

Here, we determine how the data is distributed in the space, explore what groups of similar examples exist within the data, and examine whether their characteristics can be described by behavioral theories. The HDBSCAN method identified three hidden heterogeneous clusters of consumers (Figure 5). The size of each cluster varied from a minimum of 63 (7%) for cluster 1 to a maximum of 326 (33%) and 327 (33%) for clusters 2 and 3, with 29% of data labelled as noise. Although this percentage of noise

may seem high, the literature (e.g., Chen et al. (2018) and Maurus and Plant (2016)) indicated that such a level of noise in the data is common in density-based algorithm studies. We compared the characteristics of clusters in terms of the different variables. Clusters exhibit significant differences in terms of demographics (e.g., income, education), behavioral factors (e.g., attitudes, habits, emotions), and shopping and drinking-related patterns (e.g., wine drinking, purchasing frequency), see Figure 6. We label these clusters as non-organic (Section 4.4.1), occasional organic (Section 4.4.2), and organic segments (Section 4.4.3).
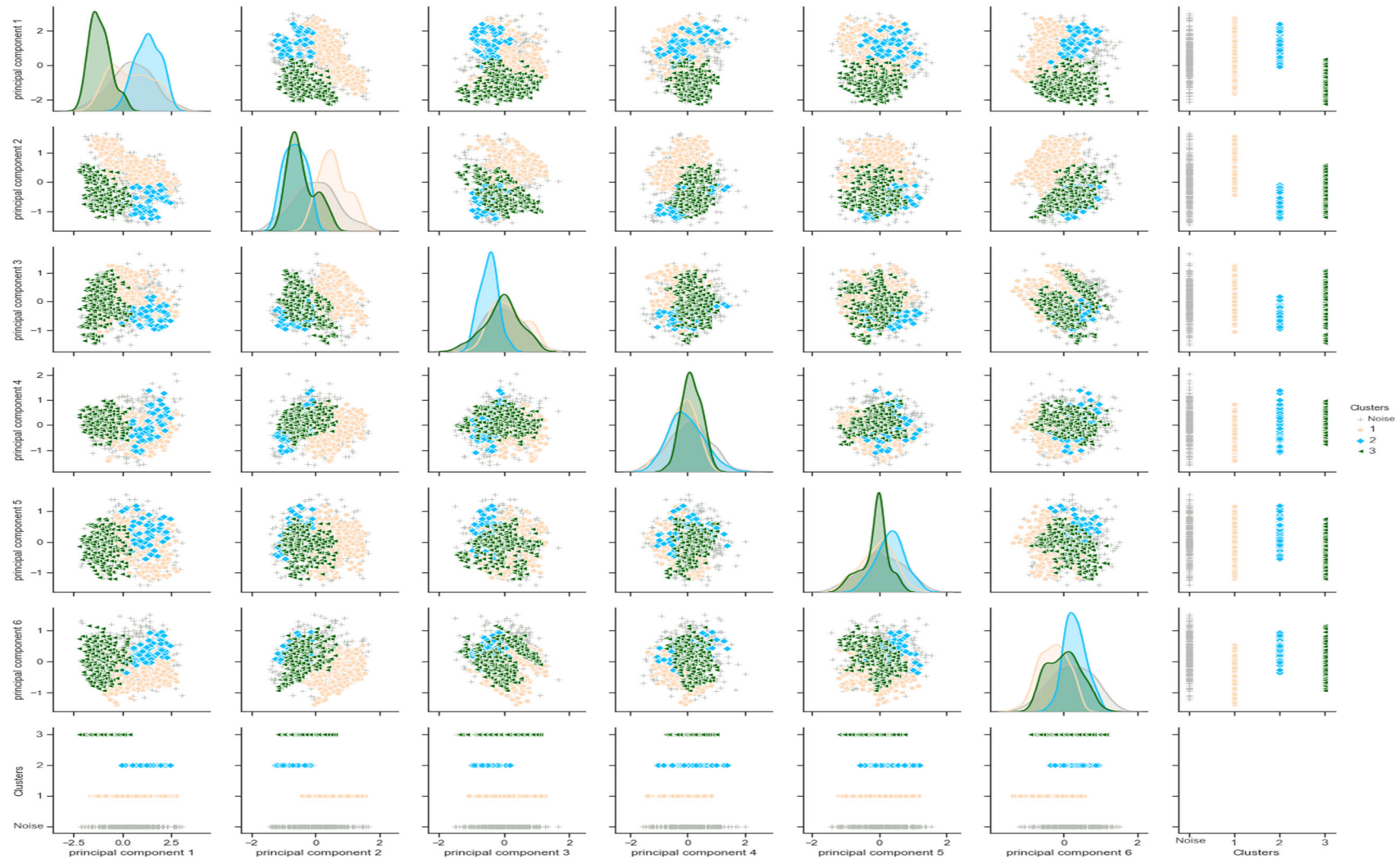
Figure 5. HDBSCAN results with three clusters (1, 2, and 3) in six dimensions. Clusters 1, 2, and 3 are represented by circles, diamonds, and triangles, respectively. Cluster 0 is noise. The distributions show a clear clustering, where data falls into three groups or types. The clouds present the density of clusters on each dimension.
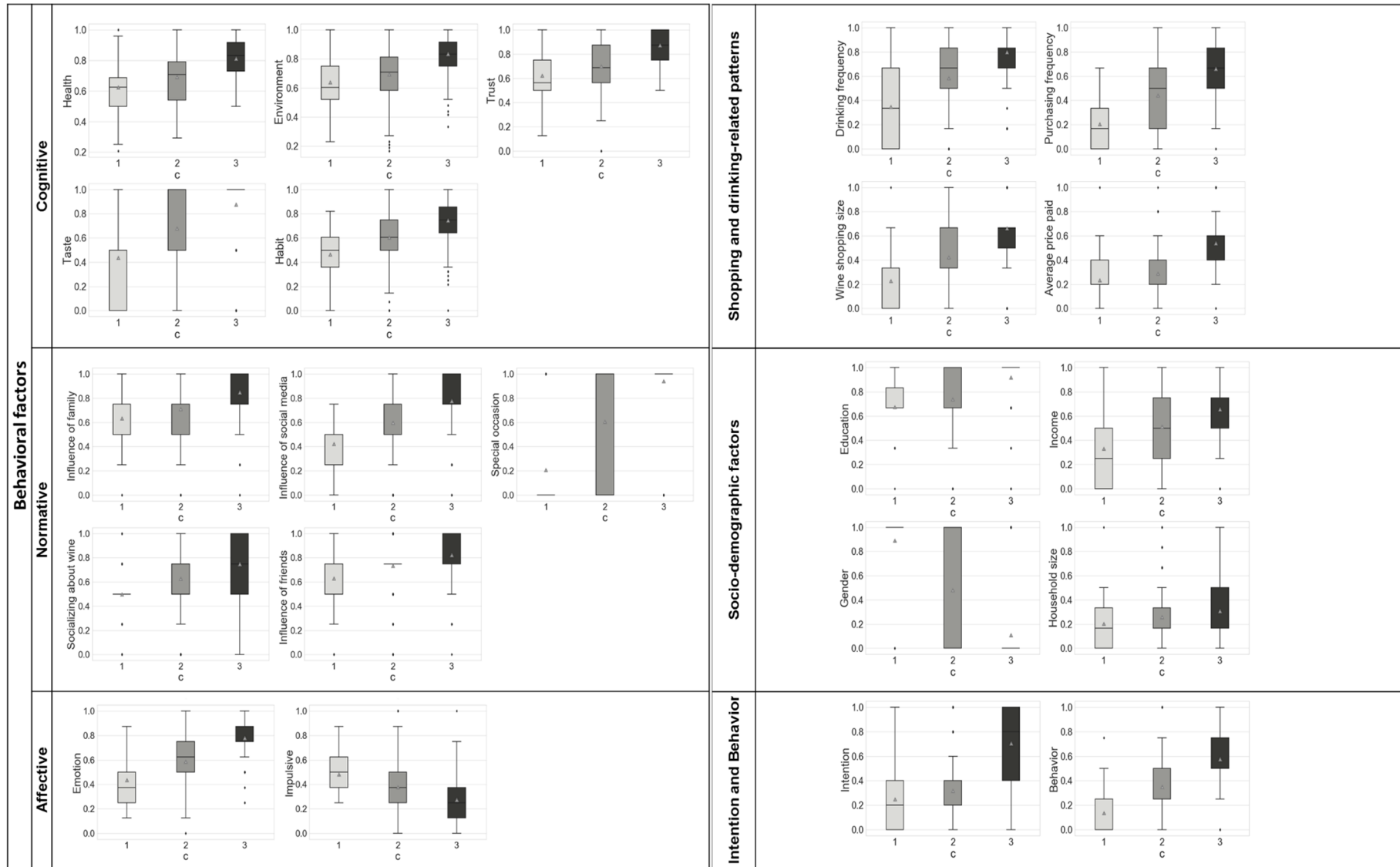
Figure 6. Variables according to which the three clusters (1, 2, and 3) are segregated. Special occasion (no=0, yes=1) and Gender are binary variables (male=0, female=1). The clusters are clearly different according to most of the variables, while there are some overlaps in others.

4.4.1. Non-organic segment: Impulsive behavior

Cluster 1, the non-organic segment, mainly represents conventional wine consumers. They report the lowest wine consumption and usually purchased items spontaneously. The gap between their higher intention (WTP 20% more) and lower organic purchasing behavior (organic wine purchasing less than 25%) is well explained by high impulsiveness, in line with IBT and affective events theory. The wine drinking and shopping frequency of this cluster were the lowest. Conventional wine consumers expressed negative feelings during shopping. They did not like the taste of organic over conventional wine or reported no distinction between the two, implying that hedonic goals were not activated. Yet, according to GFT, hedonic goal is one of the main drivers of behavior. Although they reported that health and environmental impacts are relatively important decision factors, they were less convinced that organic products have health and environmental benefits and did not trust them. Social norms influenced the wine purchasing decisions of these consumers very slightly. They stated that, in case of an increase in the price of their favorite wine, they would look for a cheaper substitute. In fact, they reported less loyalty to a certain brand of wine compared to other clusters. Regarding socio-demographic factors, consumers in this cluster were mainly poorly educated, lower-income women who had small-size households.

4.4.2. Occasional organic segment: Planned behavior

Cluster 2, the occasional organic segment, represents the bulk of the consumers with the highest potential for organic wine adoption. The intentions and behavior of these consumers were well aligned (WTP 10-20% more for organic and purchasing 25-50%

of wines organic), indicating planned wine purchasing behavior, which is aligned with TPB. For this cluster, the price was by far the main driver preventing organic wine purchasing decisions in this cluster: when the price of organic wine increases, they are unlikely to purchase it anymore (no substitution). Although the average price paid for wine in this cluster was similar to cluster 1, organic wine was mostly purchased for special occasions. In general, these consumers believed in the environmental and health benefits of organic wine consumption. Still, due to its high price, they only purchased it for celebrations or as a gift. Compared to conventional consumers, occasional organic consumers had relatively higher education, income, family size, brand loyalty, and interest in drinking organic wine and were less prone to impulsive wine shopping.

4.4.3. Organic segment: Unplanned behavior

Consumers in cluster 3, the organic segment, were mainly men with the highest education and highest income levels, living in big families. The average share of organic wine in their basket was more than 50%, higher than their reported intention (WTP varied between 20-50%). They based their choice primarily on normative goals and habits. On the one hand, the influence of family, friends, and other shoppers' choices on their wine purchasing decisions was the highest. They looked for more information about different wines from social media and sought the advice of others when selecting wine (in line with GFT). On the other hand, they were generally happy during shopping (in line with IBT) and tended to buy items habitually (in line with AT). Thus, the characteristics of this class are representative of unplanned wine purchasing behavior. Consumers in this cluster are strongly concerned with the health and environmental impacts of their food choices. Changes in the price of wine have a low

impact on their demand, and their average price acceptance is at a maximum. In other words, the price elasticity of this cluster is low, and if the prices of products increase, consumers will continue to purchase at higher prices.

5. Discussion

Our findings confirm the presence of planned, unplanned, and impulsive behaviors when shopping for wine. The following discussion of the results highlights several factors that can explain consumers' wine preferences.

Regarding the cognitive factors, RF models showed that trust adds substantially to the prediction of intentions (similar observation was made by Kim and Bonn (2015) as well). In line with D'Amico et al. (2016), the present study found that environmental consciousness and curiosity were associated with consumer WTP a premium for organic wines. When it comes to purchasing behavior, health attributes were found to be an important motivator for purchasing organic wine. This finding is consistent with the studies of Rana and Paul (2017) and Yadav (2016). Having said that, we found that consumers in cluster 2 mainly purchase conventional wine, despite their positive attitudes towards the health and environmental beliefs associated with organic products. Hence, we could not confirm that attitudes strongly predict behavior, as hypothesized by IBT and AT. Prior studies have reported contradictory results regarding the importance of taste on organic wine purchasing behavior (Mann et al. 2012). Nevertheless, our classification and clustering analyses were consistent with the study by Kim and Bonn (2015), in which American consumers reported taste as an important factor in their wine choice, favoring GFT when explaining consumers' behavior.

The influence of habits on more WTP for organic wine has not been sufficiently explained by the results of other studies in this context. Most wine-related studies define habits as the repetition of behavior as assessed by frequencies of shopping and drinking (e.g., Pomarici et al. (2016) and Vecchio (2013)), whereas, here, we considered habit as cognitively effortless and automatically initiated behavior, as assessed by the automaticity-specific index (Gardner et al. 2012). Our findings highlighted that while habitual purchasing as suggested by AT is important for promoting both organic wine purchasing intention and behavior, it could shed light on establishing stable shift towards organic consumption and its causality is not confirmed. Contrary to our expectations, habits did not override intention in directing behavior, and intentions remained significantly and equally predictive of behavior in all models: consumers choose wine mindfully rather than habitually. Gardner et al. (2015) referred to the temporal self-regulation theory to explain similar observations in terms of unhealthy snacking behavior, where strong self-control inhibits the habit. Similar to other organic wine studies (e.g., Pagliarini et al. (2013)), we found that WTP more for organic wine (intention) strongly influenced organic wine behavior. The classification results showed that, on average, consumers with higher WTP for organic wine also had a higher probability of buying it. However, our cluster analysis detected clusters of consumers with relatively higher intentions and lower behavior for organic wine (clusters 1 and 3). A similar gap between intention and behavior for organic wine has been described by Schäufele and Hamm (2018), who found that attitude and price are the barriers to organic wine adoption.

Regarding the normative factors, we found that normative support, as provided by social media and purchasing occasions, was relevant in determining consumers' organic wine purchasing behavior. This result highlights the influence of wine reviews

and recommendation systems on consumers' choices as suggested by GFT. It also highlights new potentials and opportunities for social media to assist businesses and industries to influence consumers' preferences. Szolnoki et al. (2018) and Dolan and Goodman (2017) both recently investigated the application of social media for promoting wine. Moreover, in line with the study of Boncinelli et al. (2019), in the current study, consumers valued organic wine more for special occasions rather than personal consumption. Concerning the clustering results, this statement stands true for 33% of consumers (occasional segment), while for the rest, it might not be the case, as occasions only partially influenced their wine purchasing decisions.

Regarding the affective factors, our findings demonstrate that happier, positive, and optimistic consumers are more likely to pay more for organic wine. Consistent with the study by Danner et al. (2016), positive and negative emotions were predictive of WTP more for organic wine. The influence of impulsive tendencies on organic wine purchasing decisions was more prevalent in the cluster analysis. On the one hand, consistent with IBT, impulsiveness caused by negative emotions may prompt spontaneous behavior that may, in turn, drive the consumer towards purchasing more conventional wine. On the other hand, unplanned decisions triggered by habits and normative cues may lead to higher organic purchasing if a consumer experiences positive emotions. Therefore, we can relate the effects of emotions to either impulsive or habitual behavior. Despite the importance of impulsiveness in predicting wine purchasing decisions, we only found one study, by Feldmann and Hamm (2015), that has highlighted the influence of spontaneous purchase situations.

Regarding the shopping and drinking-related patterns, the classification method indicated that the average price paid for wine was the strongest predictor and the

source of heterogeneity in the average behaviors of consumers. A higher price acceptance increases the likelihood that a consumer is more willing to pay a premium for organic wine. In the literature, the findings are mixed regarding the importance of price for buying wine (Huang et al. 2017); however, our results are in line with the studies of Schäufele and Hamm (2018) and Di Vita et al. (2019) who reported that, for the majority of consumers, price is the pivotal driver of wine choices. Another interesting result of the current study is that while consumers state they generally pay little attention to wine prices (about 70% of respondents), they actually base their organic wine purchasing decisions primarily on 'price'. While wine prices were considered to be the best predictor of organic wine purchasing behavior according to the RF model, the HDBSCAN model identified clusters that have equal average price acceptance, but the proportion of organic wine in their shopping baskets differed (refer to Figure 6, where organic wine in the shopping basket was less than 25% for cluster 1 and between 25-50% for cluster 2). The type of consumer behavior can explain this inconsistency in results; the wine purchasing decisions of cluster 2 consumers are more planned, whereas the decisions of cluster 1 consumers are more impulsive. The conventional segment consumers may change their preference for organic consumption if they experience positive emotions (like joy and contentment) during shopping and practice more planned buying as hypothesized by TIB. Interestingly, for the organic food segment, cluster 3, food price was the most important wine attribute, and that is why their high WTP more for organic wine (between 20% and 50%) cannot lead to full adoption of organic wine. The present findings seem to confirm GFT and support the findings of Janssen et al. (2020), where both conventional and organic food consumers reported that price was the most important attribute when making purchasing decisions.

Apart from the average price paid, variables such as the duration of shopping, average purchasing size, and the frequency of purchasing and drinking wine were found to be strong predictors of both intentions and behavior. It seems that consumers who spend a long time in the shop searching for products are likely to be willing to pay more for organic wine. Further, the more wines purchased per month, the higher the likelihood of intentions and behavior for purchasing organic wine. In line with previous studies, such as those by Pomarici and Vecchio (2014), higher frequencies of consuming and purchasing wine are related to a higher WTP more for organic wine.

Regarding socio-demographics, in agreement with the study by Zepeda and Deal (2009), the classification results indicated that socio-demographic factors have the lowest predictive power and are poor proxies for intention and behavior models. However, our clustering results revealed significant differences in income, education, household size, and gender between organic and conventional wine consumers.

6. Conclusions

Our findings have important implications for both theory and practice. From a theoretical perspective, they underscore the importance of considering impulsive and unplanned, as well as planned behavior, in understanding food purchasing. We argue that organic purchasing decisions result from an interplay between these factors, as explained by different social theories. Relying only on TPB and disregarding the presence of interruptive factors between intention and behavior means that we are unlikely to adequately capture the decision-making processes for organic food purchasing. TIB, AT, GFT, and IBT can explain the intention-behavior gap for different consumer segments when purchasing organic products.

6.1. Managerial implications

From a practical perspective, the classification results suggest that, for the average person, price is still an obstacle to purchasing organic food. The clustering results provide strong evidence of the influence of impulsive, habitual, and normative cues as well as the dual role of emotions in choosing organic products in three distinct consumer segments. In fact, we would have highlighted these two factors (trust and price) as the most important attributes in wine purchasing if we had only used classification algorithms.

Sales promotions and government subsidies for organic products can support organic purchasing and, at the same time, change consumer consumption habits to help the environment. Retailers can have an organic section in their stores specifically designed to facilitate this behavior. Encouraging a greater sense of joy and happiness in the store, and using social media to advertise a range of organic products, may be other effective mechanisms to change wine purchasing behavior. We may be ignoring the influence of affective factors if we rely only on the results of the classification analysis. Future research would benefit from examining the efficacy of these interventions in shifting behavior towards organic consumption.

6.2. Limitations and future directions

This study has a number of limitations that suggest several potential directions for future research. Firstly, the reliance on self-reported behavior rather than conducting observational experiments is a limitation. Survey respondents are prone to social desirability bias in reporting their intention for organic products, and their behavior can only be interpreted as a reported preference; it is not their real behavior. Thus, the

findings of this study are based on stated preferences and are experimental in nature. One possible future direction is to compliment survey data with real market transactions that provide realistic, robust results. Another limitation of this study is relevant to the geographical constraints of the sample and the generalizability of the results. Our data were collected from one region of a major city in Australia and there is a possibility that the results are more closely aligned to the perspectives of these particular residents. Thus, the results cannot be generalized to the entire national population. Future research may choose to broaden the participant recruitment process or conduct a comparative study on the differences and similarities between organic wine preferences of Australians across different states and other populations. Finally, the impact of packaging, region of origin, grape variety, and other extrinsic characteristics on organic wine purchasing can be explored in future research.

This said the article provides important contributions to the literature. Using explainable AI techniques, we advance the methodological principles of empirical research on retailing and consumer behavior. Besides, our study provides new insights into the role of emotions and norms in decisions making process for food with sustainability characteristics. Further, the presented ML algorithms can be used to inform the extended supply chain framework (Taghikhah et al. 2019) to predict consumer motivation and behavior for green products across a variety of segments. The results also allow us to further calibrate and test the agent-based model, ORVin (Taghikhah et al. 2020a), developed to quantify the cumulative impacts of organic wine preference changes among heterogeneous consumers prone to behavioral biases and social interactions.

Acknowledgments

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes, 50*(2), 179-211. doi:https://doi.org/10.1016/0749-5978(91)90020-T

Boncinelli, F., Dominici, A., Gerini, F., & Marone, E. (2019). Consumers wine preferences according to purchase occasion: Personal consumption and gift-giving. *Food Quality and Preference, 71*, 270-278. doi:https://doi.org/10.1016/j.foodqual.2018.07.013

Campello, R. J., Moulavi, D., & Sander, J. (2013). *Density-based clustering based on hierarchical density estimates.* Paper presented at the Pacific-Asia conference on knowledge discovery and data mining.

Chen, Y., Tang, S., Bouguila, N., Wang, C., Du, J., & Li, H. (2018). A fast clustering algorithm based on pruning unnecessary distance computations in DBSCAN for high-dimensional data. *Pattern Recognition, 83*, 375-387. doi:https://doi.org/10.1016/j.patcog.2018.05.030

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science, 1*(3), 98-101.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*: Academic press.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning, 20*(3), 273-297.

D'Amico, M., Di Vita, G., & Monaco, L. (2016). Exploring environmental consciousness and consumer preferences for organic wines without sulfites. *Journal of cleaner production, 120*, 64-71. doi:https://doi.org/10.1016/j.jclepro.2016.02.014

Danner, L., Ristic, R., Johnson, T. E., Meiselman, H. L., Hoek, A. C., Jeffery, D. W., & Bastian, S. E. (2016). Context and wine quality effects on consumers' mood, emotions, liking and willingness to pay for Australian Shiraz wines. *Food Research International, 89*, 254-265. doi:https://doi.org/10.1016/j.foodres.2016.08.006

Di Vita, G., Pappalardo, G., Chinnici, G., La Via, G., & D'Amico, M. (2019). Not everything has been still explored: Further thoughts on additional price for the organic wine. *Journal of cleaner production, 231*, 520-528. doi:https://doi.org/10.1016/j.jclepro.2019.05.268

Dolan, R., & Goodman, S. (2017). Succeeding on social media: Exploring communication strategies for wine marketing. *Journal of Hospitality and Tourism Management, 33*, 23-30.

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). *A density-based algorithm for discovering clusters in large spatial databases with noise.* Paper presented at the Kdd.

Feldmann, C., & Hamm, U. (2015). Consumers' perceptions and preferences for local food: A review. *Food quality and preference, 40*, 152-164. doi:https://doi.org/10.1016/j.foodqual.2014.09.014

Gardner, B., Abraham, C., Lally, P., & de Bruijn, G.-J. (2012). Towards parsimony in habit measurement: Testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. *International Journal of Behavioral Nutrition and Physical Activity, 9*(1), 102.

Gardner, B., Corbridge, S., & McGowan, L. (2015). Do habits always override intentions? Pitting unhealthy snacking habits against snack-avoidance intentions. *BMC psychology, 3*(1), 8. doi:10.1186/s40359-015-0065-4

Gentile, F., La Torre, G. L., Potortì, A. G., Saitta, M., Alfa, M., & Dugo, G. (2016). Organic wine safety: UPLC-FLD determination of Ochratoxin A in Southern Italy wines from organic farming and winemaking. *Food control, 59*, 20-26.

Helmefalk, M., & Hultén, B. (2017). Multi-sensory congruent cues in designing retail store atmosphere: Effects on shoppers' emotions and purchase behavior. *Journal of Retailing and Consumer Services, 38*, 1-11. doi:https://doi.org/10.1016/j.jretconser.2017.04.007

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence, 20*(8), 832-844.

Huang, A., Dawes, J., Lockshin, L., & Greenacre, L. (2017). Consumer response to price changes in higher-priced brands. *Journal of Retailing and Consumer Services, 39*, 1-10.

Janssen, M., Schäufele, I., & Zander, K. (2020). Target groups for organic wine: The importance of segmentation analysis. *Food quality and preference, 79*, 103785.

Kim, H., & Bonn, M. A. (2015). The moderating effects of overall and organic wine knowledge on consumer behavioral intention. *Scandinavian Journal of Hospitality and Tourism, 15*(3), 295-310. doi:https://doi.org/10.1080/15022250.2015.1007083

Lindenberg, S., & Steg, L. (2007). Normative, gain and hedonic goal frames guiding environmental behavior. *Journal of social issues, 63*(1), 117-137. doi:http://10.1111/j.1540-4560.2007.00499.x

Loose, S. M., & Lockshin, L. (2013). Testing the robustness of best worst scaling for cross-national segmentation with different numbers of choice sets. *Food quality and preference, 27*(2), 230-242. doi:https://doi.org/10.1016/j.foodqual.2012.02.002

Mann, S., Ferjani, A., & Reissig, L. (2012). What matters to consumers of organic wine? *British Food Journal, 114*(2), 272-284. doi:https://doi.org/10.1108/00070701211202430

Maurus, S., & Plant, C. (2016). *Skinny-dip: Clustering in a Sea of Noise.* Paper presented at the Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining.

Ogbeide, O. A. (2013). *Consumer willingness to pay premiums for the benefits of organic wine and the expert service of wine retailers.*

Pagliarini, E., Laureati, M., & Gaeta, D. (2013). Sensory descriptors, hedonic perception and consumer's attitudes to Sangiovese red wine deriving from organically and conventionally grown grapes. *Frontiers in psychology, 4*, 896. doi: https://doi.org/10.3389/fpsyg.2013.00896

Panzone, L. A. (2014). Why are discounted prices presented with full prices? The role of external price information on consumers' likelihood to purchase. *Food Quality and Preference, 31*, 69-80. doi:https://doi.org/10.1016/j.foodqual.2013.08.003

Pomarici, E., Amato, M., & Vecchio, R. (2016). Environmental friendly wines: a consumer segmentation study. *Agriculture and agricultural science procedia, 8*(2016), 534-541.

Pomarici, E., & Vecchio, R. (2014). Millennial generation attitudes to sustainable wine: an exploratory study on Italian consumers. *Journal of cleaner production, 66*, 537-545. doi:https://doi.org/10.1016/j.jclepro.2013.10.058

Quinlan, J. R. (1990). Probabilistic decision trees. In *Machine learning* (pp. 140-152): Elsevier.

Rahmah, N., & Sitanggang, I. S. (2016). *Determination of optimal epsilon (eps) value on dbscan algorithm to clustering data on peatland hotspots in sumatra.* Paper presented at the IOP Conference Series: Earth and Environmental Science.

Rana, J., & Paul, J. (2017). Consumer behavior and purchase intention for organic food: A review and research agenda. *Journal of Retailing and Consumer Services, 38*, 157-165.

Rugani, B., Vázquez-Rowe, I., Benedetto, G., & Benetto, E. (2013). A comprehensive review of carbon footprint analysis as an extended environmental indicator in the wine sector. *Journal of cleaner production, 54*, 61-77. doi:https://doi.org/10.1016/j.jclepro.2013.04.036

Russell, S. V., Young, C. W., Unsworth, K. L., & Robinson, C. (2017). Bringing habits and emotions into food waste behaviour. *Resources, Conservation and Recycling, 125*, 107-114. doi:https://doi.org/10.1016/j.resconrec.2017.06.007

Ryan, J., & Casidy, R. (2018). The role of brand reputation in organic food consumption: A behavioral reasoning perspective. *Journal of Retailing and Consumer Services, 41*, 239-247.

Schäufele, I., & Hamm, U. (2017). Consumers' perceptions, preferences and willingness-to-pay for wine with sustainability characteristics: A review. *Journal of cleaner production, 147*, 379-394. doi:https://doi.org/10.1016/j.jclepro.2017.01.118

Schäufele, I., & Hamm, U. (2018). Organic wine purchase behaviour in Germany: Exploring the attitude-behaviour-gap with data from a household panel. *Food Quality and Preference, 63*, 1-11. doi:https://doi.org/10.1016/j.foodqual.2017.07.010

Stern, H. (1962). The significance of impulse buying today. *Journal of marketing, 26*(2), 59-62. doi:https://doi.org/10.1177/002224296202600212

Stolz, H., & Schmid, O. (2008). Consumer attitudes and expectations of organic wine. http://www.isofar.org/modena2008/prceedings.html

Szolnoki, G., Dolan, R., Forbes, S., Thach, L., & Goodman, S. (2018). Using social media for consumer interaction: An international comparison of winery adoption and activity. *Wine Economics and Policy, 7*(2), 109-119.

Taghikhah, F., Voinov, A., & Shukla, N. (2019). Extending the supply chain to address sustainability. *Journal of cleaner production*(229), 652-666.

Taghikhah, F., Voinov, A., Shukla, N., & Filatova, T. (2020a). Exploring consumer behavior and policy options in organic food adoption: Insights from the Australian wine sector. *Environmental science & policy, 109*, 116-124.

Taghikhah, F., Voinov, A., Shukla, N., Filatova, T., & Anufriev, M. (2020b). Integrated modeling of extended agro-food supply chains: A systems approach. *European Journal of Operation Research*(Accepted).

Triandis, H. C. (1977). *Interpersonal behavior*: Brooks/Cole Pub. Co.

Trinh, G., Corsi, A., & Lockshin, L. (2019). How country of origins of food products compete and grow. *Journal of Retailing and Consumer Services, 49*, 231-241.

Vecchio, R. (2013). Determinants of willingness-to-pay for sustainable wine: Evidence from experimental auctions. *Wine Economics and Policy, 2*(2), 85-92. doi:https://doi.org/10.1016/j.wep.2013.11.002

Verplanken, B., & Orbell, S. (2003). Reflections on past behavior: a self-report index of habit strength 1. *Journal of applied social psychology, 33*(6), 1313-1330. doi:https://doi.org/10.1111/j.1559-1816.2003.tb01951.x

Vrček, I. V., Bojić, M., Žuntar, I., Mendaš, G., & Medić-Šarić, M. (2011). Phenol content, antioxidant activity and metal composition of Croatian wines deriving from organically and conventionally grown grapes. *Food Chemistry, 124*(1), 354-361.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology, 54*(6), 1063.

Yadav, R. (2016). Altruistic or egoistic: Which value promotes organic food consumption among young consumers? A study in the context of a developing nation. *Journal of Retailing and Consumer Services, 33*, 92-97.

Yang, Y., & Paladino, A. (2015). The case of wine: understanding Chinese gift-giving behavior. *Marketing Letters, 26*(3), 335-361. doi:https://doi.org/10.1007/s11002-015-9355-0

Zepeda, L., & Deal, D. (2009). Organic and local food consumer behaviour: Alphabet theory.

International Journal of Consumer Studies, 33(6), 697-705.

doi:https://doi.org/10.1111/j.1470-6431.2009.00814.x