Journal of Food Distribution Research
Volume 49, Issue 2

The Use of Time-Series Analysis in Examining Food Safety Issues: The Case of the Peanut Butter Recall

Rafael Bakhtavoryan,[a①] Oral Capps,[b] Victoria Salin,[c] and Aramayis Dallakyan[d]

[a]*Assistant Professor, College of Agricultural Sciences and Natural Resources, Texas A&M University – Commerce, P.O. Box 3011 Commerce, TX 75429-3011 USA*

[b]*Regents Professor and Executive Professor, Department of Agricultural Economics, Texas A&M University, 600 John Kimbrough Blvd, Suite 371C, 2124 TAMU College Station, TX 77843-2124 USA*

[c]*Professor, Department of Agricultural Economics Texas A&M University, 600 John Kimbrough Blvd, Suite 369, 2124 TAMU College Station, TX 77843-2124 USA*

[d]*PhD Student, Department of Agricultural Economics Texas A&M University,600 John Kimbrough Blvd, Suite 326 2124 TAMU College Station, TX 77843-2124 USA*

Abstract

This study presents a time-series analysis of the demand for peanut butter in the wake of the product recall involving Peter Pan and Great Value brands. A 2-lag directed acyclic graphs/Bernanke vector error correction model was estimated using weekly time-series data. The outbreak variable was negatively related to the demand for peanut butter, supporting the hypothesis that foodborne illness reduces consumer demand for a food product category. Hence, time-series models should be complementary to structural/econometric models in examining the impacts of food safety incidents as a check on the robustness of the results.

Keywords: directed acyclic graphs, peanut butter, recall, vector error correction model

①Corresponding author: Tel: (903) 886-5367
Email: Rafael.Bakhtavoryan@tamuc.edu

Introduction

In 2006–2007, the U.S. Centers for Disease Control and Prevention (CDC) and state departments of health investigated a multistate outbreak of salmonellosis. Subsequent investigation concluded that the foodborne illnesses had been caused by the consumption of two brands of peanut butter: Peter Pan and Great Value (a Wal-Mart store brand), both manufactured by ConAgra Foods, Inc., at its Sylvester, Georgia, processing plant (CDC, 2007). As a result, ConAgra ceased the production of peanut butter at this plant, destroyed all affected products in their possession, and voluntarily issued a nationwide recall of Peter Pan and Great Value peanut butter products produced since May 2006 through a news release distributed on February 14, 2007 (CDC, 2007). Following the recall, ConAgra not only redesigned this processing plant but also initiated an unprecedented marketing campaign concerning their Peter Pan brand (Bakhtavoryan, Capps, and Salin, 2014b).

A large body of literature has been dedicated to providing empirical evidence for the impacts of food safety issues on demand for various products (Swartz and Strand, 1981; Smith, van Ravenswaay, and Thompson, 1988; van Ravenswaay and Hoehn, 1991; Burton and Young, 1996; Verbeke and Ward, 2001; Marsh, Schroeder, and Mintert, 2004; Piggott and Marsh, 2004; Pritchett et al., 2007). All of these studies found a statistically significant negative relationship between the food safety incident and demand for the product in question.

Using Nielsen Homescan panels for household purchases from 2006 through 2008, Bakhtavoryan, Capps, and Salin (2012, 2014a,b,c) analyzed the influence of the Peter Pan and Great Value recall on various aspects of demand for peanut butter. In particular, these investigations used structural/econometric models to analyze spillover effects, competition among brands, and structural change in demand for peanut butter brands in the wake of the Peter Pan and Great Value recall. Both a single-equation model and a demand systems model were employed.

The objective of this study is to furnish findings on the impact of the Peter Pan and Great Value product recall on the demand for the peanut butter category using a time-series approach, in contrast to the structural/econometric approach previously employed by Bakhtavoryan, Capps, and Salin (2014b). In this way, a check of the robustness of the results between the two alternative approaches can be made, contributing to the extant literature dealing with examining impacts of food safety incidents. Specifically, our aim is to compare empirical results from the structural/econometric model employed by Bakhtavoryan, Capps, and Salin (2014b) to those generated by the use of a vector error correction (VEC) model. The set of variables considered and the data used in this comparison are the same as in the previous study.

Another contribution of this study is that, to the best of our knowledge, most studies dealing with food safety incidents have employed structural models as opposed to time-series models. This work then adds to the extant literature in this capacity. Further, except for the structural/econometric models employed by Bakhtavoryan, Capps, and Salin (2014a,b), previous research on food safety issues has not used the number of confirmed cases of *Salmonella* reported by the CDC as a measure of the outbreak. Instead, previous research has commonly used various types of media variables to account for the impacts of food safety incidents. Finally, this study contributes by utilizing a modified approach in its application of directed acyclic graphs while dealing with casual relationships among variables when addressing the issue of contemporaneous correlations for generating representative impulse-response functions and forecast error variance decompositions.

Literature Review

Prior studies have paid much attention to the problem of consumer response to food safety issues by employing various econometric approaches, including single-equation structural models and demand system models. In particular, Swartz and Strand (1981) investigated the impact of information concerning oyster contamination due to kepone (an insecticide) on the demand for shucked oysters in Baltimore, Maryland. They estimated a single-equation structural model with second-order and 4-lag polynomial distributed lag (PDL) structure, using biweekly data from 1973–1976. The variable reflecting the negative information was constructed based on articles from the four major Baltimore and Washington newspapers. The estimation results showed that the lags of the media variable were statistically significant, negatively impacting the consumption of oysters.

In their study, Smith, van Ravenswaay, and Thompson (1988) analyzed the response of fluid milk sales to negative newspaper coverage related to the heptachlor (an insecticide) contamination of fresh fluid milk in Oahu, Hawaii, by applying a single-equation structural model with second-degree PDL specification and 3 lags. Their study used monthly time-series data from January 1977 to June 1983. A negative media variable was developed using newspaper articles regarding the food contamination incident from two major Honolulu newspapers during the period that contained negative information on milk quality, the level of government protection, and the integrity of milk processors in handling the incident problem. The estimation results suggested a statistically significant negative relationship between the current and lagged negative media variables and fluid milk sales.

Van Ravenswaay and Hoehn (1991) studied the influence of Alar (a carcinogenic chemical sprayed on fruit) on the demand for apples by estimating a single-equation PDL model with 3 lags and employing monthly data from January 1980 to July 1989. The risk information variable concerning Alar was constructed based on the monthly number of articles in *The New York Times*. The empirical findings of the study indicated that the current and the third lag of the risk information variable were significant and negatively impacted the demand for apples.

Burton and Young (1996) investigated the effects of bovine spongiform encephalopathy (BSE) on the demand for beef and other meat products in Great Britain by applying a dynamic Almost Ideal Demand System (AIDS) model and using quarterly data from January 1961 to March 1993. They developed a variable capturing consumer awareness of BSE based on the number of published newspaper articles that contained information on BSE. Their empirical results showed that consumer awareness of BSE resulted in a loss in market shares of beef producers both in the short run and in the long run.

Verbeke and Ward (2001) analyzed consumer response to the negative public media coverage regarding food safety issues associated with fresh meat in Belgium. Their study estimated a linear approximation of the AIDS model for beef and veal, pork and meat mixtures, and poultry, employing panel data on monthly observations from January 1995 to December 1998. The mass media index, which was anticipated to capture consumer awareness of meat-related health issues, was developed by subtracting the number of positive TV reports from the number of negative TV reports associated with the effects of meat consumption on human health. The empirical findings showed that the impact of adverse publicity, primarily concerning BSE, was statistically significant and had a negative influence on the consumption of beef and veal and a positive influence on the consumption of pork and meat mixtures.

Marsh, Schroeder, and Mintert (2004) studied the effects of meat product recall events on the demand for beef, pork, poultry, and other products in the United States by estimating the absolute price version of the Rotterdam model using quarterly data on beef, pork, chicken, and turkey from 1982–1998. Two measures of meat product recalls were constructed using Food Safety Inspection Service (FSIS) reports and media reports from the popular press. The empirical results revealed that, unlike newspaper reports, FSIS reports on recall events negatively influenced the demand for beef and pork and positively influenced the demand for poultry and other products.

Piggott and Marsh (2004) estimated a Generalized Almost Ideal Demand System model to evaluate the effects of public information concerning food safety issues related to beef, pork, and poultry reported in the media on meat demand. The study employed quarterly meat data from the first quarter of 1982 through the third quarter of 1999. They developed food safety indices for each meat type by aggregating the number of newspaper articles regarding food safety issues. The estimation results established a statistically significant relationship between consumer demand and contemporaneous media coverage of health hazards.

Pritchett et al. (2007) evaluated consumer demand for meat cuts of beef, pork, and chicken in light of the announcements associated with BSE in Canada and the United States by estimating the AIDS model and using a dataset derived from monthly retail scanner data for 191 meat products sold in U.S. retail grocery stores from January 2001 through February 2005. They constructed an information variable accounting for the influence of media coverage based on the reported articles. The estimation results indicated that the BSE events negatively affected the demand for ground beef and chuck roasts and positively affected the demand for center-cut pork chops.

The 2007 Peter Pan and Great Value peanut butter recall has been analyzed by prior studies using single-equation structural model and demand systems. In particular, Bakhtavoryan, Capps, and Salin (2012) used the Barten synthetic model to estimate the pre- and post-recall demand elasticities for a statistical comparison using weekly observations from Nielsen Homescan panel data on household purchases of major peanut butter brands from January 2006 through December 2008. The estimation results revealed that demand elasticities statistically increased across the two recall periods, thus contributing to a structural change in the demand for peanut butter brands.

Using the same dataset, Bakhtavoryan, Capps and Salin (2014a) estimated the Barten synthetic model with a PDL specification applied to the variable measuring the impact of the recall to ascertain possible spillover effects among major peanut butter brands in the wake of the Peter Pan and Great Value peanut butter recall. They constructed the recall variable based on the number of confirmed cases of *Salmonella* due to the consumption of contaminated peanut butter reported by the CDC. The empirical findings revealed that the demand for Peter Pan was negatively impacted by the recall, while the demand for Jif enjoyed positive spillover effects as a result of the recall.

In another study by Bakhtavoryan, Capps, and Salin (2014b), a single-equation structural demand model was estimated to study the influence of the 2007 Peter Pan and Great Value peanut butter recall on the demand for peanut butter at the product-category level. A second-degree and a 3-lag PDL structure were imposed on the variable capturing the recall effects and constructed using the number of confirmed cases of *Salmonella* from the consumption of contaminated peanut butter. Contrary to expectations, the impact of the recall variable on the demand for peanut butter was found to be positive, suggesting that the recall had demand-enhancing effects for peanut butter at the product-category level. This unexpected finding was explained by households' restocking behavior, in which jars of tainted peanut butter were substituted with other brands, leading to an overall increase in the consumption of peanut butter.

Finally, Bakhtavoryan, Capps, and Salin (2014c) estimated a multinomial logit models to identify household socioeconomic factors that influenced three consumption patterns associated with the Peter Pan peanut butter. The three consumption patterns were buying Peter Pan in the pre-recall period only, buying Peter Pan in the post-recall period only, and buying Peter Pan in both the pre- and post-recall periods. The estimation results revealed that characteristics such as employment status of the household head, region of residence, race, ethnicity, age and presence of children in the household were statistically significant drivers associated with the actions taken by households in light of the Peter Pan recall. In the same study, findings from the Heckman sample selection model indicated that the change in price, region of residence, race, age and presence of children in the household, and household size were key drivers impacting the change in quantity of Peter Pan purchased across the pre- and the post-recall periods.

The present analysis is similar to prior studies reviewed in that it also attempts to evaluate the impact of a food safety issue on the demand for a product. However, the distinct feature of the present analysis is reflected in its use of a time-series approach complemented with the analysis of the directed acyclic graphs and its inclusion of the number of confirmed cases of *Salmonella* reported by the CDC as a measure of the outbreak.

Methodology

In a single-equation model, constructed based on economic theory, it is implicitly assumed that there is a unidirectional cause-and-effect relationship between the dependent variable and the set of independent or explanatory variables, with the causal flow implying that the set of independent variables is the cause and the dependent variable is the effect. But sometimes there are cases when a unidirectional relationship is not viable. An advantage of estimating the equations as a system rather than individually is the resulting improvement in efficiency, which is obtained because error terms are typically correlated among equations. In the present study, as an initial step, a system of equations in the form of the vector autoregression (VAR) model was estimated, in contrast to the single-equation structural model in Bakhtavoryan, Capps, and Salin (2014b).

Sims (1980) developed and introduced the VAR model, which—along with its variants—has become popular in applied time-series analysis (Brandt and Bessler, 1984; Bessler, 1984a; Awokuse and Bessler, 2003; Capps, Bessler, and Williams, 2016). One reason for the acceptance of the VAR approach is that the identification conditions of structural-equation modeling are relaxed. In a single equation or system of structural equations, the analyst must specify variables as exogenous or endogenous. To estimate the parameters of the system, either exact identification or over-identification conditions have to be fulfilled. The identification conditions are often fulfilled by specifying particular exogenous variables to appear in some equations, while they are omitted from other equations (Gujarati, 2004). This approach was not deemed appropriate by Sims, who maintained that there should not be any postulated distinction between endogenous and exogenous variables and that all variables should be treated equally (Sims, 1980).

Subsequent extensions of time-series methods took into account that, in some situations, variables share common stochastic trends; when they do, they are said to be cointegrated (Granger, 1981; Engle and Granger, 1987). Once a system of variables is determined to have cointegrating relationships, Lütkepohl and Kratzig (2004) suggested considering a specific parameterization supporting the analysis of the cointegration structure, leading to VEC models. The VEC model is sensitive to autocorrelation of the residuals, which may arise during the optimal lag selection procedure (Phoong, Ismail, and Sek, 2014). The residual autocorrelation problem applies to VAR models too. However, the VEC model has the additional imposed restriction that the variances and covariances of the error-correction terms are assumed to be constant (Phoong, Ismail, and Sek, 2014). Just as in a structural single-equation model, VAR models are developed by including variables that are suggested by the economic theory.

In our model, the variables included in the VAR (and subsequently the VEC model) are based on economic theory, as are the variables incorporated in the corresponding structural single-equation model. In particular, consumer theory hypothesizes that quantity demanded of a product is influenced by its own price, price of a substitute or a complement good, and consumer income. Hence, in the final VEC model, the quantity demanded of peanut butter was hypothesized to be affected by the price of peanut butter, the price of jelly as a complement good, and consumer income. Coupons are price reductions and, as such, impact quantity demanded of a product. Hence, in the final VEC model, a variable associated with coupon values for peanut butter was incorporated. Per theory, negative information is expected to decrease quantity demanded of a product. In our case, the outbreak variable was incorporated in the VEC model to capture the effects associated with the recall on the quantity demanded of peanut butter. A dummy variable was included in the VEC model to capture the possible structural change in the demand for peanut butter in the wake of the recall. While the theory does not say anything about the seasonality in the consumption of a product, quarterly dummy variables were incorporated in the VEC model to account for potential seasonality in the demand for peanut butter.

The initial step in accomplishing the objective of this study is to specify a VAR model:

$$\mathbf{X_t} = \boldsymbol{\beta} + \sum_{i=1}^{k} \mathbf{A_i} X_{t-i} + \boldsymbol{\varepsilon_t}, \quad (1)$$

where $\mathbf{X_t}$ is a vector of series corresponding to quantity of peanut butter purchased, real price of peanut butter, real price of jelly, coupon redemption for peanut butter, real income, and the number of confirmed cases of *Salmonella* reported by the CDC (i.e., the outbreak variable). Additionally, $\boldsymbol{\beta}$ is a drift vector, $\mathbf{A_i}$ is a coefficient matrix, $\boldsymbol{\varepsilon_t}$ is a vector of stochastic white noise error terms, *i* represents lags, and *k* is the maximum length of lag. The model was augmented by including seasonal dummies and a dummy variable to control for a structural shift in the demand for peanut butter. In this analysis, a natural logarithm transformation was applied to all the variables except for the number of confirmed cases of *Salmonella* reported by the CDC. For the outbreak variable, a square root transformation was applied to capture diminishing marginal returns associated with the possible nonlinear relationship between the quantity of peanut butter purchased and the outbreak variable (Capps, Bessler, and Williams, 2016).

Per the law of demand, a negative relationship was anticipated between the quantity purchased of peanut butter and own price (Rimal, Fletcher, and Deodhar, 2001). According to economic theory, a negative relationship was anticipated between the quantity purchased of peanut butter and the price of jelly because of the complementary relationship associated with these products (He et al., 2004; Smith, Rossi, and Allenby, 2016; Caine-Bish and Scheule, 2007). A positive relationship was expected between coupon values and the quantity purchased of peanut butter. Peanut butter was hypothesized to be a normal good (Rimal, Fletcher, and Deodhar, 2001). As such, a positive relationship was expected between the quantity purchased of peanut butter and income. Finally, in keeping with economic theory and empirical studies of food safety incidents (Duan, 2014), a negative relationship between the quantity purchased of peanut butter and the outbreak variable was expected.

Before estimating the model, a few practical issues need to be addressed. Augmented Dickey–Fuller (ADF) tests have to be carried out to test for stationarity in the series. If the respective variables are not stationary, then it is necessary to construct a first or second difference to render them stationary. Also, the optimal lag length must be determined based on statistical criteria, such as the Akaike Information Criterion (AIC) or the Schwarz Information Criterion (SIC). Finally, Johansen's cointegrating rank test must be carried out to identify possible cointegrating equations (Johansen, 1988; Juselius, 2006). If there is at least one cointegrating equation, a VEC model is appropriate. A VEC model in first differences with order of $k - 1$ can be written as

$$\Delta X_t = \boldsymbol{\alpha} + \sum_{i=1}^{k} \Gamma_{\boldsymbol{i}} \Delta X_{t-i} + \boldsymbol{\Pi} X_{t-1} + u_t, \quad (2)$$

where $\boldsymbol{\alpha}$ is a drift vector, $\boldsymbol{\Gamma_i}$ is a short-run coefficient matrix, $\boldsymbol{\Pi}$ is a long-run coefficient matrix, $\boldsymbol{\Pi} X_{t-1}$ is the error-correction term, and u_t is the error term.

Directed Acyclic Graphs

In general, VAR and VEC models say little about contemporaneous time correlation among variables. However, ignoring causal orderings among the respective variables in the VEC model in contemporaneous time may not produce representative impulse-response simulations and forecast error variance (FEV) decompositions (Bessler, 1984b; Sims, 1980).

The econometric literature dealing with the use of VAR and VEC models has traditionally accounted for contemporaneous correlations in three ways. The first is the use of Choleski factorization, in which contemporaneous correlations are established by imposing theory-based and recursive causal ordering on the variance/covariance matrix of the error terms (Bessler, 1984b; Sims, 1980; Bessler and Akleman, 1998). The problem with this approach is that situations usually are not recursive and, in general, results from impulse responses and FEV decompositions vary noticeably with the ordering chosen by Choleski factorization. The second approach rests on the use of the structural VAR method (Bernanke, 1986), in which prior notions of evidentially based and/or theoretically grounded, contemporaneously causal orderings may be imposed on the variables that make up the VAR (Bessler and Akleman, 1998). The problem here is that the true contemporaneous orderings that analysts claim to know may not be correct. The third approach developed by Pesaran and Shin (1998), a generalized impulse-response analysis for VAR models (and for cointegration or VEC models as well), avoids orthogonalization of shocks and therefore generates order-invariant results (Babula, Bessler, and Payne, 2004). The use of the third approach requires caution (Doan, 2002) because of difficulty in interpreting impulses from highly correlated shocks within a nonorthogonalized setting.

In this study, the Bessler and Akleman (1998) procedure was used to optimally choose a set of causal relations among six variables and then impose the evidentially supported causal relations on a Bernanke-type structural VAR. In following this procedure, an attempt is made to avoid choosing arbitrarily among competing but otherwise theoretically consistent sets of contemporaneous orderings inherent in Choleski-ordered or Bernanke structural VARs. This is accomplished with the help of directed acyclic graphs and the PC algorithm.[1] Pioneers in applying a graph-theoretical approach to the problem of determining the order of structural VAR were Swanson and Granger (1997), Bessler and Loper (2001), Bessler and Lee (2002), Demiralp and Hoover (2003), and Hoover, Demiralp, and Perez (2009). To address the issues associated with the VAR and VEC models in assessing the contemporaneous time correlation among variables, the present analysis is complemented with directed acyclic graphs and the PC algorithm, explained in the next section.

Directed Graphs and the PC Algorithm

A graph is a data structure, $\mathcal{G}$, consisting of a set of nodes and a set of edges. A pair of nodes $\mathrm{X_i}, \mathrm{X_j}$ can be connected by a directed edge, $\mathrm{X_i} \rightarrow \mathrm{X_j}$, or an undirected edge, $\mathrm{X_i} - \mathrm{X_j}$. Thus, the set of edges, ξ, is a set of pairs in which each pair is one of $\mathrm{X_i} \rightarrow \mathrm{X_j}, \mathrm{X_i} \leftarrow \mathrm{X_j}$, or $\mathrm{X_i} - \mathrm{X_j}$. Whenever $\mathrm{X_i} \rightarrow \mathrm{X_j} \epsilon\ \xi$, we call $\mathrm{X_j}$ child of $\mathrm{X_i}$ and $\mathrm{X_i}$ parent of $\mathrm{X_j}$. We say that $\mathrm{X_1}, \cdots \mathrm{X_k}$ form a path in graph $\mathcal{G}$ if, for every $i = 1, \dots, k-1$, we have that either $\mathrm{X_i} \rightarrow \mathrm{X_{i+1}}$, or $\mathrm{X_i} - \mathrm{X_j}$. A path is directed, if, for at least one i, we have $\mathrm{X_i} \rightarrow X_{i+1}$. X is an ancestor of Y in $\mathcal{G}$ and Y is a descendant of X if there exists a directed path $\mathrm{X_1} \cdots \mathrm{X_k}$ with $\mathrm{X_1} = \mathrm{X}$ and $\mathrm{X_k} = \mathrm{Y}$. A cycle in $\mathcal{G}$ is a directed path $\mathrm{X_1} \dots \mathrm{X_k}$, where $\mathrm{X_1} = \mathrm{X_k}$. A graph is acyclic if it contains no cycles. We call these graphs directed acyclic graphs (DAGs). DAGs are the fundamental graphical representation that underlies Bayesian Networks. A Bayesian Network structure $\mathcal{G}$ is a DAG whose nodes represent random variables $\mathrm{X_1} \dots \mathrm{X_n}$. Denote $\mathrm{Pa}^{\mathcal{G}}_{\mathrm{X_i}}$ the parents of $\mathrm{X_i}$ in $\mathcal{G}$, and $\mathrm{NonDescendants}_{\mathrm{X_i}}$ the

[1] PC stands for the initials of its inventors: Peter Spirtes and Clark Glymour.

variables in the graph that are not descendants of $\mathrm{X_i}$. Then, $\mathcal{G}$ encodes the following set of conditional independence assumptions, called the local independencies:

(3) $$(\mathrm{X_i} \perp \mathrm{NonDescendants_{X_i}} | \mathrm{Pa}^{\mathcal{G}}_{\mathrm{X_i}})$$

or

(4) $$\mathrm{P}(\mathrm{X_1}, \dots, \mathrm{X_n}) = \prod_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{P}\,(\mathrm{X_i}||\mathrm{Pa}^{\mathcal{G}}_{\mathrm{X_i}}).$$

Basically, equation (3) says that each node $\mathrm{X_i}$ is conditionally independent of its nondescendant given its parents. That is, the other information is irrelevant as long as we can identify the parents of the node.

Equation (4) is a direct consequence of an assumption about equation (3). In other words, since the joint distribution can always be written as a product of conditional probabilities, $\mathrm{P}(\mathrm{X_1}, \dots, \mathrm{X_n}) = \mathrm{P}(\mathrm{X_1})P(X_2|X_1) \dots . P(X_n|X_1 \dots X_{n-1})$, then using the independence assumption on equation (3) and that the graph is acyclic (i.e., there exists at least one node which does not have parents), equation (4) holds true. Equations (3) and (4) are the fundamental ideas behind constructing the DAGs and d-separation (Pearl 1986). Geiger, Verma, and Pearl (1990) show the soundness and completeness of d-separation. By soundness they mean that any independence reported by d-separation is satisfied by the underlying distribution. The completeness of d-separation requires the notion of faithfulness. A distribution is faithful to $\mathcal{G}$ if any independence in distribution is reflected in the d-separation properties of the graph. It can be shown that faithfulness holds for almost all distributions that satisfy equation (4) over $\mathcal{G}$. In other words, for almost all possible choices of conditional probability distributions for the variables, d-separation precisely characterizes the independencies of the underlying distribution (Koller and Friedman, 2010).

Having these tools available, Spirtes, Glymour, and Scheines (2000) incorporated the notion of d-separation into an algorithm (PC algorithm) for building DAGs. The PC algorithm is an ordered set of commands that begins with a set of relationships among variables (in our case innovations [i.e., error terms] from each VAR equation) and proceeds stepwise to remove edges between variables so as to direct causal flow in contemporaneous time (Spirtes, Glymour, and Scheines, 2000; Bessler and Akleman, 1998). The goal is to impose a directed edge among sets of variables $\{\mathrm{X_1}, \mathrm{X_2}, \mathrm{X_3}\}$ in a vertex set (variable set) $\mathcal{X}$: $\mathrm{X_1} \rightarrow \mathrm{X_2} \rightarrow \mathrm{X_3}$, $\mathrm{X_1} \leftarrow \mathrm{X_2} \leftarrow \mathrm{X_3}$, $\mathrm{X_1} \rightarrow \mathrm{X_2} \leftarrow \mathrm{X_3}$.

The algorithm begins with a complete, undirected graph that places an undirected edge between every variable in the system (every variable in graph $\mathcal{G}$ vertex set $\mathcal{X}$). Edges between variables are removed sequentially on the basis of zero correlations or zero partial (conditional) correlations. These conditioning variables on removed edges between variables comprise Bessler and Akleman's (1998) "sepset" of the variables whose edge has been removed.

Data

This study employs weekly time-series data from the Nielsen Homescan Panel on quantities purchased, prices, and coupons from July 26, 2006, through December 30, 2018, for a total of 127 weekly observations.[2] In addition, the dataset included a variable measuring income and a variable measuring the impact of the recall. Table 1 reports descriptive statistics on the variables used in the analysis.

Table 1. Descriptive Statistics (N = 127)

Variable	Description	Units	Mean	Std. Dev.
Quantity_PB	Quantity of peanut butter	oz	33.54	1.15
Price_PB	Real unit value of peanut butter	cents/oz	5.01	0.25
Price_Jelly	Real unit value of jelly	cents/oz	3.21	0.24
Coupon_PB	Real coupon of peanut butter	cents	5.42	2.91
Income	Weekly real income	dollars	614.18	8.46
CDC_cases	No. of CDC-confirmed cases	cases	3.79	7.92

Notes: Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

The quantity of peanut butter purchased was calculated by first summing the weekly total ounces of peanut butter brands across households and then by dividing that sum by the number of unique households that actually purchased peanut butter in any given week. Unit values were used as proxies for prices, which were not directly observed. The unit values for peanut butter and jelly were computed by dividing total expenditures by total ounces for each week. The coupon variable for peanut butter was constructed by first summing weekly values of coupons used and then dividing this sum by the number of unique households to express the variable on a per household basis. Weekly interpolations of real disposable personal income reported by the U.S. Department of Commerce (2011) were used as a proxy for household income.

To adjust for inflation, all unit values, coupon values, and income were deflated using the consumer price index (CPI) available from the Bureau of Labor Statistics (BLS) of the U.S. Department of Labor. The base-year CPI corresponded to the period 1982–1984. The variable accounting for the influence of the recall (hereafter referred to as the outbreak variable) was constructed based on the weekly number of CDC-confirmed cases of *Salmonella* Tennessee infection due to the consumption of tainted peanut butter (CDC, 2007). Consistent with previous research, quarterly dummy variables were included in the model to capture potential seasonality in the demand for peanut butter (Rimal, Fletcher, and Deodhar, 2001), utilizing the fourth quarter as the base or reference category. Finally, a potential permanent structural change in the demand for peanut butter was captured by a dummy variable that assumed a value of 0 before the issuance of the recall and a value of 1 afterward.

[2] The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

Empirical Results

The presence of stationarity in the historical series was tested with the ADF test. Table 2 presents the results from the ADF tests at the 5% significance level.[3] The null hypothesis of nonstationarity was rejected for the quantity purchased of peanut butter and the coupon value of peanut butter. However, the remaining variables were nonstationary.

Table 2. Augmented Dickey–Fuller Tests for Stationarity Regarding the Natural Logarithms of the Respective Variables in the Time-Series Model

Variable	Test Statistic	Decision (at 5% significance level)
Quantity_PB	−6.405	Reject nonstationarity
Price_PB	−3.179	Fail to reject nonstationarity
Price_Jelly	−3.113	Fail to reject nonstationarity
Coupon_PB	−3.803	Reject nonstationarity
Income	−1.638	Fail to reject nonstationarity
sqrt_CDC_cases	−1.840	Fail to reject nonstationarity

Notes: Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Table 3 reports the results from the ADF tests for the first differences of all the series at the 5% significance level. As shown in Table 3, all the variables were stationary in first differences except for the income variable, which was stationary using second differences.

Table 3. Augmented Dickey–Fuller Test for the First Differences of the Natural Logarithms of the Series

Variable	Test Statistic	Decision (at 5% significance level)
d_Quantity_PB	−10.558	Reject nonstationarity
d_Price_PB	−8.757	Reject nonstationarity
d_Price_Jelly	−9.585	Reject nonstationarity
d_Coupon_PB	−9.987	Reject nonstationarity
d_Income	−1.473	Fail to reject nonstationarity
d_sqrt_CDC_cases	−10.505	Reject nonstationarity

Notes: d_ indicates first differences. Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

The appropriate number of lags to be included in the model was determined based on AIC and SIC metrics (Table 4). Based on the AIC and SIC, the appropriate lag length was 2 lags because the AIC and SIC values were minimized at lag two.

Johansen's (1995) cointegrating rank tests were performed: A sequence of trace tests and maximum eigenvalue tests were carried out, producing the optimal number of cointegrating

[3] Results from the ADF tests and ADF tests for the first differences were also supported by results from the KPSS tests.

Table 4. Akaike and Schwarz Information Criteria for the Appropriate Number of Lags Selection

Lag	AIC	SIC
0	−11.2969	−11.159
1	−21.6462	−20.6809
2	−23.5858*	−21.7931*
3	−23.5076	−20.8875
4	−23.2668	−19.8193
5	−23.2341	−18.9592

Notes: Single asterisk (*) indicates the appropriate lag length. Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

equations. Table 5 presents the results of the trace tests and the maximum eigenvalue tests at the 5% significance level. The null hypothesis of zero cointegrating vectors was rejected. However, the null hypothesis of two cointegrating vectors was not rejected.

Table 5. Johansen's Cointegrating Rank Tests

Maximum Rank, r	Trace Statistic	5% Critical Value	Eigenvalue	Maximum Eigenvalue	5% Critical Value
= 0	122.4639	94.15	.	52.9442	39.37
≤ 1	69.5197	68.52	0.34528	37.4055	33.46
≤ 2	32.1142*	47.21	0.25862	14.4832*	27.07
≤ 3	17.6310	29.68	0.10940	9.5950	20.97
≤ 4	8.0360	15.41	0.07389	6.3566	14.07
≤ 5	1.6794	3.76	0.04958	1.6794	3.76

Notes: Single asterisk (*) indicates the cointegrating rank. Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Next, the VEC model parameters were estimated. Basically, the main interest lies in the equation with the dependent variable related to first differences of the quantity purchased of peanut butter. The STATA 12 software package was used to perform the estimation. Table 6 presents the results from the VEC estimation for the equation pertaining to the quantity purchased of peanut butter at the 5% significance level.

The R^2 was 0.5149, indicative of a reasonably good fit. Several coefficients were significantly different from zero. In particular, the estimated coefficient of the price of peanut butter lagged two periods was negative, as anticipated, and was statistically different from zero. This result was consistent with Bakhtavoryan, Capps, and Salin (2014b). The estimated coefficients associated with the first and the second lags of the price of jelly were negative and statistically different from zero, as expected. However, Bakhtavoryan, Capps, and Salin (2014b) did not find this variable to be statistically significant. In addition, the estimated coefficient of the second lag of income was positive and statistically significant, as expected. This finding compared favorably with that of Bakhtavoryan, Capps, and Salin (2014b). Additionally, the estimated

Table 6. Estimation Results for the Quantity Purchased of Peanut Butter Equation from the Vector Error Correction Model, $N = 123$

	Coefficient	*p* Value
_ce1		
L1	−0.737*	0.000
_ce2		
L1	0.050*	0.000
ln_Quantity_PB		
LD	−0.010	0.934
L2D	−0.065	0.523
ln_Price_PB		
LD	−0.051	0.724
L2D	−0.282*	0.049
ln_Price_Jelly		
LD	−0.225*	0.001
L2D	−0.131*	0.033
ln_Coupon_PB		
LD	−0.001	0.909
L2D	−0.005	0.554
ln_Income		
LD2	−0.897	0.919
L2D2	19.128*	0.030
sqrt_CDC_cases		
LD	−0.009	0.089
L2D	−0.009	0.077
Q1	−0.002	0.807
Q2	−0.019*	0.019
Q3	−0.012	0.160
DUMMY	−0.024*	0.005
Constant	0.024*	0.004

Notes: Single asterisk (*) indicates statistical significance at the 5% level. Log-likelihood = 1,573.709. L1 indicates that the variable is lagged one period, LD indicates lagged first differences, and _ce corresponds to the respective error-correction terms. Q1–Q3 are seasonal dummies and DUMMY is a dummy variable controlling for the structural shift in the demand for peanut butter. The estimation results of the remaining equations are available from authors upon request. Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

coefficient of the second-quarter seasonal dummy variable was negative and significantly different from zero, in accordance with the estimation results by Bakhtavoryan, Capps, and Salin (2014b), who also found seasonality to be a statistically significant factor. Consistent with the previous study, the estimated coefficient associated with the dummy variable was negative and statistically significant, indicating a structural change in the demand for peanut butter. Moreover, as in the previous study, the estimated coefficients associated with the coupon variable were statistically insignificant.

Based on one-tailed tests, the estimated coefficients of the first and second lags of the outbreak variable were negative and statistically significant, supporting the hypothesis of negative impacts associated with food safety incidents. However, this result was at odds with the finding by Bakhtavoryan, Capps, and Salin (2014b) that the parameter estimates associated with the outbreak variable were positive, implying that the outbreak positively influenced the quantity purchased of peanut butter. Differences between the time-series VEC model and the structural/econometric model likely account for the difference in the estimation results in regard to the outbreak variable. This discrepancy provides empirical evidence that alternative model specifications may generate nonrobust results. As such, the use of time-series models as well as conventional structural/econometric models is recommended when analyzing food safety issues.

DAG Application

Before estimating and discussing impulse-response functions and FEV decompositions, it is necessary to illustrate the application of the DAGs to find how the six variables were ordered in contemporaneous time using the R package pcalg (Kalisch et al., 2012). The starting point is Figure 1, the complete undirected graph of all possible edges among the six variables. Figure 2 provides the edges that the algorithm suggested as statistically significant at the 10% level.

Contemporaneous causal ordering was discovered in several steps. First, the algorithm based on unconditional correlations eliminated all statistically zero edges and retained those that were statistically nonzero (Spirtes, Glymour, and Scheines, 2000). Then, the algorithm checked all the remaining conditional correlations and retained the ones that were statistically nonzero. If the edges were fully one-side directed,[4] a unique set of correlations could have been imposed on Bessler and Akleman's DAG/Bernanke VAR model. However, per Figure 2, one edge is bi-directional, which indicated that there existed systems of observationally equivalent contemporaneous causality relationships. In that case, there was a need to find "the best" Bayesian Network that represented the data.

Although finding the best Bayesian Network structures is NP-hard (Chickering, Meek, and Heckerman, 2003),[5] feasible techniques exist for small networks (e.g., Singh and Moore, 2005;

[4] That is, it is not true that $X_i \rightarrow X_j \in \xi$ and $X_i \leftarrow X_j \in \xi$.

[5] In computational complexity theory NP-hardness (nondeterministic polynomial-time hardness) is the property that defines a class of problems. Formally, a decision problem H is NP-hard when for every problem L in NP, there is a polynomial-time reduction from L to H. NP (non-deterministic polynomial time) is a complexity class used to describe certain types of decision problems. For more information, see the work by Cormen et al. (2009).

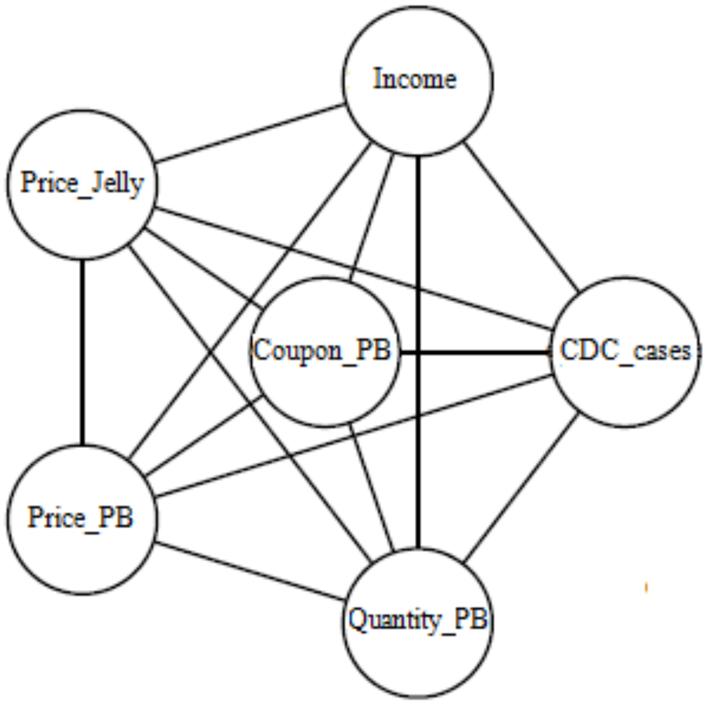

Figure 1. Complete Undirected Graph on Innovations from the VEC Model
Notes: Natural logarithmic transformation was used on all variables, with the square root transformation applied to the CDC_cases variable.

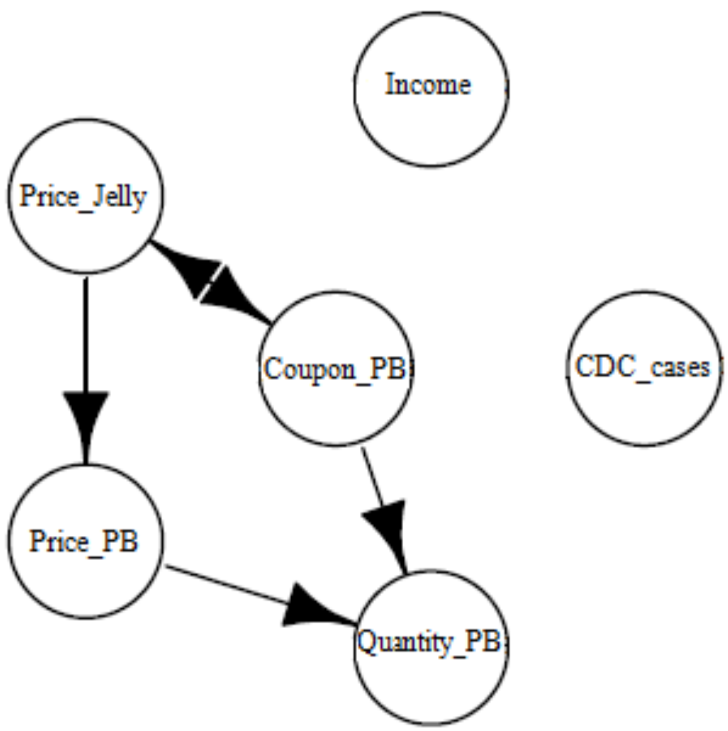

Figure 2. Generated DAG on Innovations from the VEC Model
Notes: Natural logarithmic transformation was used on all variables, with the square root transformation applied to the CDC_cases variable. Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Silander and Myllymaki, 2006; Haigh and Bessler, 2004). Haigh and Bessler (2004) modified and applied Schwarz's loss metric to the alternative systems of causality and then chose the system of causality that minimizes the Schwartz metric. This study followed the method suggested by Silander and Myllymäki (2006), rather than the Haigh and Bessler approach, to find the best Bayesian Network structure. To use the Silander and Myllymäki method, the scoring functions have to be modular (i.e., given the data, the score of a Bayesian Network structure G = (G_1, …, G_n) for variables X = (1, …, n) must be decomposable to local scores:

$$\text{score}(G) = \sum_{i=1}^{n} \text{score}_i(G_i). \tag{5}$$

The score of the network was the sum of the local scores that depend only on the conditional probability for one variable and its parents. Most of the known scores, such as SIC and AIC, are decomposable (Chickering, 1995). By measuring the local scoring function, the goodness of the parents of X_i is found. This idea naturally leads to finding the best parents for a variable X_i in any given parent candidate set C:

$$g_i^*(C) = \arg_{g \subseteq C} \max \text{score}_i(g). \tag{6}$$

The Bayesian Information criterion (BIC) is used in this study as a scoring rule, following Silander and Myllymäki (2006) and using the method discussed above. Based on Figure 2, there existed two possible relationships in the Bernanke structural VAR to form the DAG/Bernanke VAR model. Therefore, two local scores had to be estimated:

1. ln(Price_Jelly) → ln(Coupon_PB) (i.e., the ln(Price_Jelly) variable is the parent for the ln(Coupon_PB) variable);
2. ln(Coupon_PB) → ln(Price_Jelly) (i.e., the ln(Coupon_PB) variable is the parent for the ln(Price_Jelly) variable).

A choice had to be made between these two possible and competing systems of causal relations based on the provided maximum value. The highest score was provided by the option in which ln(Coupon_PB) was the parent for ln(Price_Jelly) (Table 7). Imposing these relationships, resolved the problem of contemporaneous correlation. Figure 3 shows the final DAG after this imposition.

Table 7. Two Alternative (Observationally Equivalent) Systems of Contemporaneous Causal Relations

Type	**System 1**	**System 2**
Parent	ln(Price_Jelly)	ln(Coupon_PB)
Child	ln(Coupon_PB)	ln(Price_Jelly)
Score Value	80.60	300.61

Notes: Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Having addressed the issue of contemporaneous correlation, dynamic interrelationships among the variables in the VEC model can be analyzed using methods of innovation accounting such as FEV decompositions and impulse-response functions. FEV decompositions assist in quantifying the importance of each shock in explaining the variation in each variable in the model. This

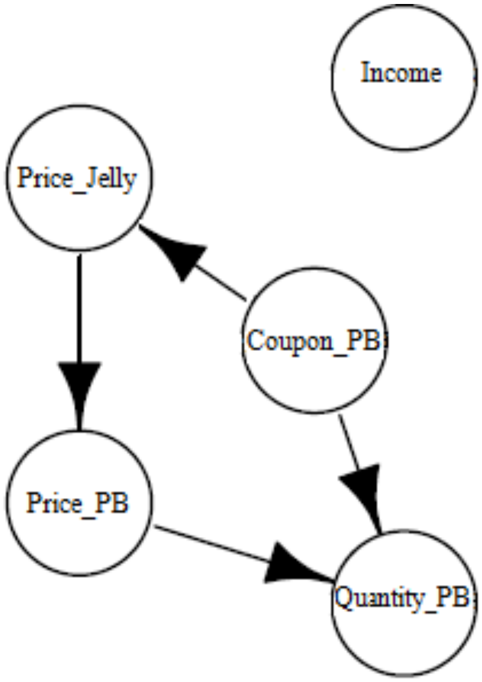

Figure 3. Final DAG Based on Innovations from the VEC Model
Notes: Natural logarithmic transformation was used on all variables, with the square root transformation applied to the CDC_cases variable. Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

metric was calculated as a fraction of the FEV of each variable at different forecast horizons. Impulse-response functions showed the impacts of unit innovations in a particular variable on all variables in the model over time.

Table 8 gives the FEV decomposition from the 2-lag VEC model for the quantity of the peanut butter purchased for 1-, 8-, 16-, 26-, and 52-week forecast horizons. If an innovation of a particular variable accounted for a high percentage of the FEV, then it was considered to be a determinant of the quantity purchased of peanut butter.

Table 8. Forecast Error Variance Decomposition for the Quantity of Peanut Butter Purchased in Percentages

Horizon in Weeks	Quantity_PB	Price_PB	Price_Jelly	Coupon_PB	Income	CDC_cases
1	72.95	24.08	1.58	1.39	0.00	0.00
8	59.46	21.32	7.79	6.65	0.79	3.98
16	57.87	20.85	7.77	6.77	1.87	4.87
26	57.16	20.61	7.77	6.70	2.65	5.09
52	56.64	20.38	7.84	6.63	3.41	5.09

Notes: Rows do not add up to 100% due to rounding errors. Natural logarithmic transformation was used on all variables, with the square root transformation applied to the CDC_cases variable. Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

About 73% and 24% of the 1-week FEV of the quantity purchased of peanut butter were accounted for by innovations in the quantity of peanut butter purchased and the real price of peanut butter, respectively. For longer-term horizons, approximately 57% and 20% of the error variance was accounted for by innovations in the quantity purchased of peanut butter and the real price of peanut butter, respectively. For the 1-week horizon, innovations in the real price of jelly and real coupon values contributed less than 2% to the FEV of the quantity purchased of peanut butter. At the same time, innovations in the real price of jelly and real coupon values contributed about 8% and 7%, respectively, to the FEV of the quantity purchased of peanut butter for longer-term horizons. Innovations in real income and the number of confirmed cases of illnesses began impacting the quantity purchased of peanut butter at a horizon of 8 weeks. In particular, innovations in real income and number of confirmed cases of illnesses accounted for about 3% and 5% of the FEV, respectively, for longer-term horizons.

Figure 4 presents DAG/Bernanke impulse-response functions in graphic format in an attempt to quantify the impact of a 1-standard-deviation shock in the error term or innovation of the variables on the quantity purchased of peanut butter. By applying this one-time exogenous shock to each variable, it was possible to trace out a dynamic picture of how the variables responded over a period of 52 weeks. In Figure 4, the impulse responses for all variables were normalized by dividing them by the historical standard deviation of the corresponding error term (innovation) in the VEC model to make the graphs comparable with each other irrespective of measurement units. In Figure 4, the responses are listed at the top of each column, given a one-time-only shock in the variables listed at the beginning of each row.

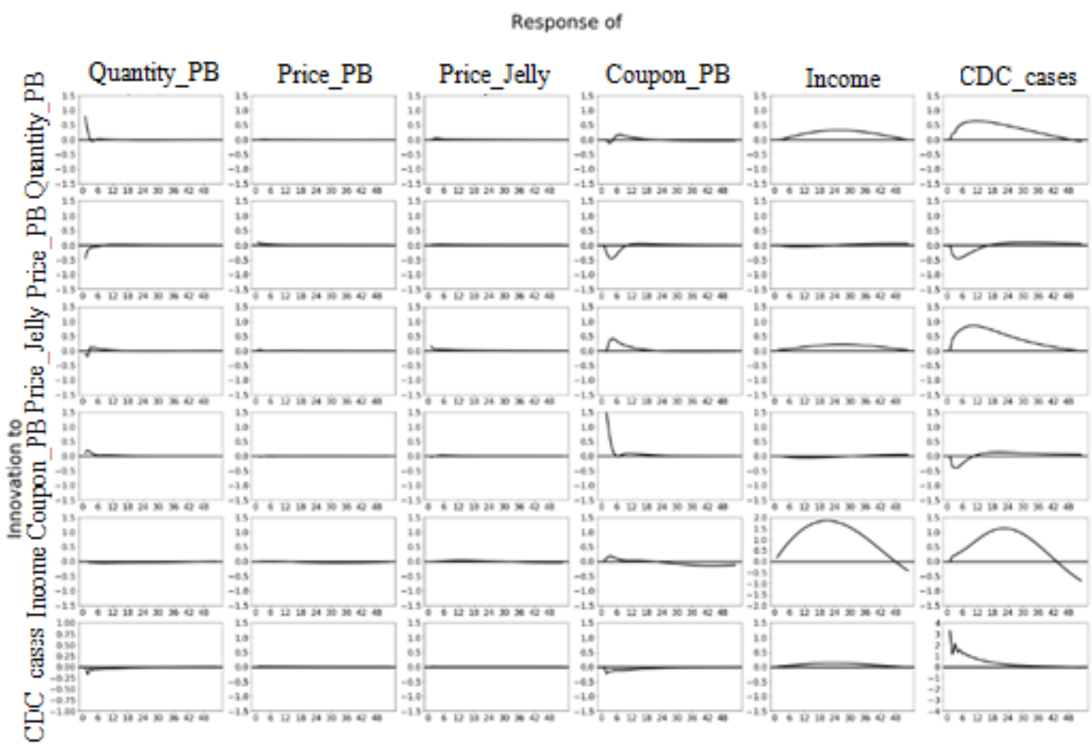

Figure 4. Impulse-Response Functions Generated by the Vector Error Correction Model
Notes: Natural logarithmic transformation was used on all variables, with the square root transformation applied to the CDC_cases variable. Calculated based on data from The Nielsen Company (U.S.), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Our primary interest lies in the response of the quantity of peanut butter purchased (the first column of Figure 4) following an initial one-time-shock only in the respective variables. According to Figure 4, the impacts dampened out over the 52-week period. The response of the quantity of peanut butter purchased to its own shock was positive and peaked in week 1. As expected, the response of the quantity of peanut butter purchased to the shock in the real price of peanut butter was negative with the peak taking place in week 1 as well. The response of the quantity of peanut butter purchased to the shock in the real price of jelly started out negative, as anticipated, for the first 2 weeks following the shock, but subsequently turned positive from weeks 3 through 14. The peak of the impact of the real price of jelly took place in week 2. The response of the quantity of peanut butter purchased to the shock in the real income was positive, peaking at week 2. The response of the quantity of peanut butter purchased to the shock of coupon values was negligible. Finally, the response of the quantity of peanut butter purchased to the shock in the number of confirmed cases of illnesses due to peanut butter consumption was negative throughout the 52-week period, with the peak occurring in week 2.

Concluding Remarks

This study presented an alternative methodological approach of time-series analysis, in contrast to a structural analysis by Bakhtavoryan, Capps, and Salin (2014b), to investigate the demand for peanut butter in the wake of a product recall. This study estimated a 2-lag DAG/Bernanke VEC model using weekly time-series data from July 26, 2006, through December 30, 2008, and using the number of confirmed cases of illnesses due to peanut butter consumption to account for the effects of the recall. The estimation results identified the real price of peanut butter, real price of jelly, real income, the outbreak variable, a structural dummy variable, and seasonality as statistically significant determinants of the quantity purchased of peanut butter. In particular, consistent with previous research, the real price of peanut butter negatively influenced the quantity purchased of peanut butter (Rimal, Fletcher, and Deodhar, 2001), the real price of jelly negatively impacted the quantity purchased of peanut butter (He et al., 2004; Smith, Rossi, and Allenby, 2016; Caine-Bish and Scheule, 2007), real income positively affected the quantity purchased of peanut butter (Rimal, Fletcher, and Deodhar, 2001), and the recall negatively impacted the quantity purchased of peanut butter (Swartz and Strand, 1981; Smith, van Ravenswaay, and Thompson, 1988; van Ravenswaay and Hoehn, 1991; Burton and Young, 1996; Verbeke and Ward, 2001; Marsh, Schroeder, and Mintert, 2004; Piggott and Marsh, 2004; Pritchett et al., 2007, Duan, 2014), with the last empirical finding being consistent with the results from the prior studies reviewed. Also, in accordance with previous research, a structural change in the demand for peanut butter was found in the wake of the recall (Bakhtavoryan, Capps, and Salin, 2012), and seasonality emerged as a statistically significant driver of the quantity purchased of peanut butter (Rimal, Fletcher, and Deodhar, 2001).

In addition, all findings compare favorably with those by Bakhtavoryan, Capps, and Salin (2014b), with two exceptions. First, the previous study found the real price of jelly to be a statistically insignificant driver of the quantity purchased of peanut butter. Second, and more importantly, the two studies are at odds concerning the impact of the outbreak variable on the quantity purchased of peanut butter. In particular, while Bakhtavoryan, Capps, and Salin (2014b) found that the outbreak variable positively affected the quantity purchased of peanut butter, the present study found that the outbreak variable had a negative impact on the quantity purchased of peanut butter. The discrepancy can likely be attributed to differences in the methodological approach (i.e., the use of a VEC model as opposed to a structural/econometric model). The use of time-series models in analyzing the impacts of food safety incidents has been sparse in the extant literature. Hence, using time-series models as well as structural/econometric models is recommended when examining impacts of food safety incidents as a check on the robustness of the results.

Foodborne illnesses remain a topical issue, and the empirical finding showing the negative impact of the recall on the peanut butter category has implications for public regulatory institutions responsible for assuring the safety of the nation's food supply. Moreover, food manufacturers' strategic decisions about quality control programs are informed by this research. Given the cost associated with food recalls, the empirical findings from this study provide further incentive for government regulatory bodies to design and implement recall-preventing policies as well as commit more effort and resources to enhancing their capacity to identify and prevent food safety issues. For peanut butter manufacturers, the extent of spillover from an implicated brand to the entire category constitutes an important and interesting element. As such, the empirical results are essential in that they provide manufacturers with an incentive to adopt and invest in safe production practices as well as closely follow food safety standards to avoid experiencing potential losses in sales in the wake of recalls. In any case, the success of these efforts is inextricably linked with a proper understanding of the economic consequences resulting from food safety issues and the welfare benefits stemming from food safety measures. Finally, the causal relationships that emerge from the study of the peanut butter product market are generalizable to the management of food safety events, and similar case studies can also be replicated for other products implicated in food safety issues.

References

Awokuse, T., and D.A. Bessler. 2003. "Vector Autoregressions, Policy Analysis and Directed Acyclic Graphs: An Application to the U.S. Economy." *Journal of Applied Economics* 6:1–24.

Babula, R., D. Bessler, and W. Payne. 2004. "Dynamic Relationships among U.S. Wheat-Related Markets: Applying Directed Acyclic Graphs to a Time Series Model." *Journal of Agricultural and Applied Economics* 36:1–22.

Bakhtavoryan, R., O. Capps Jr., and V. Salin. 2012. "Impact of Food Contamination on Brands: A Demand Systems Estimation of Peanut Butter." *Agricultural and Resource Economics Review* 41:327–339.

Bakhtavoryan, R., O. Capps Jr., and V. Salin. 2014a. "The Impact of Food Safety Incidents Across Brands: The Case of the Peter Pan Peanut Butter Recall." *Journal of Agricultural and Applied Economics* 46:559–573.

Bakhtavoryan, R., O. Capps Jr., and V. Salin. 2014b. "Dynamics of Consumer Response to Food Contamination: The 2007 Peanut Butter Recall." *Journal of Food Distribution Research* 45:66–81.

Bakhtavoryan, R., O. Capps Jr., and V. Salin. 2014c. “A Household-Level Perspective on the 2007 Peanut Butter Food Safety Incident.” College Station, TX: Texas A&M University Agribusiness, Food, and Consumer Economics Research Center, Research Report No. CM−02-14.

Bernanke, B. 1986. “Alternative Explanations of the Money Income Correlation.” *Carnegie-Rochester Conference Series on Public Policy* 25:45–100.

Bessler, D.A. 1984a. “Relative Prices and Money: A Vector Autoregression on Brazilian Data.” *American Journal of Agricultural Economics* 66:25–30.

Bessler, D.A. 1984b. “An Analysis of Dynamic Economic Relationships: An Application to the U.S. Hog Market.” *Canadian Journal of Agricultural Economics* 32:109–124.

Bessler, D.A., and D. Akleman. 1998. “Farm Prices, Retail Prices, and Directed Graphs: Results for Pork and Beef.” *American Journal of Agricultural Economics* 42:1144–1149.

Bessler, D.A., and S. Lee. 2002. “Money and Prices: U.S. Data 1869–1914.” *Empirical Economics* 27:427-446.

Bessler, D.A., and N. Loper. 2001. “Economic Development: Evidence from Directed Acyclical Graphs.” *Manchester School* 69(4):457–476.

Brandt, J.A., and D.A. Bessler. 1984. “Vector Autoregressions on U.S. Hog Prices.” *North Central Journal of Agricultural Economics* 6:29–36.

Burton, M., and T. Young. 1996. “The Impact of BSE on the Demand for Beef and Other Meats in Great Britain.” *Applied Economics* 28: 687–693.

Caine-Bish, N., and B. Scheule. 2007. “Food Preferences of School Age Children and Adolescents in an Ohio School District.” *Journal of Child Nutrition & Management* 31:532–540.

Capps, O. Jr., D.A. Bessler, and G.W. Williams. 2016. “The Ramifications of Nearly Going Dark: A Natural Experiment in the Case of U.S. Generic Orange Juice Advertising.” *Agricultural and Resource Economics Review* 45:68–97.

Centers for Disease Control and Prevention. 2007. “Multistate Outbreak of *Salmonella* Serotype Tennessee Infections Associated with Peanut Butter – United States, 2006–2007.” *Morbidity and Mortality Weekly Report* 56:521–524.

Chickering, D. 1995. “A Transformational Characterization of Equivalent Bayesian Network Structures.” *Proceedings of the 11th Annual Conference on Uncertainty in Artificial Intelligence*, pp. 87–98.

Chickering, D., C. Meek, and D. Heckerman. 2003. “Large-Sample Learning of Bayesian Networks is NP-hard.” *Proceedings of the 19th Annual Conference on Uncertainty in Artificial Intelligence*, pp. 124–133.

Cormen, H.T., C.E. Leiserson, R.L. Rivest, and C. Stein. 2009. *Introduction to Algorithms*, 3rd ed. Cambridge, MA: MIT Press.

Demiralp, S., and K. D. Hoover. 2003. "Searching for the Causal Structure of a Vector Autoregression." *Oxford Bulletin Economics Statistics* 65:745–767.

Doan, T. 2002. "Generalized Impulse Responses." *RATSletter* 15:4.

Duan, W. 2014. "Effects of the 2007 Peanut Butter Recall on Consumers' Brand Choices." MS thesis, Rutgers University.

Engle, R.F., and C.W.J. Granger. 1987. "Cointegration and Error Correction: Representation, Estimation and Testing." *Econometrica* 55:251–276.

Geiger, D., T. Verma, and J. Pearl. 1990. "Identifying Independence in Bayesian Networks." *Networks* 20:507–534.

Granger, C. W. J. 1981. "Some Properties of Time Series Data and Their Use in Econometric Model Specification." *Journal of Econometrics* 16:121–130.

Gujarati, D. 2004. *Basic Econometrics*, 4th ed. New York, NY: McGraw-Hill.

Haigh, M., and D. Bessler. 2004. "Causality and Price Discovery: An Application of Directed Acyclic Graphs." *Journal of Business* 77:1097–1121.

He, S., S. M. Fletcher, M. S. Chinnan, and Z. Shi. 2004. "Factors Affecting School Students' Consumption of Peanut Butter Sandwiches." *Journal of Food Distribution Research* 35:112–117.

Hoover, K.D., S. Demiralp, and S.J. Perez. 2009. "Empirical Identification of the Vector Autoregression: The Causes and Effects of US M2." In J.L. Castle and N. Shephard, eds. *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*. Oxford, UK: Oxford University Press, pp. 37–58.

Johansen, S. 1988. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control* 12:231–254.

Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford, UK: Oxford University Press.

Juselius, K. 2006. *The Cointegrated VAR Model: Methodology and Applications*. UK: Oxford University Press.

Kalisch, M., M. Maechler, D. Colombo, M.H. Maathuis, and P. Buehlmann. 2012. "Causal Inference Using Graphical Models with the R Package pcalg." *Journal of Statistical Software* 47:1–26.

Koller, D., and N. Friedman. 2010. *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA: MIT Press.

Lütkepohl, H., and M. Kratzig, eds. 2004. *Applied Time Series Econometrics*. Cambridge, UK: Cambridge University Press.

Marsh, T.L., T.C. Schroeder, and J. Mintert. 2004. "Impacts of Meat Product Recalls on Consumer Demand in the USA." *Applied Economics* 36:897–909.

Pearl, J. 1986. "Fusion, Propagation and Structuring in Belief Networks." *Artificial Intelligence* 29:241–288.

Pesaran, H., and Y. Shin. 1998. "Generalized Impulse Response Analysis in Linear Multivariate Models." *Economic Letters* 58:17–29.

Phoong, S.-W., M.T. Ismail, and S.-K. Sek. 2014. "Linear Vector Error Correction Model versus Markov Switching Vector Error Correction Model to Investigate Stock Market Behaviour." *Asian Academy of Management Journal of Accounting and Finance* 10: 133–149.

Piggott, N.E., and T.L. Marsh. 2004. "Does Food Safety Information Impact US Meat Demand?" *American Journal of Agricultural Economics* 86:154–174.

Pritchett, J., K. Johnson, D. Thilmany, and W. Hahn. 2007. "Consumer Responses to Recent BSE Events." *Journal of Food Distribution Research* 38:57–68.

Rimal, A.P., S.M. Fletcher, and S. Deodhar. 2001. "Dynamic Effects of Peanut Butter Brand Advertising." *Journal of Food Products Marketing* 6:45–62.

Silander, T., and P. Myllymäki. 2006. "A Simple Approach for Finding the Globally Optimal Bayesian Network Structure." *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence*, pp. 445–552.

Sims, C. 1980. "Macroeconomics and Reality." *Econometrica* 48:1–48.

Singh, A., and A. Moore. 2005. "Finding Optimal Bayesian Networks by Dynamic Programming." Pittsburgh, PA: Carnegie Mellon University, Technical Report.

Smith, A.N., P.E. Rossi, and G.M. Allenby. 2016. "Inference for Product Competition and Separable Demand." Los Angeles, CA: University of California-Los Angeles, Working Paper. Available online: http://www.anderson.ucla.edu/Documents/areas/fac/marketing/Seminars/Fall%202016/SMITH-ADAM-JMP.pdf

Smith, M.E., E.O. van Ravenswaay, and S.R. Thompson. 1988. "Sales Loss Determination in Food Contamination Incidents: An Application to Milk Bans in Hawaii." *American Journal of Agricultural Economics* 70:513–520.

Spirtes, P., C. Glymour, and R. Scheines. 2000. *Causation, Prediction, and Search*, 2nd ed. Cambridge, MA: MIT Press.

Swanson, N.R., and C.W.J. Granger. 1997. "Impulse-Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions." *Journal of the American Statistical Association* 92:357–367.

Swartz, D.G., and I.E. Strand Jr. 1981. "Avoidance Costs Associated with Imperfect Information: The Case of Kepone." *Land Economics* 57:139–150.

van Ravenswaay, E.O., and J.P. Hoehn. 1991. "The Impact of Health Risk Information on Food Demand: A Case Study of Alar and Apples." In J.A. Caswell, ed. *Economics of Food Safety*. New York, NY: Elsevier, pp. 155–172.

Verbeke, W., and R.W. Ward. 2001. "A Fresh Meat Almost Ideal Demand System Incorporating Negative TV Press and Advertising Impact." *Agricultural Economics* 25:359–374.

U.S. Department of Commerce. 2011. "Personal Income and Its Disposition." Washington, DC: U.S. Department of Commerce, Bureau of Economic Analysis. Available online: http://www.bea.gov/newsreleases/national/pi/pinewsrelease.htm [Accessed January 10, 2012]