Protective Consumption Behavior Under Smog: Using a Data-driven Dynamic Bayesian Network

Yu Yuan (✉ yuan24@sjtu.edu.cn)

Shanghai Jiao Tong University

Bo Fan

Shanghai Jiao Tong University

Research Article

Keywords: smog, protective behavior, Bayesian network, dynamic Bayesian model

Posted Date: July 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-636251/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

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, critical factors regarding health protective behavior can be identified. However, limited research is available regarding probabilistic interdependencies between the factors and protective behavior. In this study, we adopt a data-driven Bayesian network (BN) approach to measure impacts of various factors on protective behavior under smog crisis. We model the sales of air purifiers and related factors that might incur protective product purchase behavior as a BN 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 and propose the dynamic Bayesian model. Results indicate that PM2.5 concentration and product price are the two leading factors affecting public protective behavior of buying air purifiers. Additionally, media-related factors play an important role in the protective behavior. Ultimately, this study contributes to the fields of impact evaluation of public protective behavior analysis and links risk management with key factors.

1. Introduction

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 (Kim et al., 2015; 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; S, 2011). 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, and so on (Xiaodan Pan et al., 2020; Wang et al., 2021). Regarding the stream of literature of marketing, researchers often pay attention to consumer- and product-related

factors(Tsao et al., 2019). Althoughthese studies focused solely on 1D factors, they also highlighted possible dependencies among the factors within the complex environment.

Certain limitations in the survey data and statistical approach existin identifying complex relationships between multidimensional factors and target variables(Simsekler & Qazi, 2020). As a method with powerful reasoning, Bayesian network (BN) is widely adopted due to its efficient interpretation of uncertain and nonlinear characteristics. Considering that the time-variability issue also greatly impacts the performance of protectivebehavior evolution, the static BN may lead to misleading results in a dynamic environment.

These concerns have led us to adopt a data-driven static BN and Dynamic Bayesian Network (DBN) model to help explore multidimensional factors that may influencethe public's health protectivebehavior of buying anti-smog air purifiers. Static BN can be extended into DBN by introducing relevant temporal dependencies betweenrepresentations of the static network at different times, which allows modeling the dynamic evolution of the observation objects(Hosseini et al., 2020). Thus, DBN is more appropriate for monitoring and predicting public buying behaviorand capable of representing the risk states at any time with respect to BN.

Identifying factors that contribute to protective behavioris vital toward knowing how to provide sustainable anti-smog products and maintain social stability. This study mainly aims to evaluate the relationship between multidimensional factors and protectivebehavior and establish the relative importance of the factors through the lens of BN.On the basis ofsocialamplification framework of risk, we selected threemedia- and four product-related factors from consumer perspective. Given thatour 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 decisionduring smog crisis. To the best of our knowledge, this is the first study to systematically evaluate the impact of the factors on consumers' buying decisionsforairpurifiersunder smog risk using objective data. The findingshave implications for our understanding of consumers'behavior evolutionduring smog period andprovide 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.

The outline of the study is as follows. Sections 3 presents the literature review on factors of healthprotective 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, Section 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 productsand itsinfluential

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 towards 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. 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. through online questionnaire after the onset of COVID-19 epidemic in Slovenia, Terpstra (2011) explored the relationships between information seeking, perceived credibility of information sources and individuals' engagement in self-protective behaviors. However, 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 (X 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 (Lindell & Prater, 2002). 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 (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 PM_{2.5} concentration. Among all the contributing factors of smog, coal-dominated energy structure is the worst offender. 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. 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 (Shaw et al., 2004; Wachinger et al., 2013). The official reports and non-government information (refers to 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. The work of Terpstra (2011) found that medical professionals and scientists received relatively higher degree of credibility than mass media and officials when individuals adopted preventive and protective behavior during COVID-19 pandemic. 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 towards 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), 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 (Dahlén et al., 2009; 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. The other avoiding behavior and alternative mitigation options are allowed because individuals can gather enough information and afford sometime cost. Therefore, the sales of air purifiers tend to be higher on some promotional days and festivals. 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 help explain health protective behavior towards buying anti-smog products, the proposed

approach in Section 3 may be useful in understanding the 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.

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 (Aven, 2017; Shortridge et al., 2017). On the other hand, BN provides a graphical representation of complex relationships among factors, and decisionmakers can identify critical components because the model helps visualize the propagation impact of the variable, that is how a change of 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; Ale et al., 2014; Aven, 2016). 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 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 $t = 2, 3, \dots, T$ remain unchanged. The dotted lines are forward moving between the successive time slices and can never go back to the prior time slices. Fig.2 shows a simple DBN of three timeslides. The blue nodes (indicators) are the parent nodes of 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 towards buying anti-smog product 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. 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^[1] 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 per cent 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 1,400 pieces of data, including peaks and valleys of air purifier sales.

4. Results

4.1 BN

We developed the BN model in GeNIe(academic version)and testedsix 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.

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

Fig.4shows 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). InFig.4, the ovals denote thecauses and sale (yellow) of this project, and arrows reflect causality among different variables. One can see that all of theidentified variables can directly or indirectly affect the results of the final product sales. The probability distribution of factors associated with sales is shown in Fig.5. Regarding this model, 22% of the cases were associated with the State 4 (high) of protective consmption. The underlying distribution was highly skewed in the case of PM2.5, whereasothers were uniform distribution.

4.2 DBN

Generally, the protecitiveconsumption 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 protecitiveconsumption behavior in a dynamic manner.

On the basis of the previously constructed BN structure, we have added three time-related variables to the DBN, andestablished the DBN of a protective behavioras shown in Fig.6. Among the variables,

urbanization rate, population density, and CPI remain constant with time of concern. WE assigned them 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 $t=12$. 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 consumption behavior ($p(S_4)=100\%$), and calculated the posterior probability of all the root causes. Fig.7 shows the increase of 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; Hosseini & Ivanov, 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 towards 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., $s_2 \rightarrow s_4$), 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 towards 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 towards protective consumption. Our results revealed that the

factors, including smog concentration (PM_{2.5}), 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 and 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 PM_{2.5} 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 PM_{2.5} increases (Liu et al., 2018). 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 affects 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 towards 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 our knowledge, this is the first study evaluating a DBN model in this particular health protective consumption context. Identifying the importance of related factors is essential for government and suppliers to regulate unreasonable behavior and balance market supply and demand. In this study, the data-driven DBN model has implications 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. Further, when dynamic features are exploited, the DBN analysis goes beyond basic static analysis and potentially reveals deeper insights and valuable findings.

Finally, the results of this study offer insights into the evolution of health protective consumption process considering the time-varying factors by DBN model. 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 towards 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 BN 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

In this study, we use a BN to evaluate certain impacts of factors on the health protective behavior of 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 the smog scenario, we develop a DBN model on the basis of the proposed BN, where in we explore the protective behavior of

buying air purifiers and related factors that could incur these behaviors in successive time slices. This study effectively identifies critical factors that significantly affect protective consumption behavior. Although many factors affect protective consumption behavior under risk, practicing government and product managers must prioritize them regarding their importance.

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. Future studies may also integrate data not only from providers but also from consumers to investigate how consumers' perception is related to protective consumption.

A few limitations must be acknowledged. First, the aggregate data for model learning is based on a weekly scale; therefore, more fine-grained time scale may provide more timely and rapid decision-making guidance in the relative importance ranking of the factors. Second, the 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) 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.

7. Declarations

Funding

There is no funding.

Conflicts of interest/Competing interests

The authors declare no competing interests.

Availability of data and material

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

Code availability

Data was analyzed in GeNIe (academic version).

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Acknowledgments

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

8. References

1. 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.
2. Adedipe, T., Shafiee, M., & Zio, E. (2020). Bayesian network modelling for the wind energy industry: An overview. *Reliability Engineering & System Safety*, 107053.
3. Ale, B., Van Gulijk, C., Hanea, A., Hanea, D., Hudson, P., Lin, P.-H., & Sillem, S. (2014). Towards BBN based risk modelling of process plants. *Safety Science*, 69, 48–56.
4. 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.
5. Anwar, A., Gulzar, A., Sohail, F. Bin, & 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.
6. Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253(1), 1–13.
7. Aven, T. (2017). How some types of risk assessments can support resilience analysis and management. *Reliability Engineering & System Safety*, 167, 536–543.
8. Aven, T. (2019). The call for a shift from risk to resilience: What does it mean? *Risk Analysis*, 39(6), 1196–1203.
9. 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.
10. Bickerstaff, K., & Walker, G. (1999). Clearing the smog? Public responses to air-quality information. *Local Environment*, 4(3), 279–294.

11. 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>
12. 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.
13. Dahlén, M., Granlund, A., & Grenros, M. (2009). The consumer-perceived value of non-traditional media: effects of brand reputation, appropriateness and expense. *Journal of Consumer Marketing*.
14. Dianshu, F., Sovacool, B. K., & Minh Vu, K. (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>
15. 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.
16. Forsyth, T. (2014). Public concerns about transboundary haze: A comparison of Indonesia, Singapore, and Malaysia. *Global Environmental Change*, 25, 76–86.
17. Graff Zivin, J., & Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, 51(3), 689–730.
18. 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.
19. 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.
20. Hosseini, S., & Ivanov, D. (2019). A new resilience measure for supply networks with the ripple effect considerations: A Bayesian network approach. *Annals of Operations Research*, 1–27.
21. 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.
22. 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).
23. 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.
24. 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.
25. 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.
26. 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.

27. Kim, K.-H., Kabir, E., & Kabir, S. (2015). A review on the human health impact of airborne particulate matter. *Environment International*, 74, 136–143.
28. 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.
29. Lindell, M. K., & Prater, C. S. (2002). Risk area residents' perceptions and adoption of seismic hazard adjustments 1. *Journal of Applied Social Psychology*, 32(11), 2377–2392.
30. Liu, T., He, G., & Lau, A. (2018). 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), 325–363.
31. 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*, 1–23.
32. Naeem, M., & Ozuem, W. (2021). Customers' social interactions and panic buying behavior: Insights from social media practices. *Journal of Consumer Behaviour*.
33. Pan, X, Dresner, M. E., Mantin, B., & Zhang, J. (2019). Product Availability, Consumer Stockpiling, and Hurricane Events: Empirical Evidence From a Natural Experiment. In *Consumer stockpiling, and hurricane events: empirical evidence from a natural experiment (January 1, 2019)*.
34. Pan, Xiaodan, 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.
35. 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.
36. 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*, 123676.
37. 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.
38. Rook, D. W., & Fisher, R. J. (1995). Normative influences on impulsive buying behavior. *Journal of Consumer Research*, 22(3), 305–313.
39. S, P. (2011). *Rumours trigger panic buying of salt across mainland, HK*. Post S; Post S.
<http://www.scmp.com/article/741250/rumours-trigger-panic-buying-saltacross-mainland-hk>
40. Samson, A., & Voyer, B. G. (2014). Emergency purchasing situations: Implications for consumer decision-making. *Journal of Economic Psychology*, 44, 21–33.
41. 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*, 1–34.

42. Shaw, R., Kobayashi, K. S. H., & Kobayashi, M. (2004). Linking experience, education, perception and earthquake preparedness. *Disaster Prevention and Management: An International Journal*.
43. Shortridge, J., Aven, T., & Guikema, S. (2017). Risk assessment under deep uncertainty: A methodological comparison. *Reliability Engineering & System Safety*, 159, 12–23.
44. Simsekler, M. C. E., & Qazi, A. (2020). Adoption of a Data-Driven Bayesian Belief Network Investigating Organizational Factors that Influence Patient Safety. *Risk Analysis*.
45. 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>
46. Terpstra, T. (2011). Emotions, Trust, and Perceived Risk: Affective and Cognitive Routes to Flood Preparedness Behavior. *Risk Analysis*, 31(10), 1658–1675. <https://doi.org/10.1111/j.1539-6924.2011.01616.x>
47. 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.
48. 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.
49. 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.
50. 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.
51. 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.
52. 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.
53. 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.
54. 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.
55. 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.
56. 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.

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

58. 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.

59. Zheng, R., Shou, B., & Yang, J. (2020). Supply disruption management under consumer panic buying and social learning effects. *Omega*, 102238.

60. 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.

Figures

Figure 1

Generic BN

Figure 2

Simple DBN over three time periods

Figure 3

Factors influencing the occurrence of protective consumption behavior

Figure 4

Network structure of BN of sales of anti-smog air purifiers

Figure 5

Probability distribution of factors associated with protective consumption behavior

Figure 6

DBN of protective consumption considering time factors

Figure 7

Back propagation impact assessment given the protective consumption in "s4" state

Figure 8

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