Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Structural Model of a Firm's Operating Cash Flow with Applications

Kashish Arora, Vishal Gaur

To cite this article:
Kashish Arora, Vishal Gaur (2025) A Structural Model of a Firm's Operating Cash Flow with Applications. Management Science 71(10):8727-8749. https://doi.org/10.1287/mnsc.2021.03790

informs.
https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 71, No. 10, October 2025, pp. 8727–8749
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

A Structural Model of a Firm's Operating Cash Flow with Applications

Kashish Arora,[a] **Vishal Gaur**[b,*]

[a] CUHK Business School, The Chinese University of Hong Kong, Hong Kong; [b] Cornell SC Johnson College of Business, Cornell University, Ithaca, New York 14853

*Corresponding author

Contact: kashish1701@gmail.com, https://orcid.org/0000-0003-0442-2924 (KA); vg77@cornell.edu, https://orcid.org/0000-0002-3095-4310 (VG)

Received: November 16, 2021
Revised: December 20, 2023
Accepted: May 30, 2024
Published Online in Articles in Advance: February 12, 2025

https://doi.org/10.1287/mnsc.2021.03790

Abstract. Effective management of a firm's operating cash flow is essential for supporting growth, servicing debt, and maintaining overall financial health. Mismanagement of cash flows can result in severe liquidity challenges and even business failure. However, managing operating cash flow is complex because of its intricate, endogenous relationships with operational variables, like sales, operating costs, inventory, payables, and the impact of exogenous macroeconomic factors on a firm. In this paper, we present a structural model of operating cash flow that untangles this endogeneity, allows us to estimate causal relationships among these variables, and provides a valuable tool for evaluating cash flow management policies. Applying our model to quarterly financial data from S&P's Compustat database spanning from 1990 to 2020 along with macroeconomic indicators, we provide empirical evidence of the endogenous nature of cash flow with other operational variables. We then showcase the practical value of our model by (i) identifying the characteristics of structural shocks and the new equilibria they induce within the system; (ii) offering a tool for evaluating alternative managerial actions or policy decisions to counteract these shocks; (iii) predicting the impacts of macroeconomic events, such as global recessions and fluctuations in economic sentiment, on firm performance; and (iv) demonstrating superior forecasting performance compared with traditional univariate models. In summary, our structural model of operating cash flow enhances our understanding of its dynamics, enabling better-informed decision making and more effective cash flow management in firms.

History: Accepted by David Simchi Levi, operations management.

Supplemental Material: The online appendix and data files are available at https://doi.org/10.1287/mnsc.2021.03790.

Keywords: empirical operations management • supply chain management • economic shocks • structural model • cash flow management and forecasting

1. Introduction

The state of a firm is characterized by a large set of financial statement variables that vary collectively over time because of managerial actions, exogenous economic shocks, and random noise. The firm's corporate planning, budgeting, and forecasting activities require modeling these dynamic relationships. For example, when a firm faces a recession, a supply chain shock, or a market change, managers need to predict the impact of this shock on the system and evaluate alternative decisions. In this paper, we propose a structural model to estimate the contemporaneous relationships among these variables and show their applications for managerial decision making. Specifically, we focus on the operating cash flow (OCF) and its relationship with the various operational variables of a firm.

Operating cash flow plays a significant role in facilitating and supporting a firm's growth; a higher OCF helps the firm to fund growth initiatives, such as developing new products or services, expanding into new markets, and acquiring other businesses. A higher OCF also helps the firm service its debt and short-term liabilities, contributing to its overall financial health. Not surprisingly, financial analysts have emphasized OCF as one of the most crucial elements for sustained growth and a top financial priority.[1] On the flip side, a poor

OCF can cause liquidity issues, limited growth, poor credit ratings, and higher risk of bankruptcy. According to a study by the U.S. Bank after the 2008–2009 Great Recession, of all the businesses that failed in 2010, cash flow was the main driver of failure for 82% of these firms.[2] Therefore, managing OCF is essential for a firm's financial health, stability, and growth.

However, managing OCF is an extremely challenging problem because OCF is an outcome of a multitude of operational variables as follows:[3]

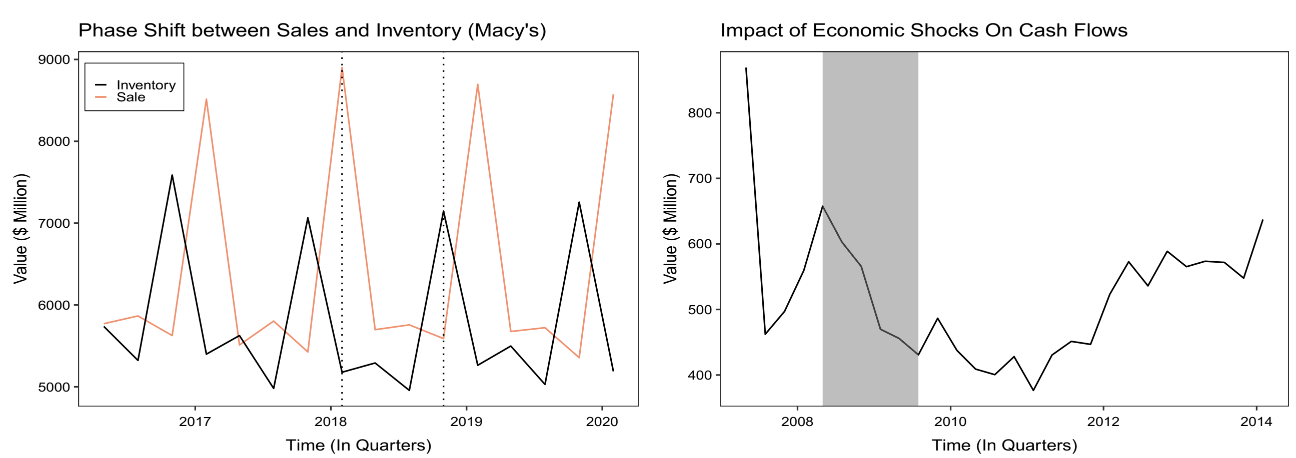
$$\begin{aligned}&\textit{Cash Flow from Operations}_t\\&\quad= \textit{Sales}_t - \textit{Operating Cost}_t - \Delta\textit{Inventory}_t + \Delta\textit{Payables}_t\\&\qquad - \Delta\textit{Recevables}_t + \Delta\textit{Other Operational Assets}\\&\qquad \textit{and Liabilities}_t.\end{aligned} \tag{1}$$

Specifically, OCF is influenced by sales and its seasonal fluctuations, operating costs, inventory management, the credit terms extended by the firm to its customers (accounts receivables), the payment terms agreed upon with its suppliers, etc. The accounts payables and receivables create timing mismatches; payments incurred today may get settled in the future. These relationships and mismatches make cash flows different from net profit and create high variability in cash flow, which is hard to manage.

The OCF management problem is further complicated because the terms on the right side of (1) are endogenously determined and are functions of the *contemporaneous* values of each other. This endogeneity makes it hard to evaluate any policies for cash flow management. Consider a chief financial officer facing an economic shock who is contemplating investing in a new inventory management policy to maintain a healthy level of cash. Equation (1) suggests that reducing one unit of inventory would increase cash flow by one unit. However, inventories affect sales, and sales are used as input to determine inventory amounts. Lower sales and inventory levels may imply that a firm can generate less cash flow from its business, which may constrain the future management of sales, inventory, and cash flow from operations. Similar endogeneity exists between sales, inventories, accounts payables, etc. Additionally, these operational and financial variables exhibit *lagged relationships.* In Figure 1, left panel, we show how Macy's inventory time series for 2015–2020 is out of phase from the sales series by one quarter (Q). Such phase differences induce a dynamic structure among the variables such that a shock to sales can impact the inventory and cash flow in the future. As an illustration, Figure 1, right panel shows how the 2008 economic recession impacted the cash flows of Macy's, Inc. significantly in the subsequent years.

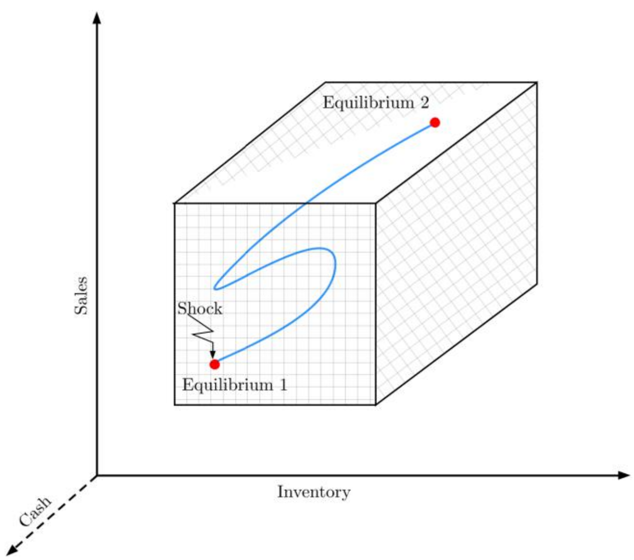
Thus, determining the impact of any of these policies requires carefully estimating the response of the system of endogenous variables to an impulse, which may include external shocks and managerial decisions, such as a sales boost or inventory management. Figure 2 shows an example of an impulse response of the endogenous system of sales, inventory, and cash variables when subjected to such a policy shock. The variables evolve and equilibrate to a new point in the three-dimensional space. The difference between the starting and ending values is the actual causal effect of the shock. In this paper, we construct a dynamic structural model of a firm's operational variables and cash flow, and we causally estimate this model to answer several questions. (i) Does cash flow have an endogenous and dynamic relationship with operational variables? (ii) What are the short- and long-term implications of economic shocks on a firm's cash flow, and how are operational variables useful in managing cash flows through

Figure 1. (Color online) Motivating Examples



Notes. The left panel shows the phase difference between inventory (lower peaked series) and sales (higher peaked series) time series of Macy's from 2015 to 2020. The right panel shows the impact of the 2008 recession on the moving average of Macy's cash flows.

Figure 2. (Color online) Impulse Response of a Sales-Inventory-Cash System



this structural model? (iii) Is there reduced-form evidence that operational variables can be useful for improving the prediction accuracy of cash flow forecasts?

Our analysis proceeds in two steps. In the first step, we formulate a firm-level structural model of the variables' relationships and their evolution over time. Our model includes sales; inventory; cash flow; accounts payables; and selling, general, and administrative (SGA) expenses. We comprehensively allow for endogenous and cross-sectional lagged relationships among variables (e.g., cash flow → sales, accounts payables → inventory, etc.). Our work adapts the structural vector autoregressive (SVAR) methodology developed in the seminal work of Sims (1980) in macroeconomics to firm-level operational data. We estimate our model for U.S. public firms in manufacturing, wholesale, and retail sectors of the economy using quarterly financial and operational data from S&P's Compustat database and data on macroeconomic indicators. Our variables form a multivariate time series of eight variables spanning 30 years from 1990 to 2020 at quarterly frequency.

The main estimation challenge in our setting boils down to appropriate identification. A common identification strategy used in the macroeconomics literature is Cholesky factorization, which works by imposing a causal hierarchy among the variables of interest. In our setting, such an approach conflicts with the vast prior literature in operations management (OM) that documents evidence against causal ordering. The OM literature has studied mechanisms such as scarcity effect, service-level effect, and impact of variety, because of which inventory and sales affect each other contemporaneously (Olivares and Cachon 2009, Cachon et al. 2019). Similarly, the analytical literature on the OM-finance interface has extensively studied the endogeneity between cash flows and other variables. Thus, we utilize an alternative identification strategy that relies on incorporating global macroeconomic variables into the model by imposing structural restrictions on the matrix of endogenous relationships. The macroeconomic variables impact the firm's financial and operational variables and are exogenous to the firm (e.g., Figure 1, right panel). The structural restrictions carefully imposed using these exogenous variables yield two advantages. First, they allow us to estimate the model without imposing any causal hierarchy on the firm-level endogenous variables, and second, they enable us to model the effect of macroeconomic shocks on firms' performance and decision making. We apply these restrictions such that our model is just identified, making the full information maximum likelihood (FIML) estimator equivalent to the Instrumental Variable (IV) estimator that is often used in causal analysis.

In the second step of our analysis, we demonstrate that our structural model can be used for cash flow management by (i) measuring structural shocks from data and computing the new equilibria induced by them; (ii) providing a tool to evaluate compensating managerial actions or policy decisions; (iii) predicting the impacts of the effects of macroeconomic events, like global recessions and varying economic sentiment, on firm performance; and (iv) jointly forecasting cash flows and operational variables to improve forecast accuracy as compared with traditional univariate time series models.

The results of our analysis are as follows. First, from a theoretical perspective, our paper yields evidence that cash flow is endogenous with sales, inventory, and other financial variables. This implies that firms can benefit from coordinating sales, operations, and finance functions. To the best of our knowledge, ours is the first empirical work in operations management that focuses on cash flow management and ties together relationships across a wide range of operational and financial variables that have previously been studied in theoretical models.

Second, we construct the impulse response functions (IRFs) from our structural estimates and show how the IRFs result in a new equilibrium for the system of variables. We show this via a stylized example of a competitor's entry into a market that shocks a focal firm's sales. Our model predicts that in response, the firm's sales decline, and the inventory increases, resulting in a new equilibrium (which is guaranteed to be unique) achieved in three to four years. Further, we show that the firm can mitigate the effects of structural shocks by taking compensating managerial actions or policy decisions. Continuing the above example, the incumbent firm facing a competitor entry can increase its inventories or provide a sales boost to recover the loss of market share. Our model yields the amount of compensating action needed every period in the future to restore market share for the focal firm. Thus, it provides a tool to evaluate such compensating actions or managerial policies for cash flow management. Finally, we show that the model can be applied in this fashion to manage the impact of macro-level phenomena on firm performance. Specifically, we study the impact of the dot-com bust of 2001, the Great Recession of 2008–2009, and the economic uncertainty of 2022–2023 on firms' sales and inventories, with Macy's as a case in point. For the two recessionary periods, using the sequential structural shocks to the Gross Domestic Product (GDP) rate during the contraction period, we trace out the system's path to equilibrium and find that the recessions severely impacted Macy's quarterly sales. For the recent economic uncertainty, we demonstrate the impact of the decline in consumer sentiment on the future projection of Macy's sales. These case studies and counterfactual analyses can be combined with existing multiperiod decision models for cash flow management and operations management under uncertainty.

Third, we use our model to generate joint forecasts for cash flows and other operational variables in our model to develop insights into the information value of these variables. Cash flow forecasting is in itself a hard problem. In the accounting literature, researchers have studied the accuracy of annual forecasts of cash flow using past values of cash flow, earnings, and disaggregated accrual terms (Nallareddy et al. 2020) and consistently found the mean absolute percentage error (MAPE) of cash flow forecasts to be an order of magnitude higher than sales or earnings forecasts. Although the statistics reported in this literature are for annual forecasts, we obtain similar findings for forecasts of quarterly cash flows by generating cash flow forecasts using the best-fitted Autoregressive Integrated Moving Average (ARIMA) model on a rolling horizon for a sample of public U.S. firms. The resulting mean absolute percentage error is of the order of 50%–60%, whereas the MAPE for similar models for sales forecasts is about 4%–5%.

By generating forecasts from our model on a rolling-horizon basis and comparing our forecasts' accuracy with that generated from corresponding autoregressive models, we find that the inclusion of cross-sectional operational variables yields better forecasts for one-quarter-ahead sales, inventory, and cash flow than the autoregressive models consistently across firms and different model specifications. Specifically, our model achieves improvement in sales, inventory, and cash flow forecasts for about 68%, 66%, and 63% of firms in our sample, respectively. The average improvement in sales MAPE over all firms is about 0.49%, and the average improvements in inventory and cash flow MAPE are about 0.29% and 19.13%, respectively. All of the above observations help make a strong case for using our model for forecasting.

To our knowledge, our paper presents the first model in OM that empirically ties together the relationships of cash flow with a wide range of operational variables at the firm level, is generalizable, and is useful in firm-level operational decision making and forecasting.

2. Literature Review

Our paper is related to the academic literature in OM, accounting, and the OM-finance interface. In this section, we summarize relevant papers from these literature streams and describe our paper's contributions.

2.1. Theoretical Research at the OM-Finance Interface

The theoretical research at the operations-finance interface provides a strong rationale for studying cash flow in conjunction with operational variables. Several research papers consider cash as a constraint in inventory replenishment in both single-period and multiperiod models of debt financing (Chod 2017), asset-based lending (Buzacott and Zhang 2004, Alan and Gaur 2018), and trade credit (Gupta and Wang 2009, Yang and Birge 2018). Luo and Shang (2015) study joint optimization of cash and inventory replenishment for a centralized multidivisional supply chain. Li et al. (2013) show the optimality of a base stock policy under financial constraints in a single-echelon setting, whereas Hu and Sobel (2007) show that echelon base stock policies are not optimal under financial constraints in a centralized multiechelon inventory system. Even when firms are not financially constrained, the timing of payments can affect inventory stocking decisions (Tong et al. 2020). Thus, cash is understood to be an integral part of operational decisions and is linked to sales and inventory. Further, Aviv (2003) makes provisions for including exogenous economic variables, such as GDP growth rate, in demand forecasting in a theoretical inventory planning model. Although these papers study optimal decision making at the product level, motivated by this research, we propose an integrative model that estimates the relationships between variables at the firm level and enables firm-level decision making.

2.2. Empirical Work at the OM-Finance Interface

Our paper is also related to the empirical work at the OM-finance interface in two ways: studying the effect of exogenous shocks on operational variables and adding to the literature on cash flow management. One stream of literature at the OM-finance interface has studied specific shocks or supply chain disruptions. For example, Carvalho et al. (2021) document the upstream and downstream propagation of shocks from the Great East

Japan earthquake. Hendricks et al. (2020) show that firms experiencing the Great Japan Earthquake lost 5.21% of their shareholder value one month after the earthquake. Barrot and Sauvagnat (2016) study major natural disasters across 30 years in the United States and find that these events have large short-term effects, not only on the sales growth of affected firms but also, on their customers. Jola-Sanchez and Serpa (2021) studied the civil war in Colombia and found that the war led to a reduction in firms' inventory levels as firms replaced inventory with cash. Agca et al. (2022a) study the role of coronavirus disease 2019 (COVID-19) in reshaping firms' sourcing decisions to more locally sourced manufacturing capacity. Similar to these papers, we study the impact of macroeconomic shocks on firm-level variables and illustrate our results using recent economic events. Our methodology contributes to the literature by providing a structural model that includes different variables simultaneously and can be applied as a tool in decision making.

Another stream of literature studies the links between supply chain structure and liquidity. Jacobson and von Schedvin (2015) study the propagation of corporate bankruptcy and the importance of trade credit chains as a channel. Agca et al. (2022b) study shock propagation in the credit default swap (CDS) market across the supply chain. The authors use supply chain link data to report the abnormal CDS spread change for firms whose customers have been shocked by credit. Wu et al. (2022) develop a gradient-boosting-based machine learning model to predict firms' credit ratings, which are essentially used as key inputs by banks and other firms to determine trade credit. Unlike these papers, we do not include supply chain data in our model, but our methodology can be extended to include upstream and downstream supply chain information. There is not much empirical work in the area of cash flow management, except our paper and the work of Osadchiy et al. (2022), who analyze the variability in inbound and net cash flows and show that firms reduce their cash flow variability by acquiring appropriate customer firms and offering them suitable trade credit terms. Like them, we also highlight the importance of managing cash flows and analyze compensating managerial actions for managing cash flows. However, we focus on a different problem of modeling the endogenous relationships among variables and the effect of external shocks on those variables.

2.3. Empirical Research in Operations Management

Our paper also builds on the rich empirical literature in OM, which has discussed many of the relationships modeled in our paper. It is well known that inventory is endogenous to sales. Gaur et al. (2005) show that inventory turnover is correlated with gross margin and sales surprise. Kesavan et al. (2010) show that a joint consideration of inventory, sales revenue, and gross margin improves sales forecasts. Olivares and Cachon (2009) present evidence of two mechanisms—a sales effect and a service-level effect—to explain how sales impact inventory levels. Conversely, Cachon et al. (2019) demonstrate the reverse relationship, showing that inventory impacts sales through a scarcity effect and a variety effect. Lagged relationships between sales and inventory have been studied by Bray and Mendelson (2012), who analyze a structural model of information-driven bullwhip effect and show that sales information of different lags influences future inventory values. Further, macroeconomic variables have been incorporated in empirical inventory models. Kesavan et al. (2016) show that high- and low-inventory turnover firms react differently to contemporaneous demand shocks, and Rumyantsev and Netessine (2007) use interest rates as an explanatory variable in an empirical model of the amount of inventory carried by firms. Although these research papers have focused on the drivers and implications of inventory levels, Jola-Sanchez and Serpa (2021) show that inventory and cash are treated as substitutes by firms, and Gao (2018) shows that the use of Just-In-Time by U.S. manufacturers is associated with a decrease in inventory and an increase in cash hoardings.

2.4. Accounting Literature on Cash Flow Forecasting

The accounting literature has underscored the importance of cash flow forecasting and studied its many aspects, including (i) whether cash flows are a better predictor of future cash flows than accrual earnings (Kim and Kross 2005, Lorek and Willinger 2009, Nallareddy et al. 2020); (ii) what is the incremental power of disaggregated accruals (i.e., changes in accounts receivables, payable, inventory, etc.) on the prediction of cash flow compared with a model with current period earnings or cash flow (Barth et al. 2001); (iii) does the method of measurement of cash flow (i.e., using balance sheet data or the cash flow statement defined according to accounting standards) make a difference to the relative performance of prediction methods (Hribar and Collins 2002); (iv) are firm-specific time series methods better at predicting future cash flow than cross-sectional methods with pooled coefficients (Lorek and Willinger 2009); and (v) has the relative prediction accuracy of different models changed over time because of shifts in accounting and operating environments, such as trends toward lower profitability, higher growth rates, increasing using of intangibles, and shrinking operating cycles (Nallareddy et al. 2020). Because this is a rich area of work, we cite only a few recent papers above for brevity. In particular, Nallareddy et al. (2020) present a comprehensive recent study of the above questions and answer them in the affirmative. For OM researchers, it is worth noting that a fundamental difference between earnings and cash flows is that earnings include both cash and accrual earnings. Accrual accounting has a temporal or "timing" effect on cash flows; for example, an increase in inventory can depress current cash flows but boost future cash flow. Further, from operations theory, an inventory increase can boost sales and earnings. Thus, it induces both endogenous and lagged effects that are included in our model.

Our study differs from the accounting literature in many respects. First, it focuses on fundamental operational variables instead of summary performance measures, such as earnings and accruals, so that the variables in our study differ from those in accounting. Further, our structural model of the endogeneity of cash flow with operational variables is new to the literature and is useful in managerial decision making for these variables. Finally, we not only show that operational variables provide improved cash flow forecasting, but also, we show in the reverse direction that cash flow is relevant for sales and inventory forecasting. This can be attributed to a liquidity effect of cash flow on operating decisions relevant to future operating performance.

3. Structural Model of Variable Relationships

We model the relationships among a firm's operational and financial variables using a structural vector autoregressive framework. Our structural framework is built on the following axioms. First, our key variables of interest—sales, inventory, cash flow from operations, accounts payable, and (SGA)—are endogenously determined. Second, additional exogenous variables can be included in the model for identification. Third, the variables and their fluctuations over time are interrelated and constitute a dynamically evolving system. We model this system using a set of equations that are tied together by a *structure*. Moreover, we assume that the firm has been operating for a sufficiently long time, so its variables can be modeled using the SVAR structure during data collection.

Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ denote the set of n endogenous variables and $\mathcal{Z} = \{z_1, \ldots, z_m\}$ denote the set of m exogenous variables. Each variable $z \in \mathcal{Z}$ impacts all variables $x \in \mathcal{X}$ but not vice versa. As such, the system consists of $n + m$ *contemporaneous* variables with n^2 relationships among the endogenous variables and mn unidirectional relationships from the exogenous to endogenous variables.

Let $\Upsilon = \mathcal{X} \cup \mathcal{Z}$ be the set of all system variables. Each variable $y_i \in \Upsilon$ evolves in a stochastic manner, and the corresponding sequence $(y_{i1}, y_{i2}, y_{i3}, y_{i4}, \ldots)$ constitutes an infinite time series $y_i = (y_{it} : t \in \mathbb{Z})$. We are interested in understanding (a) the joint evolution of this system over time and (b) how shocks to any of the variables affect the system as a whole over time. Let $\mathbf{y_t}$ be the $(n+m) \times 1$ vector of contemporaneous variables at any time t. We model the joint evolution of the variables using the structural specification $\mathcal{B}(L)\mathbf{y_t} = \mathbf{u_t}$, where L is the lag operator and $\mathcal{B}$ is a *matrix-valued polynomial* in L. Specifically, $L^k\mathbf{y_t} = \mathbf{y_{t-k}}$, and $\mathcal{B}(L)$ takes the following form: $\mathcal{B}(L) = \mathbf{B} - \mathbf{\Gamma_1}L - \mathbf{\Gamma_2}L^2 - \mathbf{\Gamma_3}L^3 - \cdots - \mathbf{\Gamma_k}L^k$. We convert the model from its lag operator notation to its matrix form and write it as

$$\underbrace{\mathbf{B}\mathbf{y_t}}_{\text{Contemporaneous Relationships}} = \underbrace{\mathbf{\Gamma_1}\mathbf{y_{t-1}} + \mathbf{\Gamma_2}\mathbf{y_{t-2}} + \cdots + \mathbf{\Gamma_k}\mathbf{y_{t-k}}}_{\text{Lagged Effects}} + \underbrace{\mathbf{u_t}}_{\text{Structural Shocks}}, \tag{2}$$

where $(\mathbf{\Gamma_1}, \mathbf{\Gamma_2}, \ldots, \mathbf{\Gamma_k})$ are $(n+m) \times (n+m)$ coefficient matrices capturing the relationships between the variables and their lagged values, $\mathbf{B}$ is an $(n+m) \times (n+m)$ matrix of full rank that captures the contemporaneous relationships among the variables in $\mathbf{y_t}$, and $\mathbf{u_t}$ is an $(n+m) \times 1$ vector of *structural shocks*. We describe the structure of matrix $\mathbf{B}$ in Section 4 when we discuss our identification strategy.

The structural shocks are assumed to be mean 0 and independent, with a diagonal covariance matrix, $\mathbf{\Sigma} = \mathbb{E}(\mathbf{u_t}.\mathbf{u_t'})$. For example, a competitor's entry can provide a structural shock to a firm's sales. Similarly, unforeseen supply chain disruptions or a change in the inventory management system can provide a structural shock to a firm's inventory; thus, shocks can be positive or negative changes to the system. A shock to a variable in any period has different types of impacts on the system. First, it has a *direct impact* on the variable because of its structural equation. Second, it has a contemporaneous *indirect impact* on other endogenous variables because of the system of relationships modeled in $\mathbf{B}$. Finally, a shock in period t affects variables in future periods through lagged relationships. For example, the competitor's entry has a direct negative impact on a firm's sales in the same period, an indirect impact on same-period inventory and cash flow because of reduced sales, and an indirect impact on all of the variables in future periods because reduced sales may lead to inventory excess and cash flow depletion in the future unless the firm undertakes mitigating actions. Thus, structural shocks govern the dynamics of the entire system.

To summarize, we are interested in estimating the impacts of these shocks: $\mathbb{E}[\mathbf{y_{t+s}} | u_{it} = 1] - \mathbb{E}[\mathbf{y_{t+s}} | u_{it} = 0]$, $\forall s = \{0, 1, ..\}$, when variable i is subjected to a unit structural shock. To do this, we first estimate the unknown matrices $(\mathbf{B}, \mathbf{\Gamma_1}, \ldots, \mathbf{\Gamma_k})$ in (2) and then, use these matrices to compute the impact of shocks. The total number of parameters to be estimated across these matrices is $(n+m)^2 \times (k+1)$, where k is the number of lags.

3.1. Reduced Form

From the structural representation laid out in (2), we can obtain the corresponding vector autoregressive or the *reduced-form* representation (assuming that $\mathbf{B}$ is full rank) as follows:

$$\begin{aligned}\mathbf{y_t} &= \mathbf{B^{-1}\Gamma_1 y_{t-1}} + \mathbf{B^{-1}\Gamma_2 y_{t-2}} + \cdots + \mathbf{B^{-1}\Gamma_k y_{t-k}} + \mathbf{B^{-1}u_t}\\ &= \mathbf{\Pi_1 y_{t-1}} + \mathbf{\Pi_2 y_{t-2}} + \cdots + \mathbf{\Pi_k y_{t-k}} + \underbrace{\boldsymbol{\xi}_\mathbf{t}}_{\text{Forecast errors}},\end{aligned} \tag{3}$$

where the coefficient matrices $\mathbf{\Pi_k}$ are the reduced-form counterparts of matrices $\mathbf{\Gamma_k}$ and $\boldsymbol{\xi}_\mathbf{t}$ is an $(n+m)\times 1$ vector of the reduced-form residuals. Intuitively, the residuals are *forecast errors* of the underlying variables and have the following covariance matrix:

$$\mathbb{E}[\boldsymbol{\xi}_\mathbf{t}\boldsymbol{\xi}'_\mathbf{t}] = \mathbf{B^{-1}}\mathbb{E}[\mathbf{u_t}.\mathbf{u'_t}]\mathbf{B^{-1'}} = \mathbf{B^{-1}IB^{-1'}} = \mathbf{\Omega},$$

where the reduced form and the structural shocks follow the relationship $\boldsymbol{\xi}_\mathbf{t} = \mathbf{B^{-1}u_t}$. In other words, the forecast errors are linear combinations of the mutually orthogonal structural shocks.

3.1.1. A Reduced-Form Model Alone Is Not Suited for Managerial Decision Making. The goal of our structural model is to provide a method to measure the change in variables in response to a shock ceteris paribus. We emphasize that $\mathbf{\Omega}$ is not a diagonal matrix. This correlation between reduced-form errors, captured by $\mathbf{\Omega}$, makes it hard to use the reduced form for managerial decision making. In other words, in a reduced-form model, one cannot hold other forecast errors constant when a shock occurs to one variable because the forecast errors are linear combinations of structural shocks. The structural model eliminates this problem by decomposing the reduced-form errors into orthogonal projections (i.e., the structural shocks). Because the structural shocks are orthogonal by construction, we can measure the impact of one of them while keeping the others constant, which can be used for suitable policy analysis. This requires the estimation of $\mathbf{B}$. Hence, we describe the identification strategy and the estimation procedure necessary for this purpose in Section 4.

3.2. Variables

A key innovation in our paper is that we do not estimate the accounting identity (1) directly but instead, model relationships among variables based on findings in the operations management literature. The previous evidence in the OM literature discussed in Section 2 shows that sales and inventory are endogenous (i.e., an increase in inventory is expected to increase sales, and higher sales are expected to cause firms to invest in more inventory). Further, the endogeneity of cash flow with sales and inventory is expected from the theoretical modeling literature on the OM-finance interface (Chod 2017, Alan and Gaur 2018); higher sales and higher inventory levels should imply that a firm can generate more cash flow from its business, and a higher cash flow from operations should enable a firm to plan for higher sales and inventory levels. Moreover, accounts payable may be endogenously determined with inventory and sales; for example, firms with constrained cash flows but facing a healthy sales season and low inventories may be able to negotiate payment terms with suppliers to finance inventory, affecting accounts payable. Similarly, selling expenses may also be related to inventories. For example, warehouse club retailers, such as Costco, are known for having low product variety and high inventory turns, which can help reduce the selling expenses at their stores. Our model consists of five endogenous variables in set $\mathcal{K}$: sales, inventory, cash flow from operations, accounts payable, and selling expenses. Here, sales, cash flow, and selling expenses are *flow* variables, and inventory and accounts payable are *stock* variables. That is, sales refers to the amount of new sales revenue generated by a business in a quarter, cash flow is the corresponding new cash generated by the business, and inventory and accounts payable refer to the *levels* of inventory and accounts payable at the end of that quarter.

3.2.1. Exogenous Variables. To identify the structural parameters in our model, we use three exogenous variables in $\mathcal{Z}$: the gross domestic product, the consumer confidence index (CCI), and the chief executive officer (CEO) confidence index. The CCI is a leading indicator of households' sentiment about the general economic conditions, unemployment, and savings capability. A higher CCI indicates an optimistic attitude of the households toward the economy, leading to increased consumer spending and hence, higher firms' sales. The CEO confidence index measures the sentiment and confidence of CEOs in the business and economic environment. As the CEO confidence index increases, CEOs are optimistic about the future, which can signal potential economic strength and growth for the firms. Although CCI is mostly a demand-side indicator of the economy, the CEO index reflects the economy's demand- and supply-side conditions. Moreover, although GDP mainly reflects the current economic climate, these two indicators also reflect the expected economic conditions in the near future. All of these variables are macro-level economic indicators outside the firm and exogenous to a single firm. The three exogenous variables provide enough identifying restrictions in the structural model.

Formally, the vector of variables in our model is $\mathbf{y_t} = (\mathfrak{s}_\mathbf{t}\mathfrak{i}_\mathbf{t}\mathfrak{c}_\mathbf{t}\mathfrak{ap}_\mathbf{t}\mathfrak{sga}_\mathbf{t}\mathfrak{gdp}_\mathbf{t}\mathfrak{cci}_\mathbf{t}\mathfrak{ceo}_\mathbf{t})^T$; see Online Appendix A for the corresponding Compustat definitions of the firm-level variables. We allow the endogenous variables to have both contemporaneous and lagged relationships. Moreover, we allow both contemporaneous and lagged values of exogenous variables to affect the endogenous variables. We also note that we use three variables as this is the minimum number required for identification.

3.3. Model Identification

Identification in our setting refers to the ability to recover the structural matrices $(\mathbf{B}, \boldsymbol{\Gamma}_1, \ldots, \boldsymbol{\Gamma}_k)$ in (2) from the reduced-form estimates. Our procedure for doing so is as follows. We first estimate the reduced form of the model by running ordinary least squares regressions for the $(n+m)$ variables in $\mathbf{y_t}$. This yields the estimates of matrices $\boldsymbol{\Pi}_\mathbf{k}$ and $\boldsymbol{\Omega}$. Then, we recover $\mathbf{B}$ from the reduced-form covariance matrix $\boldsymbol{\Omega}$ using maximum likelihood estimation. Finally, we derive matrices $\boldsymbol{\Gamma}_\mathbf{k}$ from $\mathbf{B}$ and $\boldsymbol{\Pi}_\mathbf{k}$ using $\boldsymbol{\Gamma}_\mathbf{k} = \mathbf{B}\boldsymbol{\Pi}_\mathbf{k}$. Thus, the identification problem reduces to estimating the matrix $\mathbf{B}$.

Identification of $\mathbf{B}$ is the most important structural component of the model requiring a modeling decision because the total number of free parameters in the structural model is $(k+1)(n+m)^2$, but the total number of parameters estimated in the reduced form of the model is only $k(n+m)^2 + (n+m)(n+m+1)/2$, where the latter term in the sum corresponds to the estimated covariance matrix $\boldsymbol{\Omega}$. Thus, we require at least $(n+m)(n+m-1))/2$ additional restrictions in the structural form to estimate the model. This is achieved by restricting the matrix $\mathbf{B}$ (Sims 1980). To see why, consider the relationship between $\mathbf{B}$ and the reduced-form covariance matrix $\boldsymbol{\Omega}$: $\mathbf{B}^{-1}\mathbf{B}^{-1'} = \boldsymbol{\Omega}$. Because $\boldsymbol{\Omega}$ is symmetric, it only houses $(n+m)(n+m+1)/2$ estimated parameters, whereas $\mathbf{B}$ has $(n+m)^2$ degrees of freedom. Thus, many matrices $\mathbf{B}$ will solve this system of equations.

Our method to impose identifying restrictions on $\mathbf{B}$ relies on exogenous variables. Consider the submatrices in $\mathbf{B}$ that are associated with endogenous variables $\mathbf{x_t}$ and exogenous variables $\mathbf{z_t}$. Using $\mathbf{B}$'s submatrices, we can modify the left-hand side of (2), which gives

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \begin{pmatrix} \mathbf{X}_\mathbf{t} \\ \mathbf{Z}_\mathbf{t} \end{pmatrix} = \boldsymbol{\Gamma}_1 \mathbf{y}_{\mathbf{t}-1} + \boldsymbol{\Gamma}_2 \mathbf{y}_{\mathbf{t}-2} + \cdots + \boldsymbol{\Gamma}_\mathbf{k} \mathbf{y}_{\mathbf{t}-\mathbf{k}} + \mathbf{u}_\mathbf{t}, \tag{4}$$

where $\mathbf{B}_{11}$ is the leading principal submatrix of order n of the relationships among the endogenous variables, $\mathbf{B}_{12}$ is an $n \times m$ submatrix specifying how the exogenous variables affect the endogenous variables contemporaneously, $\mathbf{B}_{22}$ is an $m \times m$ matrix of the relationships among exogenous variables, and $\mathbf{B}_{21}$ is the remaining submatrix. Because variables $z \in \mathcal{Z}$ are exogenous, we set $\mathbf{B}_{21}$ to $\mathbf{0}$. This gives us mn restrictions achieved by including exogenous variables in the model. The operational researcher can further restrict $\mathbf{B}_{12}$ or $\mathbf{B}_{22}$ to get to at least $(n+m) \times (n+m-1))/2$ restrictions.

Table 1. B Matrix for Our Model

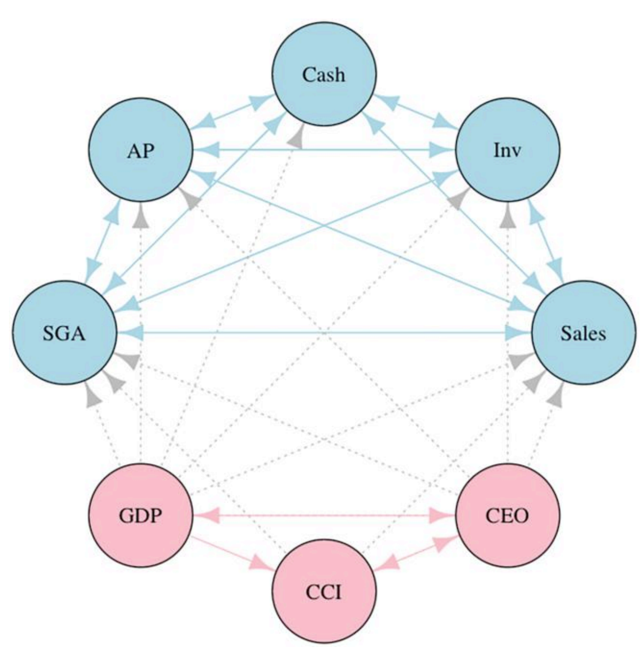
	Internal firm variables					Macrovariables		
Variable	Sales	Inv	Cash	AP	SGA	GDP	CCI	CEO
Sales	1	·	·	·	·	·	·	·
Inv	·	1	·	·	·	·	0	·
Cash	·	·	1	·	·	·	0	0
AP	·	·	·	1	·	·	0	·
SGA	·	·	·	·	1	·	·	·
GDP	0	0	0	0	0	1	0	·
CCI	0	0	0	0	0	·	1	·
CEO	0	0	0	0	0	·	·	1

Note. The center dot represents a free parameter, and zero represents a restriction.

Table 1 shows the $\mathbf{B}$ matrix for our 8×8 model, and Figure 3 shows the corresponding contemporaneous directed graph. The exogeneity of $\mathfrak{gdp}$, $\mathfrak{cci}$, $\mathfrak{ceo}$ variables yields 15 restrictions. Normalizing the diagonal of $\mathbf{B}$ gives eight more. Because the consumer confidence index is a pure demand-side variable, it only affects sales and SGA contemporaneously. The CEO confidence index reflects both the supply-side and demand-side factors. Hence, it affects all firm variables contemporaneously except cash flow—an outcome variable.

Note that all of the exogenous variables in the model also affect each other contemporaneously except $\mathfrak{cci} \longrightarrow \mathfrak{gdp}$. We posit that $\mathfrak{gdp}$ is exogenous to $\mathfrak{cci}$ (i.e., the GDP gets realized first, and the realizations of GDP impact customers' views of the economy). Hence, we add this restriction. Our final system has 28 restrictions

Figure 3. (Color online) Directed Graph of Contemporaneous Relationships



and is just identified, making the estimation equivalent to an IV estimator, which we describe in Section 3.4.

3.3.1. Alternative Identification Methods. We now highlight alternative methods for identifying $\mathbf{B}$ that we explored. A common identification technique that much of the prior SVAR literature relies on is the triangularization of the system of the variables (see, for example, Sims 1980). Given any lower triangular $\mathbf{B}_\triangle$, there is a unique solution to the equation $\mathbf{B}^{-1}\mathbf{B}^{-1'} = \mathbf{\Omega}$. $\mathbf{B}_\triangle$ can be recovered by the Cholesky decomposition of $\mathbf{\Omega}$, and hence, this identification method is also called *Cholesky* identification. Our paper departs from the prior SVAR literature in this respect because setting $\mathbf{B}$ to be lower diagonal imposes a specific causal ordering on the operational variables in the system that is not supported by the OM literature. For example, suppose that we model a triangular system of two variables, sales ($\mathfrak{s}$), and inventory ($\mathfrak{i}$). This system has a 2×2 $\mathbf{B}_\triangle$ matrix and the following structural form:

$$\begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{pmatrix} \mathfrak{s}_\mathbf{t} \\ \mathfrak{i}_\mathbf{t} \end{pmatrix} = \mathbf{\Gamma_1} \begin{pmatrix} \mathfrak{s}_\mathbf{t-1} \\ \mathfrak{i}_\mathbf{t-1} \end{pmatrix} + \mathbf{\Gamma_2} \begin{pmatrix} \mathfrak{s}_\mathbf{t-2} \\ \mathfrak{i}_\mathbf{t-2} \end{pmatrix} + \cdots + \mathbf{\Gamma_k} \begin{pmatrix} \mathfrak{s}_\mathbf{t-k} \\ \mathfrak{i}_\mathbf{t-k} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_{\mathfrak{s}_t} \\ \mathbf{u}_{\mathfrak{i}_t} \end{pmatrix}, \quad (5)$$

where the first relationship in (5) models the evolution of sales over time. Sales at any given time t depend on the lagged values of itself and of inventory but not on contemporaneous inventory. On the other hand, inventory depends on contemporaneous sales, lagged sales, and lagged inventory. Note that this structure is problematic because the OM literature has shown that the contemporaneous causal impact of inventory on sales is nonzero (Kesavan et al. 2010). The same problem persists if the ordering of the variables for triangulation was reversed. Thus, we refrain from using this technique of restricting contemporaneous impacts in our model.

Another strategy used in macroeconomics achieves identification by setting the long-term cumulative impacts of shocks to certain variables to zero. In our example, a long-run impact restriction would force the impact of a structural shock to go to zero in the long run. This also may not be supported by data. Thus, we avoid this identification method.

Note that our identification method relies on a priori theory and evidence from the operations management literature. An alternative approach for identification used in the causal inference and computer science literature aims to directly infer causal relationships using the statistical properties of data (see, for example, Pearl 2009). Examples of this approach are constraint-based algorithms for graphical models, such as Peter-Clark, Spirtes-Glymour-Scheines (SGS), and Fast Causal Inference algorithms (Spirtes et al. 1993). Typically, such algorithms start with a complete directed graph in the first step. In the second step, conditional independence relations are used to erase edges and in further steps, to direct edges (Moneta et al. 2011). The result is a set of directed acyclic graphs (DAGs) that are *Markov equivalent*. This DAG-type identification may not fit our analysis because our directed graph is not acyclic. However, such approaches may be investigated in our setup to estimate $\mathbf{B}$ in a data-driven manner. We leave the validation of such analysis for future work as an extension of the current work.

3.4. Connection to Causality and Equivalence to the IV Approach

The SVAR model identifies causal relationships among the variables by carefully placing restrictions using exogenous variables as described in Section 3.3. The causal interpretation relies on the exogeneity of the variables and the appropriateness of the structural restrictions applied (for example, in a two-variable supply-demand model for the fish market, if the quantity of fish supplied is fixed and does not depend on price contemporaneously, a researcher who knows the fish market would appropriately restrict the $\mathbf{B}$ matrix). The causal interpretation was first shown in early macro papers starting from Cash (C). Sims' Nobel prize winning work (Sims 1980), the contribution of which is to recover causal estimates from structural assumptions and restrictions in the model derived from economic theory. To summarize, given exogenous variables in the model and proper identification (careful restrictions), the dynamic estimates from SVAR are causal.[4]

This causal interpretation has also been noted in academic papers that have described the connection between SVAR and linear IVs, which are commonly used for causal analysis. For example, Hausman (1975) develops an instrumental variable interpretation of the full information maximum likelihood estimator in simultaneous equations and proves that in finite samples, the FIML estimator and IV estimator are *identical* in the just-identified case. Moreover, the two estimators are asymptotically equivalent and converge to the same limiting distribution, even for the nonjust-identified cases. We highlight that our system of equations is just identified as described in Section 3.3.

It is also noteworthy that this equivalence is used by much of the commercial software available for time series econometric analysis (like EViews), which uses the IV estimator to provide starting values for the numerical optimization problem of the Maximum Likelihood Estimation (MLE) estimator. Because many MLE algorithms suffer from slow/no numerical convergence, the IV estimator offers a reasonable set of starting values even for optimization problems for models that are not just identified; both yield the same estimates for just identified cases.

3.4.1. Example. This example illustrates the equivalence of the IV estimator to the FIML estimator by showing how to construct the IVs internally in the system using restrictions. Consider the two-variable toy model for sales and inventory in (6) and (7). The structural errors in sales and inventory equations are assumed to be orthogonal:

$$s_t = a_{11}s_{t-1} + a_{12}i_{t-1} + u_{st} \tag{6}$$
$$i_t - b_{21}s_t = a_{21}s_{t-1} + a_{22}i_{t-1} + u_{it}. \tag{7}$$

Here, the assumption is that sales is independent of inventory contemporaneously. (6) can be estimated by Ordinary Least Squares (OLS) because there are no endogenous variables on its right-hand side. (7) can also be estimated by OLS because $E(s_t u_{it}) = E(a_{11}s_{t-1}u_{it} + a_{12}i_{t-1}u_{it} + u_{st}u_{it}) = 0$ because of the shocks being orthogonal. If we were to use the shocks from (6) as an instrument for s_t in (7), the resulting moment condition would be $E[u_{st}u_{it}] = 0$ (exclusion restriction), which is the same moment condition for the OLS estimator. This instrument is a sales shifter used to identify the inventory equation. We could use u_{st} as an instrument here because of the restriction in the contemporaneous matrix that sets the inventory coefficient to zero in (6). This shows how a restriction on the **B** matrix can lead to generating an instrument in the model.

4. Model Estimation

We describe the estimation of the model in this section. We start by outlining our estimation procedure. Then, we describe the data used for estimation. Finally, we present our results, which provide direct evidence of endogeneity among the variables.

4.1. Estimation Procedure

The following steps outline our estimation procedure. We run the procedure for each firm in the sample and estimate the **B** matrix for all of the firms.

Step 1. Define series $\mathbf{y}'_t = \mathbf{y}_t - \mathbf{y}_{t-4}$, and impute any missing values through linear interpolation.

Step 2. Determine the optimal lag length k for the reduced-form model by optimizing the Akaike information criterion (AIC).

Step 3. Estimate reduced-form matrices $\mathbf{\Pi_k}$ by regressing each variable in $\mathbf{y}'_t$ on lags $\{\mathbf{y}'_{t-1}\cdots\mathbf{y}'_{t-k}\}$. Define the vector of forecast errors $\hat{\boldsymbol{\xi}}_t$ as the vector of the residuals of these regressions.

Step 4. Construct $\hat{\boldsymbol{\Omega}} = \mathbb{E}[\hat{\boldsymbol{\xi}}_t\hat{\boldsymbol{\xi}}_t{}']$, the estimate of the covariance matrix of the residuals.

Step 5. Set the diagonal of $\mathbf{B_{11}}$ to one, $\mathbf{B_{21}}$ to zero, and the diagonal of $\mathbf{B_{22}}$ to one.

Step 6. Construct the log-likelihood function $\mathcal{LL}(\mathbf{B})$, and estimate the parameters by maximizing the log-likelihood function.

In Step 1, we fourth difference all of our series to make them stationary. This is necessary because our quarterly time series are not stationary. Fourth differencing eliminates both trends and seasonal differences in our time series. We run augmented Dickey–Fuller tests on the differenced series to test stationarity and confirm the presence of a unit root in the original series but none in the differenced series; see Online Appendix B for the results of this procedure.

In Step 2, we determine the optimal lag length for our model. The long-run impacts of shocks depend on the number of lags used to fit the data. Hence, choosing the appropriate lag order k for the model is an important consideration, and its misspecification may lead to incorrect estimates of the impacts. We adopt a data-driven approach to determine the optimal lag length. Specifically, we vary $k \in \{1, \ldots, 10\}$, estimate the reduced form for each k, and compute the resulting AIC statistic. We then chose the optimal lag order k_{opt} to minimize AIC across all k.

In Step 3, we estimate the reduced form by running $n + m$ ordinary least square regressions. A key property of the reduced form is its *stability* or covariance stationarity so that the impacts of errors dissipate over time.

Definition 1 (Stability (Lutkepohl 2007)). The reduced-form Vector Autoregressive (VAR) is covariance stationary, or *stable*, if

$$\mathbf{det(I_{n+m} - \Pi_1 z - \Pi_2 z^2 - \cdots - \Pi_k z^k) \neq 0 \text{ for } |z| < 1},$$

or the reverse characteristic polynomial has no roots in the unit circle.

The above definition is equivalent to stating that the eigenvalues of the companion matrix $\boldsymbol{\Xi}$ have modulus less than one, where the companion matrix has the following form:

$$\boldsymbol{\Xi} = \begin{bmatrix} \mathbf{\Pi_1} & \mathbf{\Pi_2} & \ldots & \mathbf{\Pi_{k-1}} & \mathbf{\Pi_k} \\ \mathbf{I} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{I} & \mathbf{0} \end{bmatrix}.$$

Thus, we compute eigenvalues to verify stability. Consider the example of Macy's, one of the firms in our data set; with five lags, there are 35 eigenvalues of the companion matrix, and the largest eigenvalue for Macy's has modulus = 0.97.

In Step 4, we construct the sample estimate of the covariance matrix, $\hat{\Omega} = \mathbb{E}[\boldsymbol{\xi}_t\boldsymbol{\xi}'_t]$, after recovering $\boldsymbol{\xi}_t$. In Step 5, we impose the identifying restrictions using exogenous variables.

In Step 6, we compute the likelihood function for the reduced-form residuals and estimate **B** by maximum likelihood estimation. The vector of reduced-form residuals $\boldsymbol{\xi}_t$ follows a multivariate normal distribution $\mathbf{N}(\mathbf{0}, \boldsymbol{\Omega})$. The log-likelihood function is given by

$$\mathcal{LL}(\Omega) = -\frac{(n+m)T}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Omega}| - \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{\xi}_t'\boldsymbol{\Omega}^{-1}\boldsymbol{\xi}_t \quad (8)$$

$$= -\frac{(n+m)T}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Omega}| - \frac{1}{2}\sum_{t=1}^{T}Tr(\boldsymbol{\Omega}^{-1}\boldsymbol{\xi}_t\boldsymbol{\xi}_t') \quad (9)$$

$$= -\frac{(n+m)T}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Omega}| - \frac{T}{2}Tr\left(\boldsymbol{\Omega}^{-1}\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\xi}_t\boldsymbol{\xi}_t')\right) \quad (10)$$

$$= -\frac{(n+m)T}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Omega}| - \frac{T}{2}Tr(\boldsymbol{\Omega}^{-1}\hat{\boldsymbol{\Omega}}). \quad (11)$$

Note that $\boldsymbol{\Omega} = \mathbf{B}^{-1}\mathbf{B}^{-1'} = (\mathbf{B}'\mathbf{B})^{-1}$, and so, $\boldsymbol{\Omega}^{-1} = \mathbf{B}'\mathbf{B}$. Furthermore, $\log|\boldsymbol{\Omega}^{-1}| = -\log|\mathbf{B}'\mathbf{B}| = -2\log|\mathbf{B}|$. Using these, we can rewrite the log likelihood as a function of $\mathbf{B}$, by which we get

$$\mathcal{LL}(\mathbf{B}) = -\frac{(n+m)T}{2}\log(2\pi) + T\log|\mathbf{B}| - \frac{T}{2}tr(\mathbf{B}'\mathbf{B}\hat{\boldsymbol{\Omega}}). \quad (12)$$

The matrix $\mathbf{B}$ is estimated by minimizing the negative of the above log-likelihood function.

4.2. Data

We use publicly available firm-level financial and operational data from S&P's Compustat. The data span 30 years from 1990 to 2020, with quarterly frequency. They begin in 1990 when the Financial Accounting Standards Board (FASB) introduced the cash flow statement. In addition, we use (i) historical quarterly U.S. GDP data for the same period from the Bureau of Economic Analysis website[5] as well as (ii) the consumer confidence index and CEO confidence index data from the Conference Board. Our sample consists of retail (Standard Industrial Classification (SIC) codes 52–59), manufacturing (SIC codes 20–39), and wholesale (SIC codes 50–51) firms. We only keep firms that were active over the entire 30 years. This is because we need at least 41 quarterly data points for a firm to estimate the reduced-form model with 40 variables and no intercept. We restricted our sample to firms with 120 quarterly data points (30 years) to provide meaningful statistical inferences. Also, having a longer time horizon helps provide reliable estimates of the model parameters. The sample is not free from missing data. To minimize the effects of missing values, we impose that none of the six variable series for a firm have more than 5 missing values (of 120).[6] Finally, we discard data with negative sales or inventory values. The resulting sample includes 69,554 observations across 575 firms. Table 2 reports the summary statistics, and Online Appendix A presents the variable definitions.

4.3. Estimation Results

In this section, we provide evidence of contemporaneous causal relationships among the variables in our model by estimating the model on the data panel described in Section 4.2. Our model can be estimated separately for each firm to estimate firm-specific coefficients that can be utilized in decision making, or it can be estimated on a panel consisting of the data for all of the firms. In Sections 5 and 6, we estimate the model separately for each firm in order to obtain firm-specific response functions to shocks, and in this section, we present aggregate results from a joint estimation of the model across all of the firms. Using a panel data approach allows for common shocks across firms and also, allows for including firms that do not have a complete time series for the sample duration of 30 years. For this, we estimate the covariance matrix $\boldsymbol{\Omega}$ jointly for all of the firms and then, compute a pooled $\mathbf{B}$ matrix.[7] The contemporaneous relationships are encoded in $\mathbf{B}$, and the standard errors for the elements of $\mathbf{B}$ are calculated as $\sqrt{\text{diag}(\mathbf{H}^{-1})}$, where diag is the diagonal operator and $\mathbf{H} \in \mathbb{R}^{36\times 36}$ is the Hessian matrix of the log-likelihood function at convergence. We determine whether the parameters' estimates are significantly different from zero using a $p = 0.01$ Wald test. Table 3 shows the resulting estimates along with the standard errors. We find that all of our proposed linkages are statistically significant, providing direct evidence for the variables' endogeneity.

We also present the estimates of contemporaneous impacts for two-digit-SIC categories to allow for an interpretation of coefficients. For this, we estimate

Table 2. Sample and Business-Segment Summary Statistics

Statistic	Full sample	Retail	Wholesale	Manufacturing
# Observations	69,554	5,821	4,120	59,613
# Firms	575	48	34	493
# Industries (2-digit SIC)	29	7	2	20
Median Sales ($ million)	217	895	303	185
Median Inventory ($ million)	120	690	174	102
Median Cash Flow ($ million)	17	53	7	16
Median Accounts Payable ($ million)	57	214	80	47
Median SGA ($ million)	41	218	42	34

Notes. All financial and operational variables are quarterly and measured in millions of dollars. Note that quarterly cash flow can be negative.

Table 3. Estimates of **B** for the Sample

	Internal firm variables				
Variable	Sales	Inv	Cash	AP	SGA
Sales	1.00000	0.11581	0.07730	0.07290	0.06090
	(0.00000)	(0.00033)	(0.00026)	(0.00018)	(0.00053)
Inv	0.08111	1.00000	0.10000	0.09004	0.11161
	(0.00012)	(0.00000)	(0.00026)	(0.00018)	(0.00062)
Cash	0.06954	0.10276	1.00000	0.10844	0.15780
	(0.00012)	(0.00033)	(0.00000)	(0.00019)	(0.00062)
AP	−0.04077	0.10096	0.10594	1.00000	0.11309
	(0.00012)	(0.00032)	(0.00027)	(0.00000)	(0.00061)
SGA	0.07675	0.06909	0.10735	0.13045	1.00000
	(0.00010)	(0.00033)	(0.00027)	(0.00019)	(0.00000)

Notes. All of the linkages are significantly different than zero with respect to $p = 0.01$. The p-values are calculated from a Student's t distribution with degrees of freedom = sample size − number of covariates.

the model for each two-digit-SIC separately. Table 4 presents our estimates. For brevity, we only focus on the interactions between three of the endogenous variables: sales, inventory, and cash flow.

The Inventory (I) → Sales (S) estimates are positive for 20 of 25 SIC groups, indicating the positive contemporaneous reinforcement of inventory on sales. Our result is consistent with that in Kesavan et al. (2010). In empirical research of item-level inventories, researchers have documented both a demand-stimulating effect of inventory and a scarcity effect of inventory (Cachon et al. 2019). Our method does not disentangle these mechanisms but shows that the aggregate estimate is primarily positive. The S → I estimates are positive for 13 of 25 SIC groups, showing a positive contemporaneous association in the reverse direction. This result is also consistent with prior literature. At the item level, this result is in line with standard inventory models, like the economic order quantity model and stochastic inventory models (with mild assumptions on the probability distribution of demand).

The relationships of cash with sales and inventory are new in our paper. The C → I estimates are negative for 17 of 25 SIC groups. The negative effect on inventory may occur because the firms are able to invest more in inventory management and optimization policies through the use of extra cash flows; we do not observe an increase in inventory with an increase in cash, likely because the firms in our sample are large firms that are not cash constrained.

The S → C estimates are positive for 20 of 25 SIC groups. The effect is consistent with the cash flow accounting identity (1); an increase in sales increases cash flow in the same period, and an increase in inventory decreases cash flow. However, although the identity (1) has all coefficients equal to one, none of the estimated effects in our model are equal to one because of the contemporaneous feedback between the variables.

Overall, the estimates from our model show that cash flow is endogenous with sales and inventory. Because our model also includes lagged relationships, we now use impulse response functions to compute the short- and medium-term effects of the system of variables on each other.

5. Application to Cash Flow Management

In this section, we demonstrate the value of using the structural model developed in Section 3 for operational decision making. Because our model includes both contemporaneous and lagged covariates, it enables us to estimate the impact of a unit shock in any variable in period t on the entire system in periods $t, t+1, \ldots$. These impacts are given by the impulse response functions. With the knowledge of IRFs, a firm can determine what mitigating operational decisions to take to optimize its future performance. First, we identify the properties of structural shocks and show how we can compute the new equilibria induced by structural shocks in the system. Then, we demonstrate using a stylized example of competitor entry how to use this tool to evaluate compensating managerial actions or policy decisions against shocks. Finally, we show how the model can be used to estimate the impact of macro-level phenomena, such as periods of economic downturn and improvement in economic sentiment, on firm performance.

5.1. Determining the New Equilibria for Variables: The Case of a Competitor's Entry

We first briefly describe the construction of the IRFs from our estimated model by using the Wold representation of the system. Consider the VAR representation of the structural model in (3). By repeated substitution, we arrive at the following representation of the stable reduced-form VAR:

$$\begin{aligned}\mathbf{y_t} &= \mathbf{\Psi_0}\boldsymbol{\xi}_\mathbf{t} + \mathbf{\Psi_1}\boldsymbol{\xi}_\mathbf{t-1} + \mathbf{\Psi_2}\boldsymbol{\xi}_\mathbf{t-2} + \ldots \\ &= \mathbf{\Psi}(L)\boldsymbol{\xi}_\mathbf{t},\end{aligned} \tag{13}$$

where $\mathbf{\Psi}(L) = \sum_{m=0}^{\infty}\mathbf{\Psi}_m L^m, \mathbf{\Psi_0} = \mathbf{I}$, and $\mathbf{\Psi_m} = \sum_{j=1}^{m}\mathbf{\Psi_{m-j}}\mathbf{\Pi_j}$. Here, $\mathbf{\Psi_m}$ are the dynamic multipliers of the VAR system, and $\mathbf{\Pi_j}$ are the estimated lagged reduced-form matrices. $\mathbf{\Pi_j}$ is set to zero for $j > k$, where k is the number of lags included. The corresponding moving average (MA) representation of the structural form can be obtained by substituting $\boldsymbol{\xi}_\mathbf{t} = \mathbf{B^{-1}u_t}$ in (13), by which we get

$$\begin{aligned}\mathbf{y_t} &= \mathbf{\Psi}(L)\mathbf{B^{-1}u_t} \\ &= \mathbf{\Theta}(L)\mathbf{u_t},\end{aligned} \tag{14}$$

where $\mathbf{\Theta}(L) = \sum_{m=0}^{\infty}\mathbf{\Theta}_m L^m$, $\mathbf{\Theta_0} = \mathbf{B^{-1}}$, and $\mathbf{\Theta_m} = \sum_{j=1}^{m}\mathbf{\Theta_{m-j}}\mathbf{\Gamma_j}$. The values of $\mathbf{\Theta}$ give us the IRFs to lagged structural shocks occurring at different time periods. Intuitively, each IRF gives the effect of a shock, such as a change in demand or an adjustment to the inventory

Table 4. Industry-Specific Estimates

Industry	SIC	R/W/M	I → S	C → S	S → I	C → I	S → C	I → C
Hardware and garden	52	R	0.630	0.405	−0.13	−0.20	0.351	−0.12
General merchandise	53	R	1.409	−1.60	−0.72	−0.25	−0.03	1.409
Food stores	54	R	−0.00	−1.06	−0.07	−0.17	−0.06	0.350
Automotives	55	R	−0.02	−0.22	−0.23	0.047	0.316	0.250
Apparel	56	R	2.991	0.119	0.026	0.114	0.255	0.261
Home furnishings	57	R	0.304	0.073	0.396	−0.28	−0.11	−0.24
Durables	50	W	1.464	0.132	0.085	0.006	−0.35	0.873
Nondurables	51	W	0.732	−1.42	−0.31	−0.56	0.059	0.144
Tobacco	21	M	0.138	−0.06	0.156	−0.03	−0.26	0.676
Textile mill	22	M	0.510	−0.33	−0.11	−0.20	0.253	0.307
Apparel	23	M	0.545	−0.27	−0.08	0.246	1.139	0.216
Wood	24	M	−0.16	−0.52	0.334	−0.11	0.160	−0.06
Furniture	25	M	0.943	0.131	0.147	−0.06	0.134	0.643
Paper	26	M	0.359	−0.11	−0.00	−0.12	0.157	−0.26
Printing and publishing	27	M	0.918	−0.26	−0.04	−0.26	0.182	0.170
Chemicals	28	M	0.661	0.028	0.035	0.008	0.043	−0.11
Petroleum	29	M	2.431	−0.17	−0.00	0.047	0.051	0.867
Rubber and plastics	30	M	0.757	0.315	0.352	0.263	0.138	0.375
Leather goods	31	M	0.111	−0.13	0.868	−0.19	0.075	0.014
Stone and glass	32	M	−0.01	−0.06	0.097	−0.13	0.415	−0.40
Primary metals	33	M	−1.83	−0.14	−0.13	0.060	0.192	0.062
Fabricated metals	34	M	0.902	−0.05	0.137	−0.15	0.351	−0.11
Industrial machinery	35	M	0.508	0.142	0.114	−0.21	0.140	0.157
Electronics	36	M	1.409	−0.34	−0.07	−0.32	0.187	−0.18
Transport	37	M	0.007	−0.19	0.379	−0.00	0.177	0.188

Notes. Industries are defined by two-digit SICs. The SIC's time series are a sum of individual firms' time series. Columns on the right contain the estimated parameters for endogenous relationships. Industries are separated into retail (R), wholesale (W), and manufacturing (M) SICs.

policy, on the future state of the system so that IRFs enable us to compute the impact of such changes.

This MA representation is helpful to arrive at a new equilibrium for the system of variables in response to a structural shock to any of the endogenous variables. For example, suppose that a new competitor enters the industry. In that case, the competitor's entry provides a negative structural shock to the focal firm's sales, which would permanently affect the future performance of the firm. This is because the shock would lead to a new equilibrium of the firm's variables. We demonstrate this with a stylized example of a sales-inventory model.

Consider the moving average representation a 2×1 vector $\mathbf{y_{t+s}}$ containing the focal firm's sales and inventory:

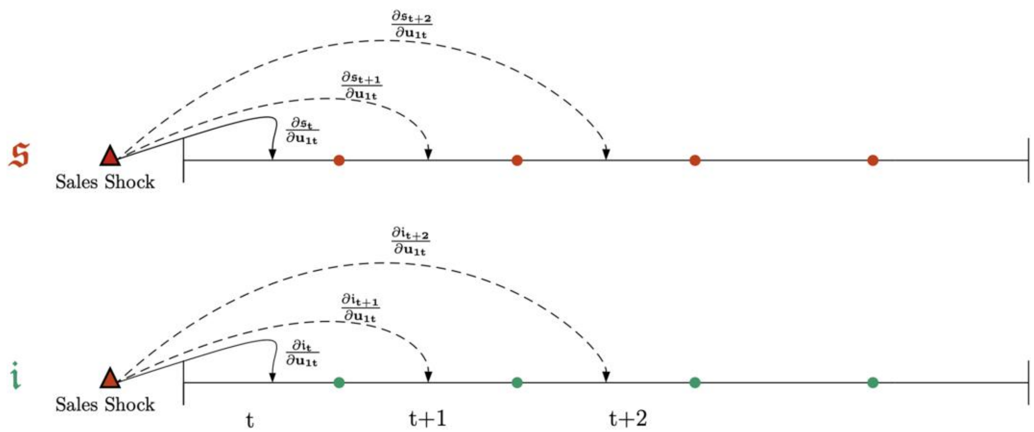
$$\begin{pmatrix} \mathfrak{s}_{\mathbf{t+s}} \\ \mathbf{i_{t+s}} \end{pmatrix} = \underbrace{\begin{bmatrix} \theta^s_{11} & \theta^s_{12} \\ \theta^0_{21} & \theta^0_{22} \end{bmatrix}}_{\text{Contemporaneous}} \begin{pmatrix} \mathbf{u_{1t+s}} \\ \mathbf{u_{2t+s}} \end{pmatrix}$$
$$+ \cdots + \underbrace{\begin{bmatrix} \theta^s_{11} & \theta^s_{12} \\ \theta^s_{21} & \theta^s_{22} \end{bmatrix}}_{\text{Dynamic}} \underbrace{\begin{pmatrix} \mathbf{u_{1t}} \\ \mathbf{u_{2t}} \end{pmatrix}}_{\text{Competitor's Entry at } t} + \cdots;$$

here, the element θ^s_{ij} can be interpreted as the impact of a unit shock in variable j at any time t on variable i at time $t+s$. These impacts fully characterize how the focal firm's variables evolve in response to the shock. Specifically, the changes in sales and inventory because of unit shocks in these variables (see Figure 4) s periods in the future are

$$\frac{\partial \mathfrak{s}_{\mathbf{t+s}}}{\partial \mathbf{u_{1t}}} = \theta^s_{11}, \quad \frac{\partial \mathfrak{s}_{\mathbf{t+s}}}{\partial \mathbf{u_{2t}}} = \theta^s_{12}, \ldots$$
$$\frac{\partial \mathbf{i_{t+s}}}{\partial \mathbf{u_{1t}}} = \theta^s_{21}, \quad \frac{\partial \mathbf{i_{t+s}}}{\partial \mathbf{u_{2t}}} = \theta^s_{22}, \ldots. \tag{15}$$

We simulate how the variables evolve when subjected to the structural shock caused by a competitor entry. For this simulation, suppose the initial sales and inventory levels are \$1,000,000 and \$250,000, respectively: $\mathbf{B}^{-1} = \begin{bmatrix} 0.94 & 0.23 \\ -0.23 & 0.94 \end{bmatrix}$ and $\mathbf{\Pi_1} = \begin{bmatrix} 0.25 & 0.75 \\ 0.1 & 0.25 \end{bmatrix}$. Suppose that a competitor entry results in a negative structural shock of -0.1 (i.e., -10%) in sales in the first quarter $\left(\text{i.e., } \mathbf{u_1} = \begin{bmatrix} -0.1 \\ 0 \end{bmatrix}\right)$. Because our model is in the differences of variables and not the levels, we trace the system's evolution using the logged difference in the variables. Table 12 in Online Appendix B shows the evolution of the firm's sales and inventory levels when subjected to the negative sales shock. This evolution is the firm's impulse response to the competitor's entry. Intuitively, the impulse response function traces the sales and inventory levels as they settle to a new equilibrium (\$900,088, \$255,319) from (\$1,000,000, \$250,000). We see how the negative shock because of the competitor's entry leads to a long-run reduction in sales and increase

Figure 4. (Color online) Impacts of a Unit Dollar Shock in Sales on Sales and Inventory



Notes. The solid arrows represent contemporaneous impacts. The dashed arrows represent impacts in the future periods.

in inventory. This is the new equilibrium if the firm has not taken any compensatory action to prevent the arrest of declining sales because of the competitor's entry. The determination of this equilibrium assumes that the data-generating process stays the same (i.e., the values of the structural parameters and the lagged coefficients remain the same under the competitor's entry).

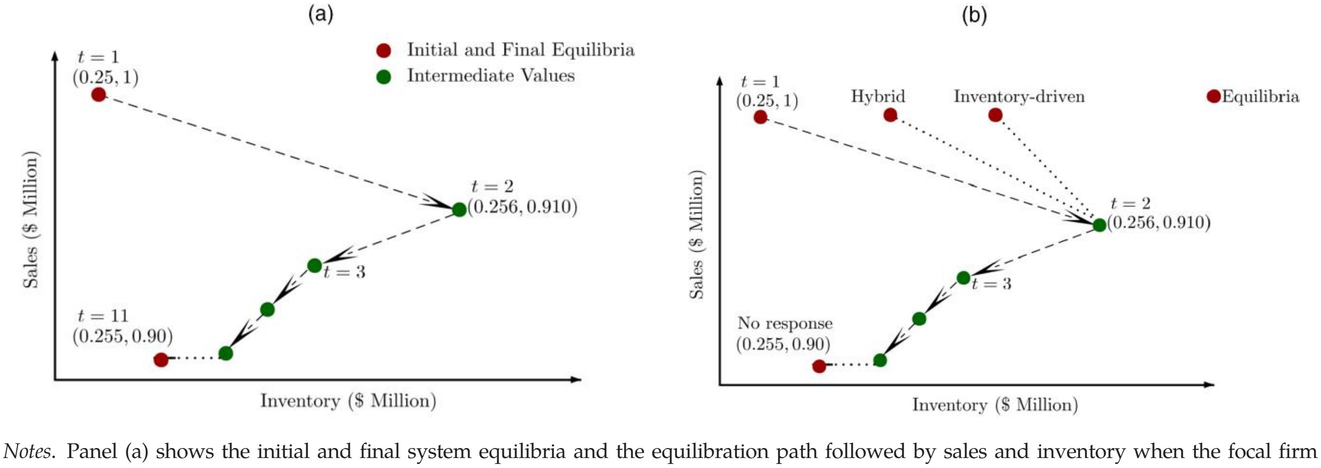
Figure 5(a) shows the initial and final equilibria and the evolution of the two variables. The new equilibrium (900,088, 255,319) is reached reasonably quickly. The existence of such equilibria follows directly from the properties of simultaneous equations models. The covariance stability of the reduced-form VAR guarantees the uniqueness of the equilibrium, which ensures that the fresh impacts in the model eventually die out in the long run. At the new equilibrium, the per-period impacts of the shocks are zero. Table 13 in Online Appendix B shows the exact intermediate numerical values of sales and inventory.

5.2. Mitigating Actions

The firm can take compensating *managerial actions or policy decisions* to mitigate the effects of structural shocks. In the above example, the incumbent firm facing a competitor entry can increase its inventories to recover the loss of market share. Our model provides a tool to evaluate such managerial actions by estimating their impact on the system of variables. To illustrate this, let us consider two mitigating actions that a firm can take: (i) increase inventory levels in the next period and (ii) use a combination of sales-boosting efforts and inventory increase in the next period.

1. Increase inventory in the next period. Let us assume that the firm increases inventory in the next period to counteract the shock provided by the

Figure 5. (Color online) System Equilibria With and Without Mitigating Actions



Notes. Panel (a) shows the initial and final system equilibria and the equilibration path followed by sales and inventory when the focal firm receives a sales shock from a Competitor's entry. Panel (b) shows the system equilibria (I-S) for hybrid and inventory-driven strategies implemented at $t = 2$ (dotted lines). The no-response equilibrium is at (0.255, 0.90). (a) Equilibration path followed by sales and inventory. (b) Mitigating actions taken by the firm.

competitor's entry altogether. To do so, the positive inventory shock needed in the next quarter is 0.075. This shock leads to a new sales-inventory equilibrium [100,000, 276,293]. This means that the competitive response leads to an equilibrium that can mitigate the sales shock but leads to a permanent increase in the firm's inventory level. Thus, the competitor's entry directly impacts the incumbent firm's cost and profitability.

2. Hybrid strategy—sales boosting and inventory increase. Let us assume that the firm can only increase its inventory levels by a little. Instead, it relies on a hybrid strategy of investing in sales-boosting efforts and partial inventory increases. For sales boosting, the firm may offer promotions, discounts, and incentives to encourage purchases. A shock profile of 0.05 to sales and 0.0375 can offset the sales shock entirely and lead to a new equilibrium [1,000,000, 262,818]. Thus, the hybrid strategy also leads to a permanent increase in inventory levels. However, the extra amount of inventory held in the system is less than strategy (1), which relies purely on an inventory-based response.

The kind of strategy to employ depends on the costs associated with increasing inventories (ordering, holding cost, etc.) and investing in sales-boosting efforts (markdowns, promotion costs, etc.) and the time needed to employ these strategies. Our model gives a clear picture of the benefits of each type of strategy, which can further be augmented with the cost parameters to arrive at the best strategy.

5.3. How Recessions and Global Economic Sentiments Affect Firm Performance

We next build on the stylized example of Section 5.2 to show the application of our model to predict and manage the effects of economic shocks on a firm. Economic recessions can severely impact a firm's operational and financial performance. Historically, macroeconomic shocks have led to disruption of supply chain operations, plummeting sales and cash flows, depletion of cash reserves, and in the worst case, businesses going bankrupt. A case in point is the COVID-19 pandemic, during which many firms witnessed some of the consequences mentioned above to various degrees (Bartik et al. 2020). For example, firms like J. C. Penny, Latam Airlines, and more than 350 other large firms filed for bankruptcy. On the other hand, positive sentiments about the economy can boost firms' sales. When consumer confidence in the economy improves, people are more likely to spend money, save less, and spend on luxurious items. This boost in consumer spending can directly lead to an increase in firms' sales and revenues, especially those in the consumer-facing sectors.

In this section, we first illustrate the use of our model to compute IRFs for Macy's, Inc.; the choice of the company is just for illustration. We then estimate the impact of macro-level phenomena on Macy's firm performance considering the 2001 dot-com bust, the 2008–2009 Great Recession, and the economic uncertainty and decline in consumer sentiment in 2022 because of fears of recession and inflation. Our results are novel as they enable the estimation of the impact of economic phenomena on any given firm.

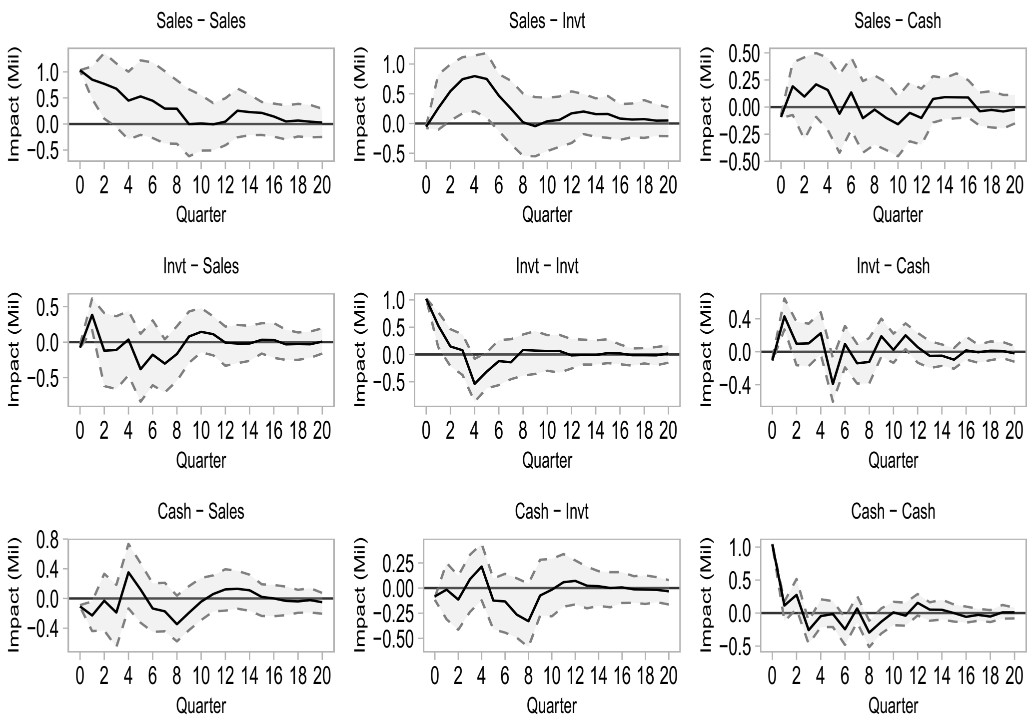
5.3.1. IRFs for Macy's, Inc. We estimate IRF matrices for the system of variables for Macy's by estimating our model for the company and then, applying the procedure outlined in Section 5.1. In Figure 6, we present the salient impulse responses for Macy's. The title of each panel in Figure 6 displays the impulse variable followed by the response variable. For example, sales-cash in the top right panel denotes the impact of a unit shock in sales on cash in the current and future periods. We extrapolate the response functions for five years into the future, and the x axis in each panel in Figure 6 represents time in quarters. All of the impacts shown are per-period (new) impacts in the future. The cumulative impacts until time t can be computed by adding the per-period impacts until t. We observe from Figure 6 that even a small shock to a variable has a long-term impact on the system of variables. Most new impacts start converging to zero in the very long run as the system equilibrates again. We emphasize that the long-term impacts are governed by the lagged relationships among the variables and not because of mere seasonality. We note the following observations from Figure 6.

5.3.1.1. Shock to Sales. A unit shock in sales results in an immediate increase in sales of about one unit. The shock has a significant effect that lasts for 20 quarters. The shock also has a positive impact on inventory, both in the short run and in the long run. A shock to sales leads to an increase in inventory in the short run. The impact is statistically significant until six quarters. The positive effect of sales shocks on inventory is consistent with the previous empirical findings in OM (see, for example, Kesavan et al. 2010).

5.3.1.2. Shock to Inventory. A unit shock to inventory results in an increase in sales in the next quarter. This demonstrates the positive reinforcement of inventory on sales. The impact on cash flow is negative contemporaneously as in the accounting identity. But, there is a positive and significant impact in the next period. Hence, inventory affects both sales and cash flows with a lag of order one. Overall, we find that positive shocks to inventory increase cash flows in the long run.

5.3.1.3. Shock to Cash. Finally, a shock to cash flow dissipates very quickly. There is a significant impact contemporaneously and in the next quarter, suggesting that large cash infusions are helpful in the short run.

Figure 6. IRF Estimates for Self-Variable and Crossvariable Impacts for Macy's, Inc.



Note. The *x* axes represent time in quarters, and the *y* axes represent the point impacts.

However, in the long run, cash infusions do not have a significant impact. This runs counter to conventional wisdom, and it suggests that operational improvement in sales and inventory is more beneficial to a firm's long-run cash flows as compared with significant cash infusions themselves.

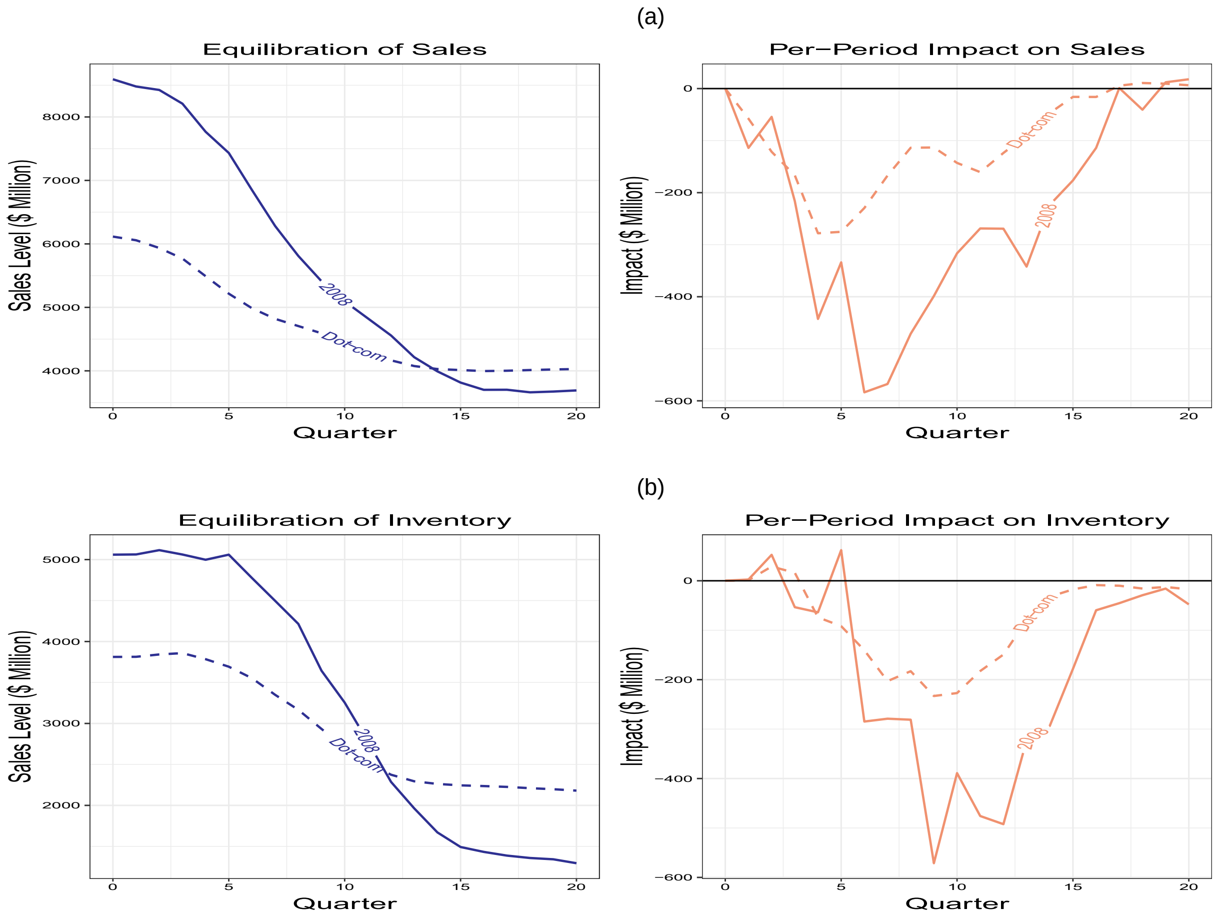
5.3.2. Case Study 1: Periods of Recession. We study the impact of the *dot-com bust of 2001* and the Great Recession of 2008–2009 on firms' sales and inventories with Macy's as a case in point. The economic shock because of the dot-com bust lasted 9 months from March to November 2001, and the shock because of the Great Recession lasted 19 months from December 2007 to June 2009.[8] To analyze the impacts of the recessions, we use the decline in the GDP growth rates across the recessions as structural shocks. Because our data have a quarterly frequency, we construct quarterly structural shocks to the GDP growth rate. The impact of the recessions over time is calculated as the sum of the impulse responses to the consecutive quarterly shocks. Table 5 summarizes the differences in the two recession periods and the corresponding GDP shocks.

Figure 7 shows the impacts of the two recessions on Macy's sales and inventories. The impact of the dot-com recession on Macy's sales is calculated as ($-0.72 \times \mathrm{IRF}_1 - 1.25 \times \mathrm{IRF}_2 - 0.718 \times \mathrm{IRF}_3 - 0.55 \times \mathrm{IRF}_4$), where the coefficients are the structural shocks to the GDP growth rate in the first, second, third, and fourth quarters of 2001 and where adding the four impulse responses gives us the impact on sales in consecutive quarters. The corresponding impact of the Great Recession is calculated as ($-1.42 \times \mathrm{IRF}_1 - 0.16 \times \mathrm{IRF}_2 + \cdots -1.344 \times \mathrm{IRF}_6$). For both of the recessions, the new sales equilibria are achieved after almost four years (Figure 7(a), left panel), by which the quarterly sales revenue of Macy's would be severely impacted (33% and 50% reduction for the dot-com bust and 2008 recessions, respectively). Note that because the IRFs are predictions to the specific shocks, they show the magnitude and longevity of the impact of each recession if a firm did not take any corrective action.

Table 5. Contrasting the Two Recessions

Time	Dot-com bust March to November 2001	The Great Recession December 2007 to June 2009
No. of quarters impacted	4	6
GDP shock vector	[−0.72, −1.25, −0.718, −0.55]	[−1.42, −0.16, −0.86, −2.86, −0.92, −1.344]

Figure 7. (Color online) Impact of Recessions on Macy's Sales and Inventory



Notes. Panel (a) shows the impacts of the 2001 dot-com and 2008 Great Recession on Macy's Sales. The left panel in panel (a) shows the equilibrium path of quarterly sales level following the recession. The right panel in panel (a) traces the per-period impact on sales. Panel (b) shows the corresponding impacts of the two recessions on Macy's Inventory.

Corrective actions can be added to this analysis and evaluated as shown in Section 5.2. Also, note that the rate of decline in sales was higher for the Great Recession than the dot-com bust.

We carry out a similar analysis for the impact of the two recessions on Macy's inventories. We find that the initial decline in the inventories is relatively slower as compared with sales (Figure 7(b), left panel), which is expected as the sales slump after the recession leads to unsold inventory. Also, we note that the per-period declines in inventory are out of phase of the per-period declines in sales, which is consistent with the fact that the inventory and sales series themselves are out of phase as shown in Figure 1. IRFs for the other endogenous variables can be computed similar to these computations.

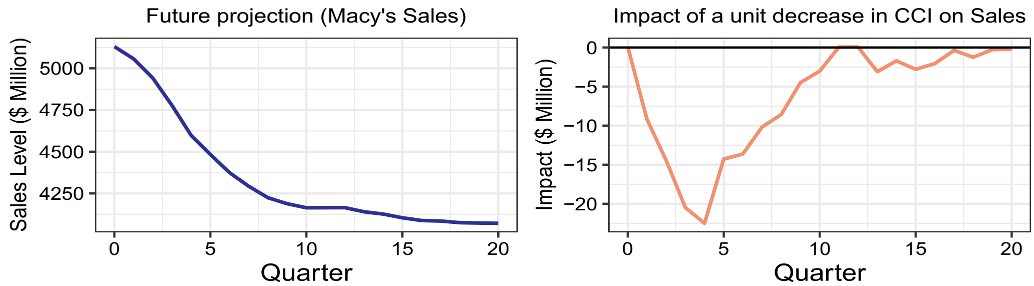
5.3.3. Case Study 2: Changing Consumer Sentiment.

As the 2022–2023 time period was characterized by economic uncertainty, the CEO of Macy's said in a press release citing its Q2 2023 earnings:

> We continue to see uncertainty in the macroeconomic environment. We are leveraging our robust data science tools to refine inventory composition, while reading and reacting to shifting consumer preferences to meet demand.

In fact, Macy's quarterly sales were down from $5,600 million U.S. dollars (USD) in Q2 2022 to $5,130 million USD in Q3 2023 because of the declining economic climate. In this section, we construct a future projection of the impact of economic uncertainty on Macy's by using the consumer confidence index as a measure of shock. We begin the projection from Q2 2023 based on a shock caused by seven points of decline in CCI. Figure 8, left panel predicts a further decline in Macy's sales as a result of this decreased index to a level of $4,777 million USD in Q2 2024, and Figure 8, right panel shows the corresponding per-period impacts of a unit decrease in CCI on sales. As noted earlier, a firm can utilize such projections to manage its corrective actions, such as boosting SGA expenses or inventory.

Figure 8. (Color online) The Future Projection of the Impact of the Decreased Consumer Confidence on Macy's Sales (Left Panel) and the Per-Period Impacts of a Unit Decrease in CCI on Sales (Right Panel)



6. Application to Cash Flow Forecasting

Forecasting OCF is an essential activity for a firm. Firms use cash flow forecasting for many purposes. Short-term forecasts for up to 1 month are used to schedule receipts and disbursements and to identify potential shortfalls; medium-term forecasts for 2–3 months are used to make operational decisions for production, stock levels, prices, and marketing spend; and longer-term forecasts for 3–12 months are required by top management, lenders, and other stakeholders to gain visibility over debt covenant positions at key reporting dates.[9] Forecasts are used to manage liquidity risk, plan future growth, and invest surplus cash. According to the Financial Accounting Standards Board, a primary objective of financial reporting is to help investors, lenders, and other stakeholders assess the amount, timing, and uncertainty of future expected cash flows (FASB 2010). However, cash flow forecasting is a challenging problem because of the complex correlation between the operational variables and the timing mismatches between revenues and the actual realization of cash flow, leading to high variability in cash flows. In this section, we use the reduced-form specification of our model to show that operational variables can play a key role in improving the forecasts for cash flows, and in turn, OCF is a useful input variable in forecasting operational variables. We first demonstrate that cash flows have high prediction error; then, we describe our forecasting procedure, and we conclude by presenting our results from out-of-sample prediction and discuss insights.

6.1. Cash Flow Variability

The quarterly forecasts of cash flows have a high forecast error. To evaluate this, we generate forecasts of cash flows using the best-fitted ARIMA model on a rolling horizon for a cross-section of firms. The resulting mean absolute percentage error for cash flow forecasts is of the order of 50%–60% (see Figure 9(a)), whereas the MAPE for similar models for sales forecasts is about 4%–5%. This observation is consistent across multiple firms and in line with prior accounting literature (Lorek and Willinger 2009, Nallareddy et al. 2020). The high forecast error for cash flows arises because the individual components of cash flows in (1) are themselves volatile. Figure 9(b) shows the four-quarter moving average of operating cash flow for Macy's, Inc. for 1990–2020, along with the corresponding values of operating profit, change in inventory, and accounts payables. The moving average of cash flow has a high coefficient of variation of 0.99. The quarterly cash flows are even more volatile than the moving average. Moreover, cash flows are correlated with operating profit (= sales − cost of goods sold − selling, general and admin expenses) and changes in inventory and accounts payables, with correlation coefficients of 0.91, 0.24, and 0.08, respectively. This high volatility at the quarterly level and the dependence on stochastic components make forecasting cash flows challenging and also, present an opportunity to improve forecasting accuracy.

6.2. Estimation Procedure

We use the reduced-form VAR to generate one-step-ahead forecasts for the multivariate time series.[10] We calculate one-step-ahead forecast values of $\mathbf{y_t}$ based on the available information until time t. Let the one-step-ahead forecast be $\mathbf{y}^*_{\mathbf{t+1}} = \mathbf{y_{t+1|t}}$. The best linear predictor of $\mathbf{y}^*_{\mathbf{t+1}}$ is given by

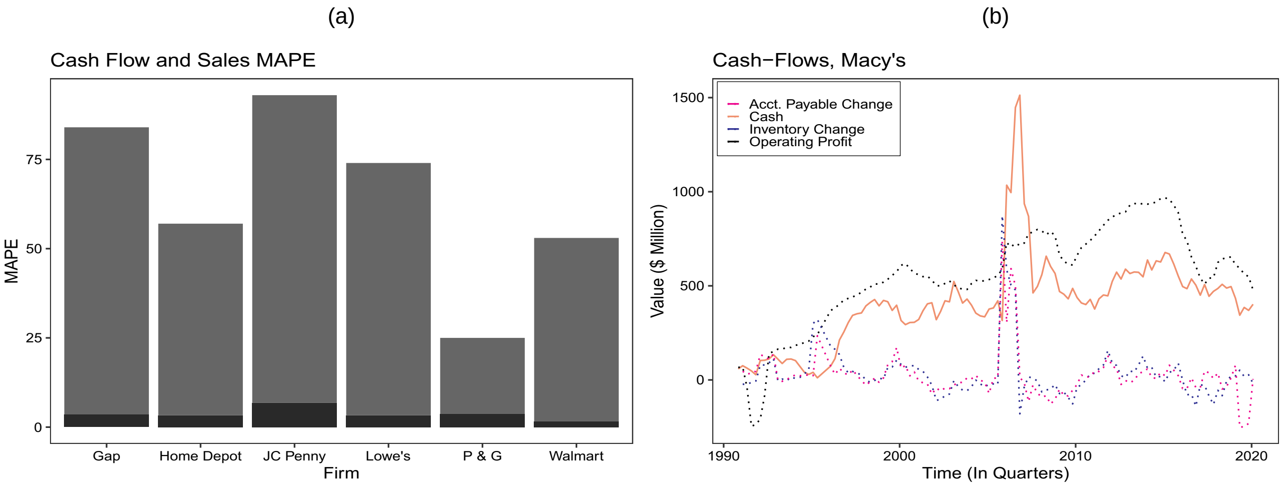
$$\mathbf{y}^*_{\mathbf{t+1}} = \mathbf{\Pi_1 y_t} + \mathbf{\Pi_2 y_t} + \cdots + \mathbf{\Pi_k y_{t-k+1}}, \tag{16}$$

where $\mathbf{\Pi} = \{\mathbf{\Pi_1}, \ldots, \mathbf{\Pi_k}\}$ are the estimated reduced-form matrices. $\mathbf{y}^*_{\mathbf{t+1}}$ minimizes the mean squared error (MSE) $= \mathbb{E}[(\mathbf{y_{t+1|t}} - \mathbf{y_{t+1}})^2]$. We measure the predictive accuracy of our forecasts by evaluating the mean absolute percentage error of the *predicted* values with respect to the *actual values*. Specifically, for variable $i \in Y$, the MAPE measure is defined as

$$\text{MAPE}_i^{\text{VAR}} = 100 \times \mathbb{E}\left[\frac{|y_{t+1|t} - y_t|}{y_t}\right]. \tag{17}$$

We generate forecasts on a rolling-window basis. First, we estimate reduced-form matrices using the first t_{test} quarters of data. Then, we generate the one-step-ahead

Figure 9. (Color online) Cash Flow Forecasts and Moving Averages of Its Components



Notes. Panel (a) shows the MAPE for cash flow (Gray) and sales (Black) forecasts for different retailers. Panel (b) shows the four-quarter moving average of cash flows for Macy's, Inc. (red line).

forecasts $\mathbf{y}_{t_{test}+1|t_{test}}$. Subsequently, we use the data from the first $t_{test}+1$ periods and repeat the procedure. We set $t_{test}=80$ (1990–2010) for our main results and compare the MAPE of the predicted values of three main endogenous variables—sales, inventory, and cash flow—with the corresponding MAPE predicted by an autoregressive model. The autoregressive model is applied in a rolling-window manner as well using the same lags as the VAR model to make the model predictions comparable. We evaluate the MAPE over the last 40 quarters. We run this forecasting procedure for different multivariate systems defined by the number of time lags and the subset of variables included. Specifically, we start with the binary system of lag 1 composed the variables sales and inventory, and we compute the MAPE from the two models. We then repeat the procedure for higher lags of the binary system. Then, we add the remaining variables to the model one by one and compute the corresponding MAPE from both models. The purpose of this procedure is to test the robustness of our results and generate comparative insights across a wide range of model specifications.

6.3. Results and Discussion

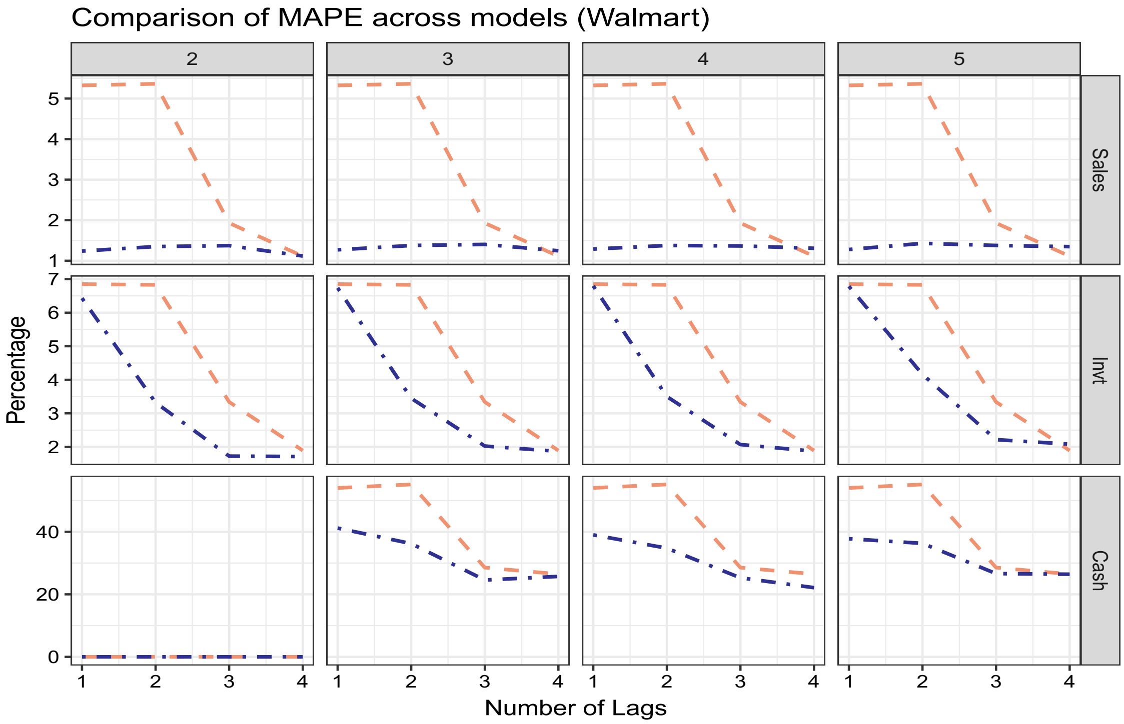
We first present out-of-sample forecasting results for all ofthe firms using both the Autoregressive (AR) and VAR models. The analysis is done on a rolling-horizon basis. We run the VAR analysis for four different lag values and four sets of variables, and we compute the lowest MAPE across these 16 specifications, which is denoted as MAPE_{VAR}.[11] Similarly, we compute the lowest MAPE for the AR models across four different lag values ($k \in \{1,2,3,4\}$) and call it MAPE_{AR}. Table 6 reports the distribution of the MAPE improvement, $\text{MAPE}_{\text{AR}} - \text{MAPE}_{\text{VAR}}$, for sales, inventory, and cash flows across all firms winsorized at the 1st and 99th percentiles. We find that the VAR model achieves an improvement in sales, inventory, and cash flow forecasts for about 68%, 67%, and 63% of the firms in the sample, respectively. The average improvement in sales MAPE is 0.74%, the average improvement in inventory forecasts is 0.14%, and the cash flow forecasts show the largest improvement of 19.13%. These observations make a strong case that adding cross-sectional variables in the forcasting model can improve performance. However, it is also worth noting that further improvements in forecast accuracy should be possible by using more sophisticated machine learning models instead of a linear regression model.

To develop further insights, we present the results of our forecasting exercise for Walmart and Home Depot in Figures 10 and 11, respectively. The columns in Figures 10 and 11 denote the number of variables in the system, starting from the two-variable system of sales and inventory and then, adding cash, accounts payable,

Table 6. Distribution of MAPE Improvement for Sales, Inventory, and Cash Flows

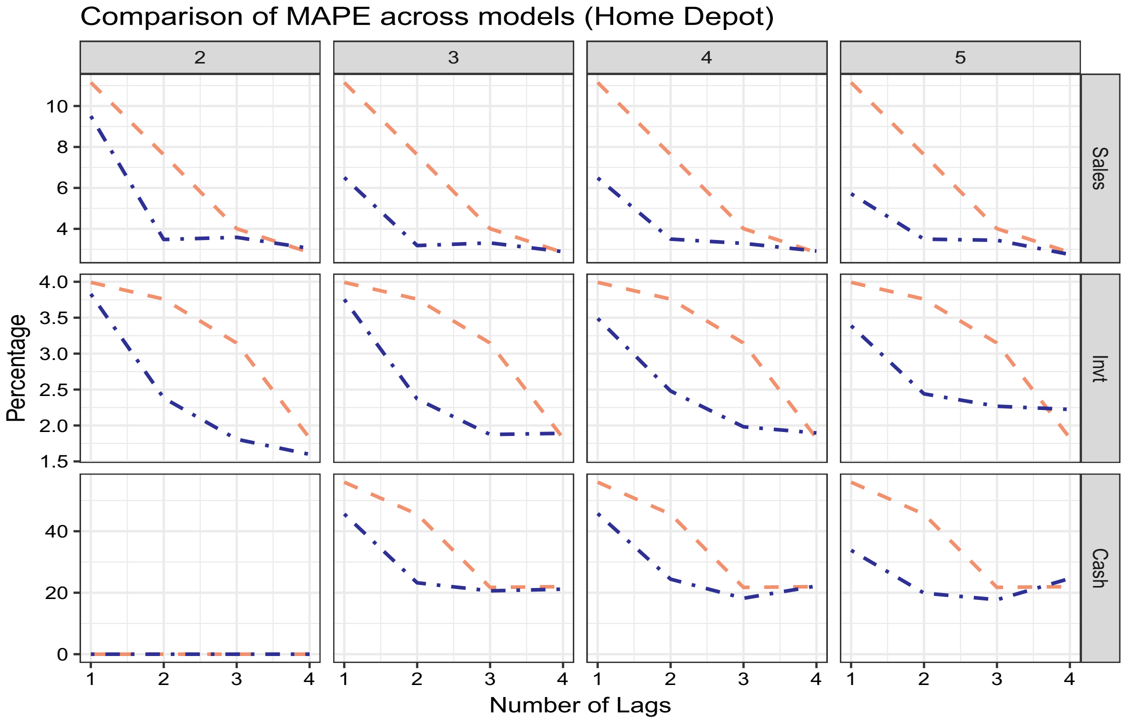
Variable	1st quarter	Median	Mean	3rd quarter	No. of firms	Age firms (better MAPE), %
Sales	−0.11	0.28	0.74	0.81	543	68
Inventory	−0.10	0.21	0.15	0.69	525	67
Cash	−3.46	3.97	19.13	23.93	526	63

Figure 10. (Color online) MAPE Values for Endogenous Variables Across Different Model Specifications for Walmart



Note. Dashdotted (dashed) line denotes the MAPE for VAR (autoregressive) model.

Figure 11. (Color online) MAPE Values for Endogenous Variables Across Different Model Specifications for Home Depot



Note. Dashdotted (dashed) line denotes the MAPE for VAR (autoregressive) model.

and SGA expenses sequentially.[12] The y axes and the x axes in the plots denote MAPE and the number of lags in the model. In total, we evaluate 4×4 model specifications for each firm.

First, consider sales forecasts. We find that the VAR model outperforms the autoregressive model across all specifications for lags 1, 2, and 3. For the lag 4 model, both models perform similarly. The top left panels in Figures 10 and 11 show the immediate improvement in sales forecast in the two-variable model of sales and inventory. The improvement is in line with the previous literature (Kesavan et al. 2010), which documents that adding inventory information improves sales forecasts. Second, inventory forecasts are also better with joint forecasting across a wide range of models. The most significant improvement in inventory forecasts occurs with the inclusion of lagged values of sales. In the basic two-variable setup, adding lagged sales values increases the accuracy of the forecasts drastically. This finding is also in line with the previous literature, demonstrating how sales information with different lead time signals impacts inventory. Third, we also document a significant improvement in cash flow forecasts across the models. Quarterly cash flows are highly variable, and the corresponding forecasting problem is hard. The lowest MAPE achieved across different models for cash flow forecasting is of the order of 40%, which is significantly higher than the MAPE achieved in sales and inventory forecasts. Nevertheless, the forecasts are improved when additional explanatory information is included. We also note that the cash flow forecasts improve by including information on accounts payables. Thus, a new finding from our analysis is that cash flow forecast accuracy also improves with lagged information from sales, inventory, and accounts payables. The intuition for this improvement is as follows. As hypothesized in our model, the per-period cash flow is influenced by lagged values of sales, inventory, and other operational variables. Thus, the VAR models yield a higher prediction accuracy than autoregressive models because the former contains additional predictive information from other variables.

The results for Home Depot in Figure 11 are similar to those for Walmart. Further, we note from Figures 10 and 11 that the relative performance of the AR model improves with a sufficient number of lags. This likely occurs because the AR model uses fewer parameters. We, however, note that the best lag VAR model outperforms the best lag AR model.

7. Conclusion

We study the problem of jointly forecasting a firm's sales, inventory, and cash flows. To this end, we propose a generalizable model of a firm's operations. We model the relationship between a firm's operational and financial variables and its evolution over time. To the best of our knowledge, ours is the first empirical work in operations management that ties together the relationships across a wide range of variables. We quantify both the *contemporaneous* as well as *dynamic* impacts of shocks in each variable on the entire system of variables. Hence, we add to the previous literature, in which different papers have considered either contemporaneous impacts only or lagged impact only or have used a subset of the variables.

We estimate our model using public financial and operational data from S&P's Compustat database. The impulse response functions generated from the model estimates provide evidence of statistically significant contemporaneous and dynamic causal impacts among the variables. In line with previous literature, we find that shocks to inventory and sales affect each other, and the impacts persist over the long run. On the other hand, shocks to cash flow have a significant impact contemporaneously but dissipate over the long run, suggesting that operational improvements in sales and inventory are more beneficial to a firm's long-run cash flows than cash infusions.

Our model has various possible applications. To demonstrate one of the applications, we estimate the impact of periods of economic recession on a firm's performance. We estimate the immediate dip in firms' sales following the recession. The sharpest decline for most firms came in the third quarter. It took about 2.5–3 years for sales and inventory to fully recover and return to equilibrium. In a second managerial application, we forecast a firm's sales, inventory, and cash flow using our model. We find consistently across firms and different model specifications that our model generates better forecasts for variables than the univariate model and those issued by equity analysts. In addition, our results show that incorporating additional information from operational and financial variables can significantly improve cash flow forecasting.

This paper provides a methodology that managers can utilize as input to their already existing decision support toolboxes. Specifically, our model can be used to estimate the causal impact of shocks, like the ongoing COVID-19 pandemic on firms' short- and long-term operational and financial performances. Such estimates can enable managers to make optimal operational policy decisions, such as inventory planning, staffing, and cash management, in light of various economic disruptions.

Acknowledgments

The authors gratefully acknowledge Department Editor David Simchi Levi, the associate editor, two anonymous referees, and Mani Sethuraman for their numerous comments, which significantly improved this paper.

Endnotes

[1] See https://tinyurl.com/cash-most-important.

[2] See https://www.synergystrategies.com/top-reasons-why-businesses-fail-by-source-jessie-hagen-u-s-bank/.

[3] Our work is focused on *cash flow from operations* or operating cash flow (Compustat field oancf) as opposed to free cash flow. For brevity, we abbreviate cash flow from operations to *cash flow* or OCF throughout this paper.

[4] As Stock and Watson (2018) state, "the macroeconomist's shock is the microeconomists' random treatment, and impulse response functions are the causal effects of those treatments on variables of interest over time, that is, dynamic causal effects."

[5] See https://www.bea.gov/data/gdp/gross-domestic-product.

[6] Thus, we impose around 96% completeness of each series. The rest of the values are interpolated as described in Section 4.

[7] It is easy to show that the estimated covariance matrix is a weighted average of the covariance matrices of the individual firms in the sample.

[8] The National Bureau of Economic Research (NBER) provides the exact months for these periods of contraction on the NBER website.

[9] See https://www.cashanalytics.com/what-is-cash-flow-forecasting/.

[10] The VAR model often provides superior estimates to those from univariate time series models and elaborate theory-based simultaneous equations models (Zivot 2006). The superior forecasting performance is because of the inclusion of additional explanatory information from other variables.

[11] The two-variables model contains sales and inventory; the three-variables model contains sales, inventory, and cash; and so on.

[12] We demonstrate our results for Walmart and Home Depot in this section. We include the results for selected other firms in Online Appendix C.

References

Agca S, Birge J, Wu J (2022a) The impact of the COVID-19 pandemic on global sourcing of medical supplies. *Medical Res. Arch.* 10(9).

Agca S, Babich V, Birge JR, Wu J (2022b) Credit shock propagation along supply chains: Evidence from the CDS market. *Management Sci.* 68(9):6506–6538.

Alan Y, Gaur V (2018) Operational investment and capital structure under asset-based lending. *Manufacturing Service Oper. Management* 20(4):637–654.

Aviv Y (2003) A time-series framework for supply-chain inventory management. *Oper. Res.* 51(2):210–227.

Barrot J-N, Sauvagnat J (2016) Input specificity and the propagation of idiosyncratic shocks in production networks. *Quart. J. Econom.* 131(3):1543–1592.

Barth ME, Beaver WH, Landsman WR (2001) The relevance of the value relevance literature for financial accounting standard setting: Another view. *J. Accounting Econom.* 31(1–3):77–104.

Bartik AW, Bertrand M, Cullen Z, Glaeser EL, Luca M, Stanton C (2020) The impact of Covid-19 on small business outcomes and expectations. *Proc. Natl. Acad. Sci. USA* 117(30):17656–17666.

Bray RL, Mendelson H (2012) Information transmission and the bullwhip effect: An empirical investigation. *Management Sci.* 58(5):860–875.

Buzacott JA, Zhang RQ (2004) Inventory management with asset-based financing. *Management Sci.* 50(9):1274–1292.

Cachon GP, Gallino S, Olivares M (2019) Does adding inventory increase sales? Evidence of a scarcity effect in U.S. automobile dealerships. *Management Sci.* 65(4):1469–1485.

Carvalho VM, Nirei M, Saito YU, Tahbaz-Salehi A (2021) Supply chain disruptions: Evidence from the Great East Japan Earthquake. *Quart. J. Econom.* 136(2):1255–1321.

Chod J (2017) Inventory, risk shifting, and trade credit. *Management Sci.* 63(10):3207–3225.

FASB (2010) Conceptual framework for financial reporting. Statement of financial accounting concepts no. 8. Accessed December 2, 2023, https://fasb.org/page/ShowPdf?path=FSP%20Staff%20draft_FASB%20Standard_Final.pdf&title=Financial%20Statement%20Presentation—Introduction%20to%20Staff%20Draft%20of%20an%20Exposure%20Draft.

Gao X (2018) Corporate cash hoarding: The role of just-in-time adoption. *Management Sci.* 64(10):4858–4876.

Gaur V, Fisher M, Raman A (2005) An econometric analysis of inventory turnover performance in retail services. *Management Sci.* 51(2):181–194.

Gupta D, Wang L (2009) A stochastic inventory model with trade credit. *Manufacturing Service Oper. Management* 11(1):4–18.

Hausman JA (1975) An instrumental variable approach to full information estimators for linear and certain nonlinear econometric models. *Econometrica* 43(4):727–738.

Hendricks KB, Jacobs BW, Singhal VR (2020) Stock market reaction to supply chain disruptions from the 2011 Great East Japan Earthquake. *Manufacturing Service Oper. Management* 22(4): 683–699.

Hribar P, Collins DW (2002) Errors in estimating accruals: Implications for empirical research. *J. Accounting Res.* 40(1):105–134.

Hu Q, Sobel M (2007) Echelon base-stock policies are financially sub-optimal. *Oper. Res. Lett.* 35(5):561–566.

Jacobson T, von Schedvin E (2015) Trade credit and the propagation of corporate failure: An empirical analysis. *Econometrica* 83(4): 1315–1371.

Jola-Sanchez AF, Serpa JC (2021) Inventory in times of war. *Management Sci.* 67(10):6457–6479.

Kesavan S, Gaur V, Raman A (2010) Do inventory and gross margin data improve sales forecasts for U.S. public retailers? *Management Sci.* 56(9):1519–1533.

Kesavan S, Kushwaha T, Gaur V (2016) Do high and low inventory turnover retailers respond differently to demand shocks? *Manufacturing Service Oper. Management* 18(2):198–215.

Kim M, Kross W (2005) The ability of earnings to predict future operating cash flows has been increasing—Not decreasing. *J. Accounting Res.* 43(5):753–780.

Li L, Shubik M, Sobel MJ (2013) Control of dividends, capital subscriptions, and physical inventories. *Management Sci.* 59(5): 1107–1124.

Lorek K, Willinger G (2009) New evidence pertaining to the prediction of operating cash flows. *Rev. Quant. Finance Accounting* 32(1):1–15.

Luo W, Shang K (2015) Joint inventory and cash management for multidivisional supply chains. *Oper. Res.* 63(5):1098–1116.

Lutkepohl H (2007) *New Introduction to Multiple Time Series Analysis* (Springer, Berlin).

Moneta A, Chlaβ N, Entner D, Hoyer P (2011) Causal search in structural vector autoregressive models. *J. Machine Learn. Res.* 12(2011):95–114.

Nallareddy S, Sethuraman M, Venkatachalam M (2020) Changes in accrual properties and operating environment: Implications for cash flow predictability. *J. Accounting Econom.* 69(2–3): 101313.

Olivares M, Cachon GP (2009) Competing retailers and inventory: An empirical investigation of General Motors' dealerships in isolated U.S. markets. *Management Sci.* 55(9):1586–1604.

Osadchiy N, Schmidt W, Wu J (2022) Customer portfolio approach to managing cash flow variability. Preprint, submitted January 4, http://dx.doi.org/10.2139/ssrn.4162190

Pearl J (2009) *Causality: Models, Reasoning and Inference*, 2nd ed. (Cambridge University Press, New York).

Rumyantsev S, Netessine S (2007) What can be learned from classical inventory models? A cross-industry exploratory investigation. *Manufacturing Service Oper. Management.* 9(4): 409–429.

Sims CA (1980) Macroeconomics and reality. *Econometrica* 48(1): 1–48.

Spirtes P, Glymour C, Scheines R (1993) *Causation, Prediction, and Search* (MIT Press, Cambridge, MA).

Stock JH, Watson MW (2018) Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *Econom. J.* 128(610):917–948.

Tong J, DeCroix G, Song J-S (2020) Modeling payment timing in multiechelon inventory systems with applications to supply chain coordination. *Manufacturing Service Oper. Management* 22(2):346–363.

Wu J, Zhang Z, Zhou SX (2022) Credit rating prediction through supply chains: A machine learning approach. *Production Oper. Management* 31(4):1613–1629.

Yang SA, Birge JR (2018) Trade credit, risk sharing, and inventory financing portfolios. *Management Sci.* 64(8):3667–3689.

Zivot E, Wang J (2003) Vector autoregressive models for multivariate time series. *Modeling Financial Time Series with S-Plus®* (Springer, New York), 385–429.