ARTICLE

Have consumers escaped from COVID-19 restrictions by seeking variety? A Machine Learning approach analyzing wine purchase behavior in the United States

Wolfram Rinke and Shuay-Tsyr Ho

1Department of Information-Technology and Information-Management, Fachhochschule Burgenland GmbH, Eisenstadt, Austria and 2Department of Agricultural Economics, National Taiwan University, Taipei, Taiwan

Corresponding author: Shuay-Tsyr Ho, email: shuaytsyrho@ntu.edu.tw

Abstract
The COVID-19 pandemic itself constitutes an environment for people to experience the potential loss of control and freedom due to social distancing measures and other government orders. Variety-seeking has been treated as a mechanism to regain a sense of self-control. Using Machine Learning model and household-level data with a focus on the wine market in the United States, this study showcases the changing variety-seeking behavior over the pandemic year of 2020, in which people’s perception of the status of restriction measures influences the degree of their use of variety-seeking behavior as a coping strategy. It is the shopping pattern and store environments that drive the behavioral responses in wine purchases to freedom-limited circumstances. Coupon use is associated with a lower variety-seeking tendency at the beginning of the stay-at-home order, but the variety level resumes when more time has passed in the restriction periods. Variety-seeking tendency increases with shopping frequency at the beginning of the social distancing measure but decreases to a level lower than all the non-restriction periods.

Keywords: COVID-19; data mining; Machine learning; purchase behavior; variety-seeking; wine

JEL classifications: C33; D12; D91

I. Introduction
The COVID-19 pandemic shaped the environment and atmosphere for individuals to feel both spatially and psychologically stuck, which could lead to a higher degree of variety-seeking behavior among consumers feeling a sense of threat (Kim, 2020). Variety-seeking behavior has been found to be used by consumers as a coping mechanism when they feel a low sense of personal control (Levav and Zhu, 2009).
In this research, we treat the COVID-19 pandemic as a threat-inducing context to examine how people's purchase behaviors respond to the distress and isolation brought about by the social distancing measures. Consumer research and the marketing field have looked into the relationship between variety-seeking and self-control or power using data from lab experiments and online surveys (Levav and Zhu, 2009; Yoon and Kim, 2018; Wang, Raghunathan, and Gauri, 2022). Empirical insights on how purchase behaviors are influenced by the global pandemic are also based on survey data (Kim, 2020; Gu et al., 2021). The purpose of this study is to expand the breadth and depth of the inquiry on variety-seeking behavior and control restoration using a detailed household-level transaction-level dataset.

COVID-19 pandemic shocked the food and beverage markets around the world (Chowdhury et al., 2022; Wittwer and Anderson, 2021). While the alcohol market remains resilient owing to the shift in purchase behaviors across different retail channels, consumption is optimistically rebalanced (NielsenIQ, 2020; Dubois et al., 2021). Food service and drinking place sales are replaced by beer, wine, and liquor sales (Castaldelli-Maia et al., 2021). The lockdown of the hospitality sector during the COVID-19 pandemic increased alcohol sales in off-premise stores and e-commerce channels. Alcohol sales through e-commerce in Australia, Brazil, China, France, Germany, Italy, Japan, Spain, the United Kingdom, and the United States increased by more than 40% (OECD, 2021). At the same time, the fluctuating monthly sales data of beer, wine, and spirits in the United States depicts a different story on the purchase patterns across states during the pandemic, indicating that the resilience of the alcohol market could be susceptible to the condition of the pandemic and different statuses of government measures (NIAAA, 2020). Anxiety during the COVID-19 pandemic could also play an important role in driving alcohol consumption among both drinking professionals and amateurs (Agnoli and Charters, 2022). Variety-seeking tendencies between alcoholic beverages could involve substitutability or complementarity among other beverages (Fogarty, 2010). This study chooses to focus on the wine market in the United States to derive the variety-seeking tendency, which entails the preference for more disaggregated items under one product category (Dodd, Pinkleton, and Gustafson, 1996; Caracciolo et al., 2022). This research aims to examine how U.S. wine consumers might have (or have not) used varied choices to cope with limited freedom under COVID-19 restrictions and how their wine shopping behaviors could have changed during the pandemic.

II. Theoretical background

Variety-seeking tendency characterizes consumers’ behavioral patterns seeking diversity in goods and services, where potential motivations encompass external situations, stimulation, and future uncertainty (Kahn, 1995). In addition, variety-seeking behavior can be used as a coping strategy to tackle stress, perceived threats, or loss of control (Kim, 2020). Choice is a representation, in Western society, of the ability to express oneself, control the environment, and determine the boundary of freedom (Stephens, Markus, and Townsend, 2007; Kim and Drolet, 2003). It is the perceived threats to lives and freedom that evoke certain behaviors related to choice selection, attempting to regain a sense of control and freedom. Events like the COVID-19 pandemic could
threaten people's self-esteem, and their freedom and control of normal life could be significantly impaired.

Regarding the underlying psychological mechanism explaining choice behaviors under control-threatening events, reactance is characterized by how people reassert their freedom and feeling of control through certain behaviors when such freedom is threatened (Brehm, 1966). Variety-seeking behaviors have been found to restore such senses when life and activity are limited. People show a higher tendency to diversify their choices in food and stationary items when they are provided with information regarding the trajectory of COVID-19 pandemic development as well as the degree of life-threatening risk of such a pandemic they self-evaluate (Kim, 2020). People may diversify their choices to “win back” their sense of personal control that is eroded by their perception of low economic mobility due to their socioeconomic status objectively judged by society (Yoon and Kim, 2018). Spatial confinement exemplified by the store and shopping environment can also influence customers’ responses to both varietal and unique purchase decisions (Levav and Zhu, 2009). Narrow aisles lead to not only more variety-seeking patterns but also more uniqueness-seeking patterns. In sum, the existing evidence showcasing the variety-seeking behavior among consumers under certain circumstances is supported by subjective experiment-related data, while the household-level trip-level data, albeit self-reported, characterizes the more actual decision-making process of consumers that the purchases are tracked in a longer time frame and preferences might have exhibited more systematic patterns.

III. Data and method

The Consumer Panel Dataset provided by NielsenIQ used in this research is a longitudinal dataset comprised of a cross-sectional panel of households that constantly provide transaction information on their grocery purchases using an in-home scanner (NielsenIQ, 2023). The dataset includes detailed information on household shopping and purchase behavior, including shopping trips, quantity and price of purchased items, product characteristics, and in-store promotions. The variety-seeking behavior is characterized by the count of different products purchased to examine how people diversify their alcohol purchases at different points in time related to social distancing measures during the COVID-19 pandemic, particularly in 2020. Two social distancing periods are considered: the closure of business and public service from March 10th to July 14th, 2020, and the limited public gathering from October 8th to November 3rd, 2020 (ACAPS, 2022). Spanning across these two windows, three periods as unlocked, prelocked (14 days before lock), and locked are classified interchangeably to showcase the time-evolving pattern of purchase behavior.

This study focuses on the states with the highest level of population, California and Florida. We chose the social distancing measure at the national level and in the selected states for two reasons. First, the use of federal orders rather than state-level government measures focuses on the overall impact of social distancing implementations on individual behavior (Hale et al., 2021). Consideration of different measures at the state level needs to be associated with state-level alcohol regulations and taxation laws to investigate how shopping behavior and variety-seeking tendencies are influenced at the localized level (Ho and Rickard, 2021). Second, we chose the most populous states
as the research focus rather than states with the highest level of alcohol use to examine the general impact of social distancing measures on wine purchase behavior, excluding the economic and societal factors driving the systemic drinking pattern and laws regulating different types of alcoholic beverages in specific states (Smith, 1982).

We empirically document how consumers’ purchase behaviors vary by the length of time during the restriction period and by the days ahead of and after lifting the social distancing measure. Holidays, in particular, are filtered out to isolate the impact of restriction periods on purchase patterns. Variety-seeking behavior is proxied by the number of counts of different disaggregated items purchased at the Universal Product Code (UPC) level, at the daily level, and at the store level. The varietal pattern is much less obvious at the household level with wide ranges of zero values, while the store-level variety-seeking pattern entails a greater level of variability in household purchases. Also, related to sampling technique, the number of households sampled at the county level is on average less than five households, such that the variety-seeking tendency at the store level could reflect the varietal preference at the locality level.

We use a Machine Learning regression model to describe household shopping behavior exemplified by the variety-seeking tendency to examine the aforementioned theory empirically. The choice between Machine Learning and traditional statistical models goes beyond whether the method hinges upon a more rigorous structure of statistical assumptions (Athey and Imbens, 2019). The neural network model mathematically and computationally mimics the decision process of the human brain and adopts the algorithm properties to train the data for which the mechanism is unknown or unassumed. The neural network model can be a more accurate tool for predicting consumer choices since it is not constrained by linear estimation procedures and compensatory decision rules (West, Brockett, and Golden, 1997). A classic feed-forward artificial neural network (ANN) perceptron model is used in this study (Niklas and Rinke, 2020; Rumelhart, Hinton, and Williams, 1986). The applied architecture is a multi-layer perceptron model with a sigmoidal transfer function in the first and second hidden layers and a linear transfer function in the output layer for scaling the output variable. Empirically, the ANN model performs better with an R2 of 0.884 than a standard linear regression model with an R2 of 0.794.

The model contains two categories of independent variables. One group describing the shopping trips taking place in certain stores include the number of trips on a per-day, per-store basis, the number of bottles of wine purchased, the number of brands, the total expenditure on wine purchases, and the value of coupons used on wine purchases. The other set of variables describe the restrictions and include an indicator value between 0.0 and 1.0 expressing how close we are to the implementation of stay-at-home orders (PL-indicator), the number of days during the stay-at-home order period (DIL-counter), the number of days till the end of stay-at-home orders (DEL-counter), and the number of mobility-restricted orders passed (L-counter). The dependent variable is the number of different UPCs bought.

IV. Results

The results are divided into two parts. Given that independent and dependent variables are correlated, behavioral insights from Principal Component Analysis (PCA)
Table 1. Significance of variables in explaining the variation of variety measure from PCA analysis

<table>
<thead>
<tr>
<th>Components</th>
<th>% Variance</th>
<th>Brand</th>
<th>Volume</th>
<th>Spending</th>
<th>Coupon</th>
<th>Trip</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>0.304</td>
<td>0.492</td>
<td>0.467</td>
<td>0.508</td>
<td>0.225</td>
<td>0.47</td>
</tr>
<tr>
<td>PC2</td>
<td>0.246</td>
<td>0.059</td>
<td>0.042</td>
<td>0.028</td>
<td>0.069</td>
<td>0.04</td>
</tr>
<tr>
<td>PC3</td>
<td>0.113</td>
<td>0.005</td>
<td>−0.056</td>
<td>0</td>
<td>0.174</td>
<td>0.021</td>
</tr>
<tr>
<td>PC4</td>
<td>0.102</td>
<td>−0.462</td>
<td>0.412</td>
<td>0.309</td>
<td>0.508</td>
<td>−0.509</td>
</tr>
<tr>
<td>PC5</td>
<td>0.099</td>
<td>0.108</td>
<td>−0.399</td>
<td>−0.313</td>
<td>0.808</td>
<td>0.214</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Components</th>
<th>PL-indicator</th>
<th>DIL-counter</th>
<th>DEL-counter</th>
<th>L-counter</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>0.066</td>
<td>0.065</td>
<td>0.036</td>
<td>0.027</td>
</tr>
<tr>
<td>PC2</td>
<td>−0.645</td>
<td>−0.527</td>
<td>−0.542</td>
<td>−0.011</td>
</tr>
<tr>
<td>PC3</td>
<td>−0.074</td>
<td>0.026</td>
<td>0.104</td>
<td>−0.974</td>
</tr>
<tr>
<td>PC4</td>
<td>0.021</td>
<td>−0.006</td>
<td>0.005</td>
<td>0.053</td>
</tr>
<tr>
<td>PC5</td>
<td>0.047</td>
<td>0.045</td>
<td>−0.019</td>
<td>0.168</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations.

are presented before the main results from the dependency matrix and the sensitivity analyses based on the partial derivatives derived from the ANN model (Rinke, 2015). Principle Component Analysis transforms the input data into linear, independent variables. It reduces the dimensionality of the model and gives a more distinct picture of the shopping behaviors separated at different phases of time. Table 1 presents the results from the PCA method, leading to five principal components and explaining 86% of the data variance. The main factors are marked in bold.

The component PC1 explains more than 30% of the variability and combines the shopping day variables, including brand, volume, spending, and number of trips; PC2 is mostly explained by the variables related to the social distancing measures; PC3 is derived from the variable of coupon values and the number of restriction orders; PC4 is derived from all the variables describing the shopping traits; PC5 is derived from the variables of quantity purchased, total wine expenditures per trip, and mainly coupon value.

Focusing on the two components that explain the most variation in the variety-seeking tendency, Figure 1 shows a scatterplot comparing PC1 and PC2. The scatterplot shows that purchase patterns distinguish themselves across different time frames involving restriction measures.

Stay-at-home orders at different points in time have differential impacts on purchase behavior. Consumers exhibit a higher degree of variety-seeking during the first stay-at-home order (marked in light green) compared to the degree during the second stay-at-home order (marked in pink), indicating a learning curve among households in dealing with the uncertainty and limited freedom brought about by the restriction orders. Also, the variety-seeking tendency is instead higher in the non-restriction period, which indicates that variety-seeking may not be regarded as a coping mechanism to restore self-control under the social distancing order. However, this line of results should not be interpreted as contradicting the aforementioned theory since further mechanisms determining the degree of variety-seeking behavior have not been established, which will be examined using the ANN model in the subsequent section.
Figure 1. Comparison of components PC1 and PC2 that explain the most of variation in the variety-seeking variable.
Source: Authors’ calculations.

Table 2 shows the dependency analysis from the ANN regression model (Rinke, 2015), which indicates the level of explaining power of key independent variables on the variety-seeking measure. Variety-seeking tendency has the strongest dependency associated with the variable “DIL-counter,” with a value of 1.0 and an average elasticity of –0.330. The variable “brand” has the second-strongest dependency association of 0.807 with an average elasticity of 0.471. The variable “DEL-counter” also shows strong dependency with variety-seeking measure with a value of 0.804 and an average elasticity of 0.193, which is consistent and expected as both counters are inverse linear dependent. The “DIL-counter” variable shows different directions in locked, pre-locked, and unlocked phases, which, on the other hand, supports the theory. A switch from unlocked to locked (a value from zero days in lockdown to one day in lockdown) immediately shows a strong tendency to reduce variety seeking level with an elasticity of –0.660 and –0.731, respectively. But as the lockdown continues, the elasticity of 0.190 shows that the variety-seeking trend is increasing.

Figures 2 and 3 present the evolving purchase pattern on a daily basis in the year 2020. They show how the partial derivative of the selected variable associated with the variety-seeking level changes over time, as calculated from the ANN regression model. Figure 2 shows that the partial derivative of coupon value in wine purchases further reduces the variety-seeking behavior as the nation enters into a stay-at-home order period at the beginning and the pattern instead reverses afterward.

Two stay-at-home orders show a similar pattern. The positive association between variety-seeking and coupon value implies that, as people have been immersed in the seemingly lockdown environment longer, the coupons would encourage individuals to buy different varietals and may have wider targeted products. At the store level, the reduced impact of coupons on the variety-seeking tendency at the beginning of
Table 2. ANN-based dependency analysis

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Dependency</th>
<th>Avg. elast.</th>
<th>Abs. avg. elast.</th>
<th>Locked</th>
<th>Prelocked</th>
<th>Unlocked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand</td>
<td>0.807</td>
<td>0.471</td>
<td>0.471</td>
<td>0.490</td>
<td>0.655</td>
<td>0.422</td>
</tr>
<tr>
<td>Volume</td>
<td>0.613</td>
<td>0.314</td>
<td>0.314</td>
<td>0.286</td>
<td>0.244</td>
<td>0.351</td>
</tr>
<tr>
<td>Spending</td>
<td>0.284</td>
<td>-0.029</td>
<td>0.029</td>
<td>-0.038</td>
<td>-0.069</td>
<td>-0.013</td>
</tr>
<tr>
<td>Coupon</td>
<td>0.251</td>
<td>-0.044</td>
<td>0.044</td>
<td>-0.088</td>
<td>-0.055</td>
<td>-0.004</td>
</tr>
<tr>
<td>Trip</td>
<td>0.514</td>
<td>0.174</td>
<td>0.174</td>
<td>0.172</td>
<td>0.057</td>
<td>0.196</td>
</tr>
<tr>
<td>PL-indicator</td>
<td>0.652</td>
<td>0.014</td>
<td>0.014</td>
<td>0.243</td>
<td>-0.016</td>
<td>-0.184</td>
</tr>
<tr>
<td>DIL-counter</td>
<td>1.000</td>
<td>-0.330</td>
<td>0.330</td>
<td>0.190</td>
<td>-0.660</td>
<td>-0.731</td>
</tr>
<tr>
<td>DEL-counter</td>
<td>0.804</td>
<td>0.193</td>
<td>0.193</td>
<td>0.201</td>
<td>0.226</td>
<td>0.181</td>
</tr>
<tr>
<td>L-counter</td>
<td>0.525</td>
<td>-0.104</td>
<td>0.104</td>
<td>-0.127</td>
<td>-0.050</td>
<td>-0.093</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations.

Figure 2. Time-evolving pattern of variety-seeking behavior associated with coupon value. Source: Authors’ calculations.

the first restriction period suggests that hoarding behavior associated with the same varietal (Baddeley, 2020) might hold for the purchase of alcoholic beverages during the pandemic. Overall, the variety-seeking tendency associated with coupon value remains quite consistent throughout the turbulent pandemic year in 2020, implying that variety-seeking has not been fully treated as a tackling strategy to regain self-control under the restriction period, particularly regarding how store promotion could influence people’s shopping behavior. An increase in consumer patronage and an increase in the number of visits to stores enhance the degree of variety-seeking behavior compared to the prelocked and unlocked periods, as presented in Figure 3.
However, the variety-seeking tendency drops as time moves further into the social distancing period, while the reversed trend does not appear during the second stay-at-home order period. The fluctuation in variety-seeking tendency associated with the visits to a store per day across various phases of social distancing measures implies that the coping mechanism changes as time enters further into the restriction period and as it approaches the end of the lockdown. Overall, there is an obvious implication that consumers may seek a greater level of variety during the restriction order period in response to the disruption of normal life. Also, in times between the first and second stay-at-home orders and even after the second federal order, the variety-seeking tendency has been pushed up to the new normal, surpassing the original purchase behavior before the very first stay-at-home order, implying the long-lasting influence of freedom-limited measures on consumers’ behavioral strategies to respond to and cope with.

V. Conclusion

The COVID-19 pandemic influences consumer purchase behavior associated with different degrees of restriction periods. Variety-seeking behavior itself evolves when people are in different stages of the implementation and lifting of stay-at-home orders. The findings of this study both elucidate and challenge the theory proposing that people seek more variety to cope with the potential loss of self-control in a freedom-restricted environment, as evidenced by studies using more subjective and context-specific data. Shopping characteristics contributed to mixed impacts on variety-seeking behavior.
They moderated the channel of how variety-seeking behavior could be used to restore self-control under social distancing measures during the pandemic.

The implications of this study are twofold. The Artificial Neural Network model used here provides an extended perspective beyond point estimates on how variety-seeking tendency as a behavioral response can be interpreted beyond the direct association with external shock related to control- or freedom-threatening events. Our findings shed light on the role of shopping or household characteristics that lead to variety-seeking behavior under specific circumstances, which can be applied to investigate dynamic consumer choices in the long run. Also, examining variety-seeking behavior in wine purchases has potential marketing implications. The more crowded environment or holiday season may have a higher likelihood of motivating variety-seeking with a wider range of products. Retailers may benefit from diversifying their inventory. The frequent wine shoppers are more aligned with psychological reactance behavior when feeling a loss of control, while the pattern is not consistent throughout the shock. Criteria for customer segmentation beyond socio-economic characteristics are nonnegligible for new product introductions and customer retention strategies in the wine market.

Acknowledgments. We are extremely grateful to the editor and anonymous referees for their valuable comments and suggestions, which have helped improve the quality of the paper. We are also indebted to the participants at the 15th Annual AAWE Conference in Stellenbousch for their many helpful comments.

Our analyses were calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the authors and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

References


NielsenIQ (2023). Consumer panel dataset, 2020. The James M. Kilts Center for Marketing, the University of Chicago Booth School of Business, Chicago, IL.


