Environment, Development and Sustainability
https://doi.org/10.1007/s10668-022-02875-6

Protective consumption behavior under smog: using a data-driven dynamic Bayesian network

Yu Yuan[1] · Bo Fan[2]

Received: 24 November 2021 / Accepted: 22 December 2022

Abstract
In the midst of the deteriorating air pollution and collective stress, people pay close attention to risk mitigation measures such as keeping indoor and purchasing anti-smog products. Through impact evaluations, factors regarding health protective behavior can be identified. However, limited research is available regarding probabilistic interdependencies between the factors and protective behavior and largely relies on subjective diagnosis. These concerns have led us to adopt a data-driven static Bayesian network (BN) and Dynamic BN model to help explore multidimensional factors that may influence the public's health protective behavior of buying anti-smog air purifiers and examine the dependencies among network nodes. Using the city-level aggregate data from an online shopping platform, the results shed new light on relationships existing among 11 factors and protective behavior of buying air purifiers. Furthermore, taking into account the dynamic nature of protective behavior, we add time-related factors on the basis of static BN to construct the dynamic BN model. Results indicate that PM2.5 concentration and product price are the two leading factors affecting the consumption behavior for air purifiers. Additionally, media-related factors play an important role in the consumption behavior. This study contributes to the fields of impact evaluation of protective consumption behavior analysis and links environment risk with public consumption by identifying key factors.

Keywords Smog · Protective consumption behavior · Bayesian network · Product price · Media

✉ Yu Yuan
yuanyu24@sjtu.edu.cn

Bo Fan
fanbo@sjtu.edu.cn

1 China Institute for Urban Governance, Shanghai Jiao Tong University, Shanghai, People's Republic of China

2 School of International and Public Affairs, Shanghai Jiao Tong University, Shanghai, People's Republic of China

Published online: 30 December 2022

1 Introduction

The protective behavior of the public under risk is worth considering (Dai et al., 2020; Quan et al., 2022). For example, the COVID-19 has had a great impact on people's physical and mental health (Torales et al., 2020), and even once there was a shortage of masks (Usman et al., 2022) and panic buying, posing a great challenge to the healthcare supply chain (Sriyanto et al., 2021). Similarly, severe air pollution has drawn global attention. Nearly one fourth of China's territory has been affected by smog, and the average national number of smog days in 2017 was 27.5 (Zhang et al., 2018). The adverse effects of smog on human health and the environment are significant worldwide due to the myriad of pollutants dispersed into the atmosphere (Raza et al., 2020). The extreme smog crisis harms people's health, causing cough, nasal irritation, shortness of breath, and respiratory infections (Raaschou-Nielsen et al., 2016). People who frequently experience city smog pay close attention to risk mitigation and self-protective measures, such as reducing outdoor activities and using green products (Yadav & Pathak, 2016). Considering the serious health and social problems, collective purchase for anti-smog products is emerging in recent years, the so-called smog economy. The respiratory tract discomfort caused by the smog has also spiked the sales of related drugs and health products.

When the risk of smog becomes a social consensus, collective pressure is likely to cause herding behavior, consumers' purchase decisions are often emotion related (Eysenck et al., 1985). However, evidence exists of a cognitive component in health protective behavior (Rook & Fisher, 1995), for example, perceived consumer effectiveness (Kautish et al., 2021), Value compatibility (Dhir et al., 2021), and social impact (Awan & Raza, 2012). Therefore, such protective behavior even panic buying phenomenon occurs within the dynamic and complex social environment due to several factors (Wu et al., 2017). To identify such factors, previous studies explored demographic characteristics, individual perceptions, and social psychological factors, such as risk-coping behavior (Pan et al., 2020; Wang et al., 2021). Regarding the stream of literature of marketing (Awan, 2011), researchers often pay attention to consumer- and product-related factors (Tsao et al., 2019). Although these studies focused solely on 1D factors, they also highlighted possible dependencies among the factors within the complex environment.

Although many factors affect protective consumption behavior under risk, practicing government and product managers must prioritize them regarding their importance. Identifying factors that contribute to protective behavior is vital toward knowing how to provide uninterrupted anti-smog products and maintain social stability (Scheffran et al., 2012). For companies to be successful in social sustainability issues (Awan et al., 2020), they must first manage strategic enablers and monitor tactical enablers to achieve sustainability goals (Awan & Sroufe, 2022). In addition, nonlinear, multi-dimensional factors such as consumer awareness (Qiao et al., 2021), disruptions and uncertain business environment (Joshi & Sharma, 2022) need to be taken into account.

This study mainly aims to evaluate the relationship between multidimensional factors and protective behavior and establish the relative importance of the factors through the lens of BN. On the basis of social amplification framework of risk, we selected three media- and four product-related factors from consumer perspective. Given that our data are at the city level, four city-related factors are also taken into consideration. Furthermore, to characterize the evolution of protective behavior over a longer period of time, we added three time-related factors to the DBN model. This study performed sensitivity and importance analyses to explore the influences of factors on public's real purchase decision during smog crisis. The findings have implications for our understanding of consumers' behavior evolution during smog period and provide useful guidelines of early interventions for suppliers to maintain balance of supply and demand from the comprehensive perspective and for authorities to regulate the irrational collective behavior.

This study has several contributions. First, we use real sales data of air purifiers to measure the public's protective consumption behavior, which is an improvement over previous studies that used self-reported behavioral intentions (Wu et al., 2018), as there is often a gap between intentions and real behavior (ElHaffar et al., 2020; Park & Lin, 2020). Second, this paper uses a data-driven approach to obtain the network structure and relationships among protective consumption behaviors and their influencing factors, which avoids the possible subjective limitations of previous subjective diagnosis based on the expert knowledge (Yang et al., 2021). Third, existing studies generally focus on considering one aspect that influences protective consumption behavior, such as marketing strategies from a firm-side perspective (Stanciu et al., 2020; Untaru & Han, 2021), risk perception on the public side (Mehiriz & Gosselin, 2022), but the protective consumption behavior of the public in the real world is influenced by multiple factors (Awan & Raza, 2012), which are not independent of each other. Certain limitations in the survey data and statistical approach exist in identifying complex relationships between multidimensional factors and target variables (Simsekler & Qazi, 2020). Adopting BN which as an approach with high inference capability (Amin et al., 2021), its effective explanation of uncertainty and nonlinear features is a more appropriate approach for protective consumption behavior influenced by multiple factors. In addition, DBN overcomes the misleading results that static BN may produce due to cross-sectional data (Liu et al., 2021), and considering time-varying data can facilitate the understanding of protective consumption behavior evolution (Hosseini et al., 2020).

The outline of the study is as follows. Sections 2 presents the literature review on factors of health protective buying. Section 3 describes the methodology, such as BN and DBN model and the data source. Section 4 presents the results and interpretations. Section 5 holds the discussion and implications of our study. Finally, Sect. 6 covers the conclusions and limitations of this study.

2 Literature review

Despite the extant studies on protective behaviors under smog crisis from academia (Bickerstaff & Walker, 1999), few studies have examined the public's buying behavior of anti-smog products and its influential factors. Buying air purifiers is a consumption decision which is driven by marketing factor as well as by social environment. Existing studies believe that individuals' health beliefs and social capital (from the perspective of social interaction) have an impact on self-protection behavior (Chuang et al., 2015; Verroen et al., 2013). Under circumstances of health threats and social panic, individuals who are uncertain about the consequences of smog often impulsively purchase anti-smog products to avoid or mitigate the risk (Wang et al., 2021). Such public herd consumption behaviors have been frequently observed when unexpected events occur. For example, when snowstorm warnings were issued, consumers rushed to stores and stockpiled excessive amounts of food.[1] During the severe smog in China, the sales of masks in online platforms rose sharply, and the sales during the red warning period were 9.3 times the usual.

Being at the intersection between consumer behavior and risk management research (Zheng et al., 2020), protective behavior toward buying anti-smog products represents a relatively unexplored research area where purchase decisions for emergency products differ from other usual consumer behaviors (Samson & Voyer, 2014). Hence, understanding the causes and their importance has important theoretical value and practical implication. The 2020 surge in the discussions on protective behavior was triggered by the COVID-19 pandemic (Kim et al., 2022; Milošević et al., 2022). Zickfeld et al. (2020) investigated the demographic and psychological factors that influence the public to adopt protective behaviors, as well as the changes in attitudes and emotions over time. Singh et al. (2021) identified factors that influenced consumer panic buying behavior during the New Crown pneumonia pandemic. Through questionnaire after the onset of COVID-19 epidemic in Fiji, consumer attitudes and subjective norms, scarcity of goods, and time pressure positively influenced consumer purchase intentions. However, the current literature of protective consumption behavior is quite limited and only a few have directly examined the causes of purchase behavior under crisis (Yuen et al., 2020).

Various factors may play a role in the occurrence of protective consumption behavior in different crises. Risk exposure significantly influences the public's responses to potential and existing threats (Pan et al., 2019). People living in high hazard areas are more likely to adopt risk mitigation measures than those living in lower earthquake probability. In the context of air pollution, personal risk exposure has consistently been confirmed to impact public risk perception which might induce coping behavior and even panic buying (Wang et al., 2021; Zeidner & Shechter, 1988). Highly polluted areas are not exempted from herd buying of anti-smog products due to reduced visibility and physical discomfort caused by high PM2.5 concentration. PM2.5 concentration is higher in the cold winter and lower in the summer because of the coal-burning-based heating method. For instance, poor air quality occurs yearly in Poland when houses and buildings require heating (Woźniak et al., 2020); the Huai River heating policy in China provided coal-based heating for cities north of the river, greatly increasing the pollution in these cities (Almond et al., 2009). In addition, the industrialization and urbanization will affect the air pollution levels of cities, for example, the famous 1952 London smog incident and the 1955 Los Angeles smog crisis are two typical cases.

Such factors may exacerbate people's perceived threats through media communication (Naeem & Ozuem, 2021). The social amplification of risk discussed by Kasperson et al. (1988) refers to the effects of the broader psychological, social, institutional, and cultural processes affecting people's perception and behavioral responses. Social information about risk is also an important explanatory variable for the public's coping behaviors (Dootson et al., 2022; Yeo et al., 2022). Media-transmitted information can amplify the risk and social effect not only by evoking personal experiences, but also by influencing those without such direct exposure (Wachinger et al., 2013). The official reports and non-government information (refers to

[1] Time. Panic shopping! how a blizzard turns us into irrational hoarders at the grocery store. 2015. http://time.com/money/3682510/blizzardpanic-shopping-groceries/

information posted and disseminated on social media in this study) have also been considered critical predictors of responses to environmental disasters (Lindell & Hwang, 2008). For example, Grothmann and Patt (2005) found that risk perception and response behavior regarding climate change are shaped by the information disseminated by the media. Wang et al. (2021) used air pollution-related media, including reports from authorities and social media to predict people's perceived threats to air pollution. Further, the negative emotion of risk, such as depression, fear of unknown, and collective stress also impacts the public's response, compensatory control theory considers the purchase behavior toward personal protective equipment as a way to compensate for the social presence, because the fear appeal is known as in anticipation of seeking affection, acceptance, and social information (Addo et al., 2020). The social-interactive in shopping may lead to panic buying, which can help individuals regain a sense of "normalcy" and tide over the health crisis (Benight & Bandura, 2004).

Although protective consumption behavior is often emotion related (Eysenck et al., 1985; Kim et al., 2022), the public must still consider product-related factors in the decision to purchase anti-smog air purifiers. The purifier market was not mature in China in the early years, and the price was much higher than that of masks, the public's purchase of purifiers still requires a certain degree of cognitive effort and information search (Wu et al., 2017). Brand image plays an important role in winning the favor of consumers in market competition (Anwar et al., 2011). Effective marketing communications, such as advertisements can increase consumers' familiarity with the brand and reduce the cost of information search for consumers (Im et al., 2016). Correspondingly, the word-of-mouth of a product, especially recommendations from peers, can increase the trust for a product. Another factor that must be considered when making a purchase is product price. Pan et al. (2020) unveiled that household income significantly affects consumers' protective consumption intention as hurricanes approach.

With smog crisis, unlike some natural hazards (e.g., earthquakes, flood), the public have some time to prepare, given that smog is a gradual risk and evolves regularly over time (Mehiriz & Gosselin, 2022). The other avoiding behavior and alternative mitigation options are allowed because individuals can gather enough information and afford some time cost. Therefore, the sales of air purifiers tend to be higher on some promotional days and festivals (Wang et al., 2021). However, ultimately, the individual can only do so much to maintain rationality and must submit to his/her inability to tolerate further health threats and collective stress.

Most studies focused solely on factors of a single dimension or some without simultaneously considering all factors in a complex environment. Therefore, their potential interdependencies have not been discussed with the use of probabilistic and graphical models. Ultimately and moving beyond the previous studies that did not consider factors nonlinear relationships and discussed only protective intention rather than real consumption behavior (Wu et al., 2018), the proposed approach in Sect. 3 may be useful in understanding the dynamic nature and causes as to why consumers impulse purchase air purifiers following a smog. The remainder of this study will present the results of the relative importance of factors to the public's protective consumption behavior over the 2016/2017 timeframe in China.

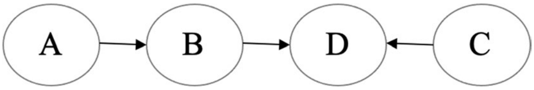

Fig. 1 Generic BN

3 Materials and methods

3.1 Bayesian network (BN)

Bayesian network is based on Bayesian statistical methods, which could fit the nonlinear and complex interactions among variables (Sener et al., 2019). BNs have been extensively explored in earlier studies, especially in risk and uncertainty-related application areas due to their advantage in identifying and assessing the impact of factors in different what-if scenarios. On the one hand, BN employs a probabilistic reasoning to capture the network characteristic and distribution which is in line with the nature of risk analysis (Shortridge et al., 2017). On the other hand, BN provides a graphical representation of complex relationships among factors, and decision-makers can identity critical components because the model helps to visualize the propagation impact of the variable, that is how a change in interested variable impacts other variables across a network (Kabir & Papadopoulos, 2019). Owing to the feature and advantage of BNs, previous study has assessed the critical components and risk forecasting (Adedipe et al., 2020). For instance, Simsekler and Qazi (2020) identified organizational factors of patient safety errors to help managers to prioritize them on the basis of their importance.

A BN model is a probabilistic model of directed acyclic graph. The model comprises a set of nodes representing variables and arcs that either represent causality or statistical dependencies among interconnected variables (Hanea et al., 2018). The strength of dependencies among interconnected variables is captured through probability distributions. The directed links ($A \rightarrow B$, $B \rightarrow D$, $C \rightarrow D$) in Fig. 1 represent the dependence of indicators on the target variable, and past beliefs on uncertain variables are efficiently updated with the evidence against distinct sources in the network.

This method has one major limitation. Empirical data and expert judgment can be used to construct a BN model (Hanea et al., 2018), however, in most scenarios the data are scarce or even unavailable, thus expert judgment is a more popular approach in existing literature although it is quite challenging (Werner et al., 2017). To explore the application of BNs in the health protective behavior field with a unique survey dataset, this study used a data-driven approach to obtain conditional probability distribution and develop our BN model.

3.2 Dynamic Bayesian networks (DBN)

Another limitation of BN is that it is static and unsuitable for representing dynamic relationships among variables. A DBN is a derivative of the static BN by adding the time dimension, which can capture the temporal relationships and predict the future probability of a variable using prediction inference. Nowadays, DBNs are widely used to model dynamic processes in many other fields, such as resilience assessment (Tong et al., 2020), reliability assessment (Rebello et al., 2018), and others. Protective consumption behavior is susceptible to a variety of factors when faced with unexpected events, irrational herd buying behavior is prone to appear under collective stress, leading to abrupt changes in demand and then substantial stock-outs, increasing-price. However, health

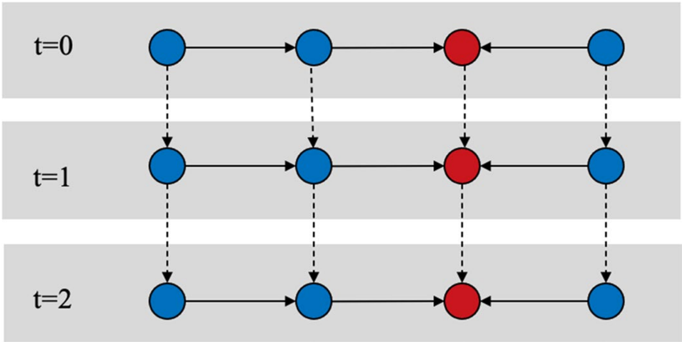

Fig. 2 Simple DBN over three time periods

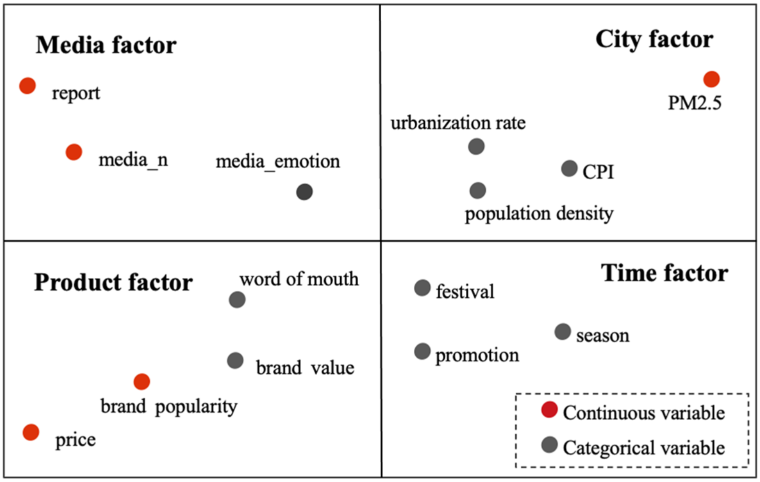

Fig. 3 Factors influencing the occurrence of protective consumption behavior

protective behavior may emerge as intervention ineffective if all the causes are treated with equal weight, and identifying the most critical variable is necessary. The nature of variables is dynamic and then the protective consumption behavior evolves over time. Hence, DBN was utilized to capture this temporal nature and identify the most important causes/variables. This study mainly aimed to visualize the complex relationships among self-protective behavior causes and identify critical causes to prevent and mitigate unreasonable behavior. Although BN can meet the requirement of identifying causes, it fails to explain the dynamic relationships among variables and even ignores time-related factors. Thus, we proposed a DBN model to achieve this goal.

A DBN comprises multiple BNs (referred to as time slices), the process in a DBN is stationary and the network structure repeats after the 2nd time slice, so the variables for the slices $t=2$, 3,..,T remain unchanged. The dotted lines are forward moving between the successive time slices and can never go back to the prior time slices. Figure 2 shows a simple DBN of three time slides. The blue nodes (indicators) are the parent nodes of

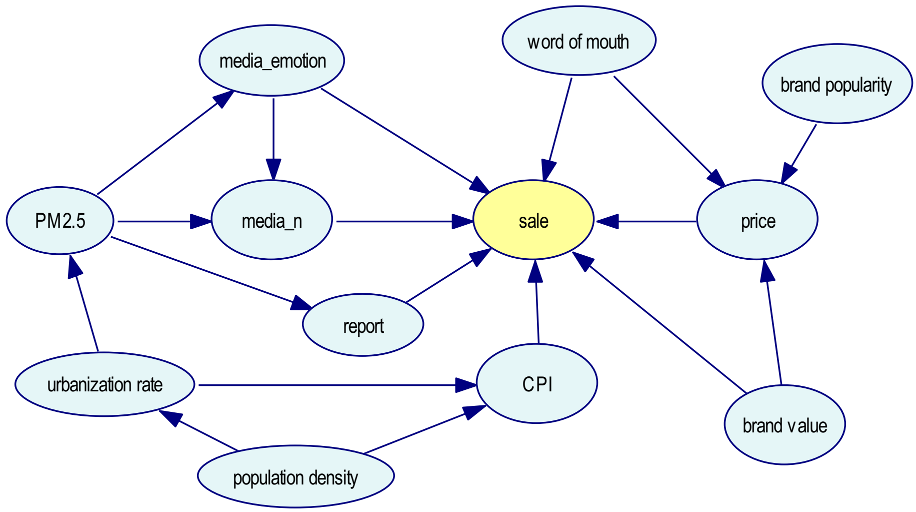

Fig. 4 Structure of BN for air purifiers sales

the red nodes (target variable) in the same time slice, and the probability of the target variable also depends on its value from prior time slices.

3.3 Data sources

In this study, we evaluated the impact of factors on health protective behavior toward buying anti-smog products on the basis of BN. As described above, the protective consumption behavior is affected by the following four dimensions and 13 variables (including time dimension), as shown in Fig. 3. Several continuous variables, including the number of reports (report) and media information (media_n), product price (price) and air quality (PM2.5) were discretized in subsequent model settings. For categorical variables, we followed the original number of categories, for example, media_emotion is a binary variable, then, we assigned two states to this variable (positive and negative) to maintain a low complexity of the network, while we assigned four states to the variables, such as CPI and season, coded as 4, 3, 2, and 1.

Given the variation of smog in each city of China, we targeted five cities (Beijing, Chengdu, Shijiazhuang, Shanghai, and Xi'an) and five product brands (Xiaomi, Midea, Panasonic, Philips, and Blueair) as our sample. The data include four parts as shown in Fig. 4. We collected the variables included in the model from exclusively renowned literary publications. The four dimensions have been proposed ensuring a good coverage of the influencing factors of the purchase decision for anti-smog air purifiers. To reduce the overlap among variables, we exerted substantial effort to remove duplicate variables. This led to a condensed list of indicators.

The first part is the media-related data which we collected using the Python algorithm. We used the key words, such as "smog," "PM2.5," and "air pollution" to search for news reports from authoritative media agencies, such as China Daily, CCTV (China Centre TV), and other official media. As for social media, we selected SINA micro-blog, the most popular social platform in China, as data source of media-related indicators. We collected all the official microblog using key words above and recorded the release time of each post.

Table 1 Sensitivity of the prediction ability of BN with different discretization schemes

Discretization Scheme (Number of States)	Sale	Price	pm2.5	Media-n	Brand popularity	Report	Accuracy rate
Scheme 1	2	2	2	2	2	2	0.722
Scheme 2	3	3	3	3	3	3	0.726
Scheme 3	4	4	3	3	3	3	0.717
Scheme 4	4	4	4	4	4	4	0.792
Scheme 5	5	5	4	4	4	4	0.774
Scheme 6	5	5	5	5	5	5	0.742

The media_n is represented by the number of comments, and for media_emotion, we used text sentiment analysis method on the basis of word vectors to identify the sentiment polarity of online comments. During data preprocessing, we only deleted duplicate comments and comments with only emoticons. Given the characteristics of the Chinese language, we must perform word segmentation for text data. In this study, we used the package jieba[2] of R language for word segmentation.

The second is the city-related data, including air quality (PM2.5) and city inherent index (urbanization rate, population density, and consumer price index (CPI)). We collected the air quality data of five target cities from the official website of the China National Environmental Monitoring Center, and obtained the data of other three indicators from the Statistical Yearbook and Government Annual Report.

As for the product-related data, we collected the data of sales and prices of five target air purifiers brands from Taobao and Tianmao, which accounted for about 70% online market share. We selected Xiaomi, Midea, Panasonic, Philips, and Blueair selected because of their popularity among consumers, and they also have a wide price range (low, medium and high prices) of air purifiers. The unit of original data we obtained is one week, thus in this study, the analysis unit of dynamic variables is one week. Similar to the variable media_n and media_emotion, we measured word-of-mouth and brand popularity using online product comments. We used Python to crawl the comments text of each brand of the online shopping platform, the number of comments reflect the brand popularity, and the sentiment polarity of comments text can represent the word-of-mouth. Another variable is brand value, the data are from the 2017 Brand Finance Global 500. On the basis of the value ranking of five brands, we divided them into three levels: high, middle, and low. Ultimately, brand value and word-of-mouth are three categorical variables, and brand popularity is a continuous variable.

Finally, we also selected three time-related variables as a supplement to the DBN model. Festival and promotion are binary variables, if the data node contains festival or promotion days, the value is 1. The season has four states, namely, spring, summer, autumn, and winter. All data are from July 2016 to July 2017. The dataset includes 1400 pieces of data, including peaks and valleys of air purifier sales.

[2] 1QIN WENFENG,YANYI W.jiebar: Chinese text segmentation[M/OL].2019.https://CRAN.R-project.org/package=jiebaR.

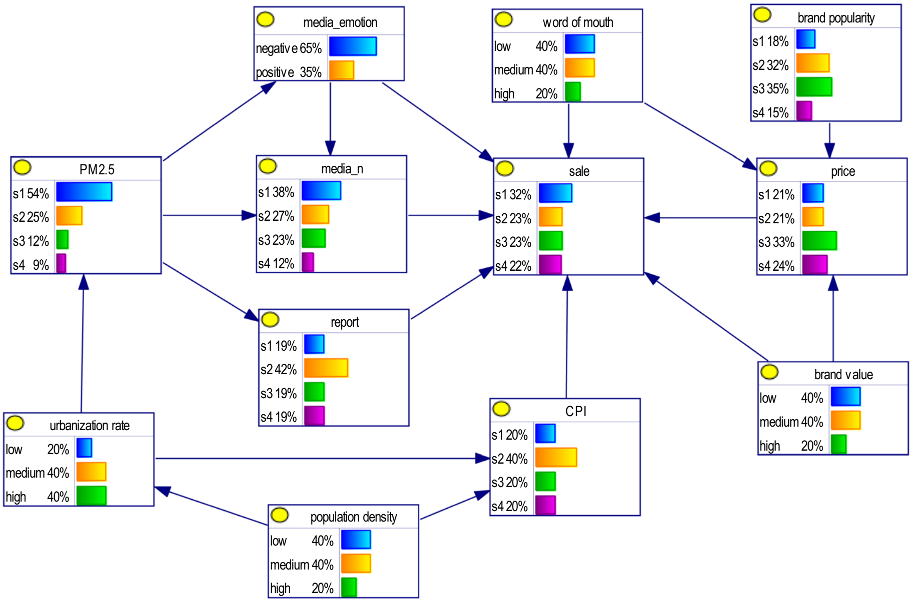

Fig. 5 Probability distribution of factors associated with sales of air purifiers

4 Results

4.1 BN

We developed the BN model in GeNIe (academic version) and tested six discretization schemes for continuous variables, as shown in Table 1. We adopted the Bayesian search algorithm to establish the sensitivity of discretization states and to provide insights into decision-making for further analysis. Resultantly, the accuracy rate of each scheme can be computed by a constructed confusion matrix, which is an output of the prediction capability of model. We divided the accuracy rate by the number of correct predictions by the total number. Further, we validated the BN model through k-fold cross-validation, this process divides the dataset into k parts of equal size, and run k times with the data of $k-1$ parts as the training set and the left part as the testing set (Marcot & Hanea, 2020). On the basis of the size and common practice, we set $k=5$ in this section and following this, the Scheme 4 captured more information on the original continuous variables distribution and highest accuracy rate.

Figure 4 shows the structure of the BN model, which comprises 12 nodes, including 11 factors that could influence engagement in protective behavior, and the study objective—purchase decision of anti-smog air purifiers (i.e., sales of each product for each city). In Fig. 4, the ovals denote the causes and sale (yellow) of this project, and arrows reflect causality among different variables. One can see that all of the identified variables can directly or indirectly affect the results of the final product sales. A priori probabilities are not determined by subjective expert assessments but are based on our objective data-driven to run the model. The probability distribution of factors associated with sales is shown in Fig. 5. Regarding this model, 22% of cases were associated with

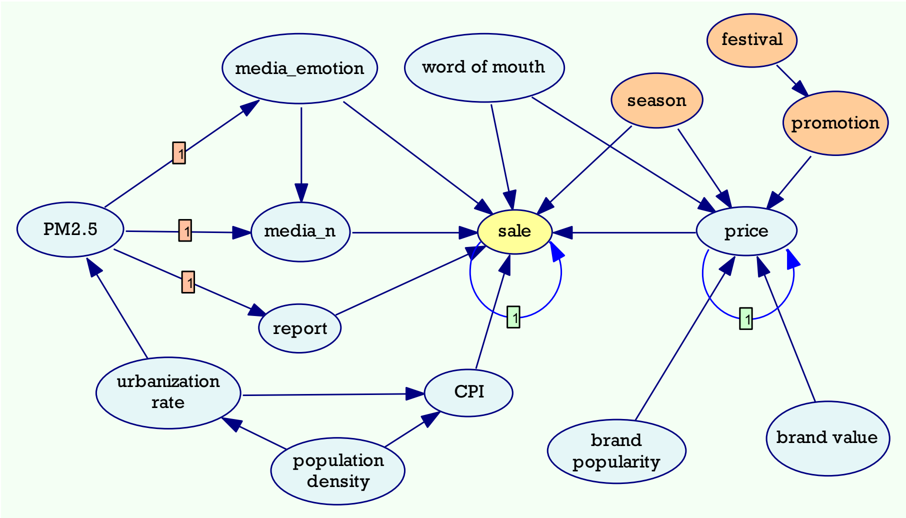

Fig. 6 DBN of protective consumption considering time factors

the State 4 (high) and 32% of cases were associated with the State 1 (low) of consumption behavior. The underlying distribution was highly skewed in the case of PM2.5, whereas others were uniform distribution.

4.2 DBN

Generally, the protective consumption phenomenon tends to be a process rather than a state; thus, considering the behavior variation of the public can be important. BN is unable to account for the time-related factors (festival, promotion, and season) in the analysis as they are limited to static systems. In this section, we propose a DBN to assess the protective consumption behavior in a dynamic manner.

On the basis of the previously constructed BN structure, we have added three time-related variables to the DBN, and established the DBN of a protective behavior as shown in Fig. 6. Among the variables, urbanization rate, population density, and CPI remain constant with time of concern. We assigned them in unconditional probability tables that do not change throughout the analysis. Dynamic variables, meanwhile, are varied by time, and consequently their state changes with time. For the sake of the study, we set each month as a time slice, and then $t = 13$. The dynamic links/arcs that join the temporal variables represent changes over time among the variables. The single digit numbers on the links and self-rolling arcs imply the temporal delay of influence (Khan et al., 2020). For instance, a link labeled as 1 between the variables PM2.5 and report imply the influence that takes one timestep. In the DBN, we assigned CPTs to variables that have father nodes in the same or different time slice. For example, we assigned "CPI" CPT that considers their father nodes (i.e., urbanization rate), whereas we assigned each dynamic variable (i.e., variable that has a temporal arc) a CPT that considers the variable itself at a previous time-slice. Given that this paper takes a data-driven approach, we can obtain CPT through learning.

The BN can also be employed in a backward analysis and provides a tool to identify the most vulnerable root cause (Kammouh et al., 2020). We set evidence of a protective

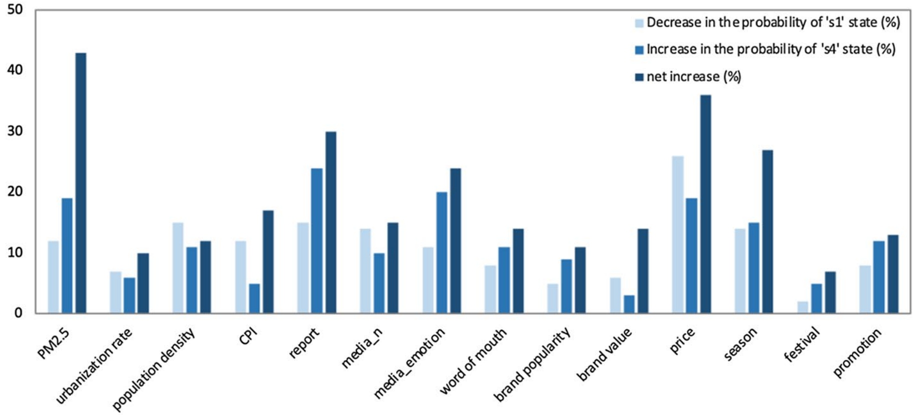

Fig. 7 Back propagation impact assessment given the protective consumption in “s4” state

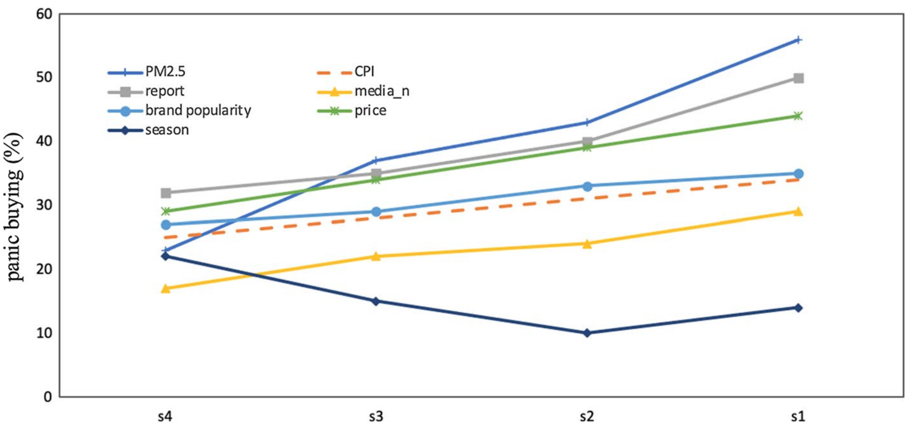

Fig. 8 Sensitivity analysis of root causes given the protective consumption in “s4” state

consumption behavior (p(S4) = 100%) and calculated the posterior probability of all the root causes. Figure 7 shows the increase in all the root causes after providing the protective consumption behavior evidence. Overall, the results showed that PM2.5, price, and season are the three leading critical factors in the back propagation impact assessment given the protective consumption in the “s4” state.

Finally, we conducted a sensitivity analysis for the seven variables with four states to identify sensitive variables in the DBN model, the results of which are shown in Fig. 8. Sensitivity assessment of factors can help prioritize factors for decision-making (Aven, 2019). The steepness of the slope can identify the most sensitive causes. The higher the steepness, the more sensitive the cause. We identified factors PM2.5, report, and price as sensitive factors. On the basis of the aforementioned importance analysis (Fig. 8), we marked PM2.5 and price as critical variables because each of them had a relatively high degree of impact on the health protective behavior toward buying anti-smog products, the two variables have a high level of sensitivity in Fig. 8, which confirms the strength of the proposed approach to identify the most vulnerable causes. Less vulnerable causes are found to be almost horizontal, such as brand popularity, CPI, and media_n. For the variable season, if considering the period from summer to winter, (i.e., s2 → s4), the season also has the steeper slope. The sensitivity analysis of other non-four-category variables is not shown in the figure. For the media_emotion with only two states (positive or negative), the change in the probability distribution of protective consumption is up to 21% concerning the extreme two states. In addition, the probability changes about 16% for word-of-mouth.

5 Discussion

Several factors may influence the health protective behavior toward buying anti-smog air purifiers. Using online sales data, we proposed BN and DBN model to explore four dimensions of factors and their interdependencies and relative importance toward protective consumption. Our results revealed that the factors, including smog concentration (PM2.5), product price, season, and media-related factors (report and media emotion) are critical factors with respect to protective consumption of air purifiers.

In the literature, media-related factors have been discussed as contributing factors to the health self-protective behavior. Specifically, media-related factors (e.g., official report and emotion) is a trending topic that its role on herd behavior, including impulse buying and health protection behavior have been investigated widely in various risk scenarios (Naeem & Ozuem, 2021). For instance, information from official channels is significantly positive to the adoption of protective behavior, such as the purchase of flood insurance (Lindell & Hwang, 2008). Evidence from China showed that individuals are likely to be affected by government and non-government advices, the information from later is easier to form a herd effect and in this process, protective consumption of air purifiers has become an inevitable phenomenon as the smog intensifies and the information is spread through the media (Wu et al., 2017). Although social media has become an increasingly widespread source of news, official reports remain the core framework for health emergencies (Dry & Leach, 2010). The results of this paper may point to a new finding that the number of official reports and social media emotion together drive protective consumption behavior.

Smog concentration is an important factor to be found as a critical one affecting protective consumption behavior. Air pollution is generally most extreme in winter, and the increase in PM2.5 concentration entails reduced visibility and respiratory diseases (Forsyth, 2014). These risk exposures have significantly attracted the attention of the public and has become a wide-ranging issue, which induces engagement of protective consumption. Further, recent empirical studies also showed that individuals search more online for anti-smog products, including masks and air purifiers when PM2.5 increases (Liu et al., 2017). Especially along with air pollution deteriorates and even exceeds the warning threshold, the sales of anti-smog products would be doubled. Other previous findings on protective behavior against air pollution also support our results (Graff Zivin & Neidell, 2013; Sun et al., 2017).

Correspondingly, many studies support the cost has negatively affected the adoption of protective behavior (Chang & Wildt, 1994). The protective consumption behavior in this article differs from general protective behavior or stockpiling buying. In our study, the air purifiers have higher prices as durable household goods rather than consumables, such as masks (Dianshu et al., 2010), the price is an important factor when making a purchase decision. Although the herd effect still exists, this herd buying phenomenon is more obvious when the price is lower.

The results of this study contribute to the fields of evaluating factors and their impact on protective consumption behavior when the public faces air pollution. Compared with other studies that have focused on partial factors that influence engagement in self-protective behavior, in this study, we identified the factors with four dimensions and analyzed the interdependencies of these factors using a BN model, and then we quantified the importance and sensitivity of the different factors. In previous studies, the impact of factors for the protective behavior toward buying anti-smog air purifiers were typically measured on the basis of regression analysis and the interdependencies of factors receive less attention. In this study, we integrate the direct and indirect impacts of factors in the importance evaluation using a network-based perspective.

In addition, this study adopts a data-driven BN approach to provide significant insights into the relative importance of the factors for protective consumption behavior. To the best of our knowledge, this is the first study to systematically evaluate the impact of the factors on consumer's protective consumption decision for air purifiers under smog using objective data. In this study, the data-driven DBN model has implication on using real-time and dynamic data to analyze protective behavior regarding the factors' relative importance. Traditional methods for impact evaluation of health protective behavior generally focus on the subjective risk perception of public. The statistical approaches might have failed to identify complex relationships between multidimensional factors and target variables because of the limitations in validity and reliability in a survey data. A DBN model represents either causal probabilistic relationships or statistical dependencies among interconnected variables.

Finally, the results of this study offer insights into the evolution of health protective consumption process considering the time-varying factors by DBN model. When dynamic features are exploited, the DBN analysis goes beyond basic static analysis and potentially reveals deeper insights and valuable findings. A key factor for improving public response to harmful smog crisis is to develop a better understanding of how individuals' protective consumption behavior fluctuates over time. This is essential to identify which periods are most at risk and serve as guidance to intervene and reduce irrational herd behavior.

Fundamentally, the findings of this study can help government and suppliers better understanding the public's health protective behavior toward buying air purifiers. First, we summarized factors of four dimensions that may lead to protective purchase behavior under smog on the basis of a literature review. These factors can be used as a checklist to identify the potential predictors of protective behavior. A comprehensive identification of protective consumption behavior can help government design communication strategies and suppliers adjust product supply and sales plan. Second, we confirmed the effectiveness of the BN for evaluating the relative importance of factors with respect to protective consumption behavior. Practitioners can focus on the identified critical factors and take appropriate early interventions. Finally, the DBN approach can also be used to predict the evolution of health protective behavior. Institutions and managers can make effective predictions on purchase decision for protective products under the risk of smog by utilizing a large amount of objective data rather than survey data.

6 Conclusions

This study effectively identifies critical factors that significantly affect protective consumption behavior. We use a BN model to evaluate certain impacts of factors on purchasing anti-smog air purifiers under severe smog, with consideration given to the dependencies among factors. To investigate the evolution of protective consumption behavior under smog, we develop a DBN model on the basis of the proposed BN. Utilizing the week-level aggregate empirical data from multiple sources, adoption of the data-driven BN and DBN model revealed that PM2.5 and price are the leading factors that drive protective consumption behavior. Additionally, we find that media-related factors have the medium level of impact on the buying of air purifiers. We expect that the findings of this study contribute to the body of knowledge regarding protective consumption behavior evaluation and intervention irrational group behavior.

A few limitations must be acknowledged. First, the aggregate data for model learning are based on a weekly scale; therefore, future research could attempt to analyze at finer time scales to provide more timely and rapid decision-making guidance. Second, this study is based on the data from the online sales of air purifiers, the generalizability and transferability of the results to other anti-smog products (for example, masks or even some anti-smog foods) and even a broader range of emergency products may be limited. Additionally, the main findings of this study have not been validated in data sample of other cities or countries. Therefore, future research could also investigate the effectiveness of the proposed model in different anti-smog products or countries to help create a more holistic conclusion regarding the interdependencies between influencing factors and protective consumption behavior.

Acknowledgements All authors are grateful to the Journal's Editor and anonymous referees whose insightful comments helped to improve the quality of the manuscript.

Funding Funding was provided by National Natural Science Fund of China (Grant Numbers 72134005, 71974128), National Natural Science Foundation of China (Grant Number 72274123) and the Ministry of Education of China (Grant Number 19JZD022).

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability Data were analyzed in GeNIe (academic version).

Declarations

Conflicts of interest The authors declare no competing interests.

Consent to participate Not applicable.

Consent for publication Not applicable.

Ethical approval Not applicable.

References

Addo, P. C., Jiaming, F., Kulbo, N. B., & Liangqiang, L. (2020). COVID-19: Fear appeal favoring purchase behavior towards personal protective equipment. *The Service Industries Journal, 40*(7–8), 471–490.

Adedipe, T., Shafiee, M., & Zio, E. (2020). Bayesian network modelling for the wind energy industry: An overview. *Reliability Engineering & System Safety*, *202,* 107053.

Almond, D., Chen, Y., Greenstone, M., & Li, H. (2009). Winter heating or clean air? Unintended impacts of China's Huai river policy. *American Economic Review, 99*(2), 184–190.

Amin, M. T., Khan, F., Ahmed, S., & Imtiaz, S. (2021). A data-driven Bayesian network learning method for process fault diagnosis. *Process Safety and Environmental Protection, 150*, 110–122.

Anwar, A., Gulzar, A., Sohail, F. B., & Akram, S. N. (2011). Impact of brand image, trust and affect on consumer brand extension attitude: The mediating role of brand loyalty. *International Journal of Economics and Management Sciences, 1*(5), 73–79.

Aven, T. (2019). The call for a shift from risk to resilience: What does it mean? *Risk Analysis, 39*(6), 1196–1203.

Awan, U. (2011). Green marketing: Marketing strategies for the Swedish energy companies. *International Journal of Industrial Marketing, 1*(2), 1.

Awan, U., Khattak, A., Rabbani, S., & Dhir, A. (2020). Buyer-driven knowledge transfer activities to enhance organizational sustainability of suppliers. *Sustainability, 12*(7), 2993.

Awan, U., & Raza, M. A. (2012). Green consumer behavior: Empirical study of Swedish consumer behavior. *Recent Researches in Economics, 1*, 89–104.

Awan, U., & Sroufe, R. (2022). Sustainability in the circular economy: Insights and dynamics of designing circular business models. *Applied Sciences, 12*(3), 1521.

Benight, C. C., & Bandura, A. (2004). Social cognitive theory of posttraumatic recovery: The role of perceived self-efficacy. *Behaviour Research and Therapy, 42*(10), 1129–1148.

Bickerstaff, K., & Walker, G. (1999). Clearing the smog? Public responses to air-quality information. *Local Environment, 4*(3), 279–294.

Chang, T. Z., & Wildt, A. R. (1994). Price, product information, and purchase intention: An empirical study. *Journal of the Academy of Marketing Science: Official Publication of the Academy of Marketing Science, 22*(1), 16–27. https://doi.org/10.1177/0092070394221002

Chuang, Y.-C., Huang, Y.-L., Tseng, K.-C., Yen, C.-H., & Yang, L. (2015). Social capital and health-protective behavior intentions in an influenza pandemic. *PLoS ONE, 10*(4), e0122970.

Dai, B., Fu, D., Meng, G., Liu, B., Li, Q., & Liu, X. (2020). The effects of governmental and individual predictors on COVID-19 protective behaviors in China: A path analysis model. *Public Administration Review, 80*(5), 797–804.

Dhir, A., Malodia, S., Awan, U., Sakashita, M., & Kaur, P. (2021). Extended valence theory perspective on consumers'e-waste recycling intentions in Japan. *Journal of Cleaner Production, 312*, 127443.

Dianshu, F., Sovacool, B. K., & Vu, K. M. (2010). The barriers to energy efficiency in China: Assessing household electricity savings and consumer behavior in Liaoning Province. *Energy Policy, 38*(2), 1202–1209. https://doi.org/10.1016/j.enpol.2009.11.012

Dootson, P., Kuligowski, E., Greer, D. A., Miller, S. A., & Tippett, V. (2022). Consistent and conflicting information in floods and bushfires impact risk information seeking, risk perceptions, and protective action intentions. *International Journal of Disaster Risk Reduction, 70*, 102774.

Dry, S., & Leach, M. (Eds.). (2010). Epidemics: science, governance and social justice. Routledge.

ElHaffar, G., Durif, F., & Dubé, L. (2020). Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions. *Journal of Cleaner Production, 275*, 122556.

Eysenck, S. B. G., Pearson, P. R., Easting, G., & Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness and empathy in adults. *Personality and Individual Differences, 6*(5), 613–619.

Forsyth, T. (2014). Public concerns about transboundary haze: A comparison of Indonesia, Singapore, and Malaysia. *Global Environmental Change, 25*, 76–86.

Graff Zivin, J., & Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature, 51*(3), 689–730.

Grothmann, T., & Patt, A. (2005). Adaptive capacity and human cognition: The process of individual adaptation to climate change. *Global Environmental Change, 15*(3), 199–213.

Hanea, A. M., McBride, M. F., Burgman, M. A., & Wintle, B. C. (2018). The value of performance weights and discussion in aggregated expert judgments. *Risk Analysis, 38*(9), 1781–1794.

Hosseini, S., Ivanov, D., & Dolgui, A. (2020). Ripple effect modelling of supplier disruption: Integrated Markov chain and dynamic Bayesian network approach. *International Journal of Production Research, 58*(11), 3284–3303.

Im, I., Jun, J., Oh, W., & Jeong, S.-O. (2016). Deal-seeking versus brand-seeking: Search behaviors and purchase propensities in sponsored search platforms. *MIS Quarterly*, *40*(1), 187–204

Joshi, S., & Sharma, M. (2022). Impact of sustainable supply chain management on performance of SMEs amidst COVID-19 pandemic: An Indian perspective. *International Journal of Logistics Economics and Globalisation, 9*(3), 248–276.

Kabir, S., & Papadopoulos, Y. (2019). Applications of Bayesian networks and Petri nets in safety, reliability, and risk assessments: A review. *Safety Science, 115*, 154–175.

Kammouh, O., Gardoni, P., & Cimellaro, G. P. (2020). Probabilistic framework to evaluate the resilience of engineering systems using Bayesian and dynamic Bayesian networks. *Reliability Engineering & System Safety, 198*, 106813.

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., & Ratick, S. (1988). The social amplification of risk: A conceptual framework. *Risk Analysis, 8*(2), 177–187.

Kautish, P., Sharma, R., Mangla, S. K., Jabeen, F., & Awan, U. (2021). Understanding choice behavior towards plastic consumption: An emerging market investigation. *Resources, Conservation and Recycling, 174*, 105828.

Khan, B., Khan, F., & Veitch, B. (2020). A Dynamic Bayesian Network model for ship-ice collision risk in the Arctic waters. *Safety Science, 130*, 104858.

Kim, J., Yang, K., Min, J., & White, B. (2022). Hope, fear, and consumer behavioral change amid COVID-19: Application of protection motivation theory. *International Journal of Consumer Studies, 46*(2), 558–574.

Lindell, M. K., & Hwang, S. N. (2008). Households' perceived personal risk and responses in a multihazard environment. *Risk Analysis: An International Journal, 28*(2), 539–556.

Liu, T., He, G., & Lau, A. (2017). Avoidance behavior against air pollution: evidence from online search indices for anti-PM 2.5 masks and air filters in Chinese cities. *Environmental Economics and Policy Studies, 20*(2), 1–39.

Liu, Z., Ma, Q., Cai, B., Liu, Y., & Zheng, C. (2021). Risk assessment on deepwater drilling well control based on dynamic Bayesian network. *Process Safety and Environmental Protection, 149*, 643–654.

Marcot, B. G., & Hanea, A. M. (2020). What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? *Computational Statistics, 36*, 1–23.

Mehiriz, K., & Gosselin, P. (2022). The Effect of perceived threats and response efficacy on adaptation to smog: An instrumental variables design. *Risk Analysis, 42*(5), 1042–1055.

Milošević, D., Middel, A., Savić, S., Dunjić, J., Lau, K., & Stojsavljević, R. (2022). Mask wearing behavior in hot urban spaces of Novi Sad during the COVID-19 pandemic. *Science of the Total Environment, 815*, 152782.

Naeem, M., & Ozuem, W. (2021). Customers' social interactions and panic buying behavior: Insights from social media practices. *Journal of Consumer Behaviour, 20*(5), 1191–1203.

Pan, X., Dresner, M. E., Mantin, B., & Zhang, J. (2020). Pre-hurricane consumer stockpiling and post-hurricane product availability: Empirical evidence from natural experiments. *Production and Operations Management, 29*(10), 2350–2380.

Pan, X., Dresner, M., Mantin, B., & Zhang, J. A. (2020). Pre-hurricane consumer stockpiling and post-hurricane product availability: Empirical evidence from natural experiments. *Production and Operations Management, 29*(10), 2350–2380.

Park, H. J., & Lin, L. M. (2020). Exploring attitude–behavior gap in sustainable consumption: Comparison of recycled and upcycled fashion products. *Journal of Business Research, 117*, 623–628.

Qiao, A., Choi, S. H., & Pan, Y. (2021). Multi-party coordination in sustainable supply chain under consumer green awareness. *Science of the Total Environment, 777*, 146043.

Quan, L., Al-Ansi, A., & Han, H. (2022). Assessing customer financial risk perception and attitude in the hotel industry: Exploring the role of protective measures against COVID-19. *International Journal of Hospitality Management, 101*, 103123.

Raaschou-Nielsen, O., Beelen, R., Wang, M., Hoek, G., Andersen, Z. J., Hoffmann, B., Stafoggia, M., Samoli, E., Weinmayr, G., & Dimakopoulou, K. (2016). Particulate matter air pollution components and risk for lung cancer. *Environment International, 87*, 66–73.

Raza, W., Saeed, S., Saulat, H., Gul, H., Sarfraz, M., Sonne, C., Sohn, Z.-H., Brown, R. J. C., & Kim, K.-H. (2020). A review on the deteriorating situation of smog and its preventive measures in Pakistan. *Journal of Cleaner Production*, *279,* 123676.

Rebello, S., Yu, H., & Ma, L. (2018). An integrated approach for system functional reliability assessment using dynamic Bayesian network and hidden Markov model. *Reliability Engineering & System Safety, 180*, 124–135.

Rook, D. W., & Fisher, R. J. (1995). Normative influences on impulsive buying behavior. *Journal of Consumer Research, 22*(3), 305–313.

Samson, A., & Voyer, B. G. (2014). Emergency purchasing situations: Implications for consumer decision-making. *Journal of Economic Psychology, 44*, 21–33.
Scheffran, J., Brzoska, M., Brauch, H. G., Link, P. M., & Schilling, J. (2012). *Climate change, human security and violent conflict: challenges for societal stability* (Vol. 8). Springer Science & Business Media.
Sener, A., Barut, M., Dag, A., & Yildirim, M. B. (2019). Impact of commitment, information sharing, and information usage on supplier performance: a Bayesian belief network approach. *Annals of Operations Research*, *303*(1), 125–158.
Shortridge, J., Aven, T., & Guikema, S. (2017). Risk assessment under deep uncertainty: A methodological comparison. *Reliability Engineering & System Safety, 159*, 12–23.
Simsekler, M. C. E., & Qazi, A. (2022). Adoption of a data-driven Bayesian belief network investigating organizational factors that influence patient safety. *Risk Analysis, 42*(6), 1277-1293.
Singh, G., Aiyub, A. S., Greig, T., Naidu, S., Sewak, A., & Sharma, S. (2021). Exploring panic buying behavior during the COVID-19 pandemic: A developing country perspective. *International Journal of Emerging Markets*. https://doi.org/10.1108/IJOEM-03-2021-0308
Sriyanto, S., Lodhi, M. S., Salamun, H., Sardin, S., Pasani, C. F., Muneer, G., & Zaman, K. (2021). The role of healthcare supply chain management in the wake of COVID-19 pandemic: Hot off the press. *Foresight, 24*, 429–444.
Stanciu, S., Radu, R. I., Sapira, V., Bratoveanu, B. D., & Florea, A. M. (2020). Consumer behavior in crisis situations. research on the effects of COVID-19 in Romania. *Annals of the University Dunarea de Jos of Galati: Fascicle: I, Economics & Applied Informatics*, *26*(1), 5–13.
Sun, C., Kahn, M. E., & Zheng, S. (2017). Self-protection investment exacerbates air pollution exposure inequality in urban China. *Ecological Economics, 131*, 468–474. https://doi.org/10.1016/j.ecolecon.2016.06.030
Tong, Q., Yang, M., & Zinetullina, A. (2020). A dynamic Bayesian network-based approach to resilience assessment of engineered systems. *Journal of Loss Prevention in the Process Industries, 65*, 104152.
Torales, J., O'Higgins, M., Castaldelli-Maia, J. M., & Ventriglio, A. (2020). The outbreak of COVID-19 coronavirus and its impact on global mental health. *International Journal of Social Psychiatry, 66*(4), 317–320.
Tsao, Y.-C., Raj, P. V. R. P., & Yu, V. (2019). Product substitution in different weights and brands considering customer segmentation and panic buying behavior. *Industrial Marketing Management, 77*, 209–220.
Untaru, E.-N., & Han, H. (2021). Protective measures against COVID-19 and the business strategies of the retail enterprises: Differences in gender, age, education, and income among shoppers. *Journal of Retailing and Consumer Services, 60*, 102446. https://doi.org/10.1016/j.jretconser.2021.102446
Usman, B., Zaman, K., Nassani, A. A., Haffar, M., & Muneer, G. (2022). The impact of carbon pricing, climate financing, and financial literacy on COVID-19 cases: Go-for-green healthcare policies. *Environmental Science and Pollution Research, 29*(24), 35884–35896.
Verroen, S., Gutteling, J. M., & De Vries, P. W. (2013). Enhancing self-protective behavior: Efficacy beliefs and peer feedback in risk communication. *Risk Analysis, 33*(7), 1252–1264.
Wachinger, G., Renn, O., Begg, C., & Kuhlicke, C. (2013). The risk perception paradox—implications for governance and communication of natural hazards. *Risk Analysis, 33*(6), 1049–1065.
Wang, F., Yuan, Y., & Lu, L. (2021). Dynamical prediction model of consumers' purchase intentions regarding anti-smog products during smog risk: Taking the information flow perspective. *Physica a: Statistical Mechanics and Its Applications, 563*, 125427.
Werner, C., Bedford, T., Cooke, R. M., Hanea, A. M., & Morales-Nápoles, O. (2017). Expert judgement for dependence in probabilistic modelling: A systematic literature review and future research directions. *European Journal of Operational Research, 258*(3), 801–819.
Woźniak, J., Krysa, Z., & Dudek, M. (2020). Concept of government-subsidized energy prices for a group of individual consumers in Poland as a means to reduce smog. *Energy Policy, 144*, 111620.
Wu, X., Hu, X., Qi, W., Marinova, D., & Shi, X. (2018). Risk knowledge, product knowledge, and brand benefits for purchase intentions: Experiences with air purifiers against city smog in China. *Human and Ecological Risk Assessment: An International Journal, 24*(7), 1930–1951.
Wu, X., Qi, W., Hu, X., Zhang, S., & Zhao, D. (2017). Consumers' purchase intentions toward products against city smog: Exploring the influence of risk information processing. *Natural Hazards, 88*(1), 611–632.
Yadav, R., & Pathak, G. S. (2016). Young consumers' intention towards buying green products in a developing nation: Extending the theory of planned behavior. *Journal of Cleaner Production, 135*, 732–739.
Yang, Z., Wan, C., Yang, Z., & Yu, Q. (2021). Using Bayesian network-based TOPSIS to aid dynamic port state control detention risk control decision. *Reliability Engineering & System Safety, 213*, 107784.

Yeo, J., Knox, C. C., & Hu, Q. (2022). Disaster recovery communication in the digital era: Social media and the 2016 southern Louisiana flood. *Risk Analysis, 42*(8), 1670–1685.

Yuen, K. F., Wang, X., Ma, F., & Li, K. X. (2020). The psychological causes of panic buying following a health crisis. *International Journal of Environmental Research and Public Health, 17*(10), 3513.

Zeidner, M., & Shechter, M. (1988). Psychological responses to air pollution: Some personality and demographic correlates. *Journal of Environmental Psychology, 8*(3), 191–208.

Zhang, S., Li, Y., Hao, Y., & Zhang, Y. (2018). Does public opinion affect air quality? Evidence based on the monthly data of 109 prefecture-level cities in China. *Energy Policy, 116*, 299–311.

Zheng, R., Shou, B., & Yang, J. (2021). Supply disruption management under consumer panic buying and social learning effects. *Omega*, *101*, 102238.

Zickfeld, J. H., Schubert, T. W., Herting, A. K., Grahe, J., & Faasse, K. (2020). Correlates of health-protective behavior during the initial days of the COVID-19 outbreak in Norway. *Frontiers in Psychology*, *11,* 564083.