A Data-Driven Model of a Firm's Operations With Application to Cash Flow Forecasting

Kashish Arora, Vishal Gaur
Cornell University, Sage Hall, 106 East Ave, Ithaca NY 14853
ka522@cornell.edu, vg77@cornell.edu

A firm's cash flow from operations is a function of the contemporaneous and lagged values of its operational variables—sales, operating cost, inventory, payables, etc. Consequently, cash flow forecasting is a challenging problem. In this paper, we propose a generalizable and data-driven model of a firm's operations to disentangle this endogeneity and estimate causal impacts among variables. By estimating our model using quarterly public financial data from S&P's Compustat database for 1990-2020, we obtain several results. First, we show evidence that cash flow has both endogenous and lagged relationships with sales and inventory. Second, we show that lagged operational variables significantly improve the accuracy of cash flow forecasts compared to an autoregressive model of prior period cash flows alone. Moreover, cash flow also helps improve forecast accuracy for sales and inventory. Third, our model helps quantify the short- and long-run impacts of structural shocks in variables on the entire system. These estimates are useful to assess the effects of exogenous macroeconomic shocks such as the Great Recession on future cash flows and operational variables and they provide a joint distribution of variables that can be used as an input in operational planning.

Key words: Empirical Operations Management, Supply Chain Management, Economic Shocks, Structural Model, Supply Chain Finance, Forecasting, Cash Flows

1. Introduction

Forecasting the cash flow from operations is an important 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.[1] Forecasts are also used to manage liquidity risk, plan for future growth, and invest surplus cash for future growth. According to the Financial Accounting Standards Board (FASB), 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). A positive and healthy cash flow can spur the growth of the firms' sales and operations, whereas an inadequate cash flow can increase its risk of insolvency.

However, cash flow forecasting is 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). A consistent observation from these studies is that the mean absolute percentage error (MAPE) of cash flow forecasts is extremely high,

[1] `https://www.cashanalytics.com/what-is-cash-flow-forecasting/`

unlike sales or earnings forecasts. For example, Lorek and Willinger (2009) examine out-of-sample forecast accuracy using annual data for 1,174 firms from 1990-2004 and find a MAPE of 61.7% for cross-sectional estimation and 48.7% for time-series estimation. While these statistics are for annual forecasts, we obtain similar findings for forecasts of quarterly cash flows. We generate cash flow forecasts using the best-fitted ARIMA model on a rolling horizon for a sample of public U.S. firms. The resulting mean absolute percentage error (MAPE) is of the order of 50-60%, whereas the MAPE for similar models for sales forecasts is about 4-5% (see Figure 1 left panel).

Recent research in Operations Management suggests an opportunity to improve cash flow forecasting using operational variables. Specifically, a firm's cash flow from operations is computed as a function of operational variables as follows:[2]

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

We know that the terms on the right-hand side of (1) are functions of contemporaneous and lagged values of each other and of cash flow. For example, inventories affect sales, sales are used as an input to determine future amounts of inventory, and these variables together determine cash flows and earnings. Conversely, financial and accounting variables such as cash flow and accounts payable are used as input in operational decisions. However, the existing research has not explored whether a joint model of these endogenous variables can be beneficial for cash flow forecasting.

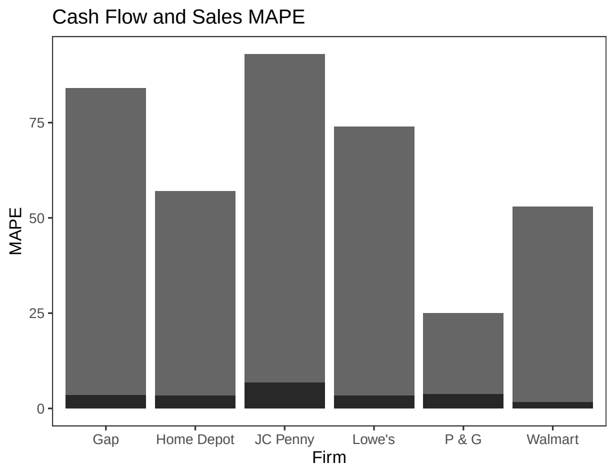
We illustrate the features of this problem using examples. First, to show the correlations of the operational variables in (1) with cash flow, Figure 1 (right panel) displays the four-quarter moving average of operating cash flow along with the corresponding values of operating profit, change in inventory, and accounts payables for Gap, Inc. for 1990-2020. We find that there is a significant temporal variation in the moving average of quarterly cash flow (coefficient of variation = 0.99). 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, for the moving averages.[3]

Second, these operational and financial variables exhibit lagged relationships. In Figure 2 (left panel), we show Gap's inventory and sales time series for 2015-2020. The inventory series are out of phase from the sales series by six months. This phase difference induces a dynamic structure between sales and inventory that has not been utilized in existing cash forecasting models. Finally,

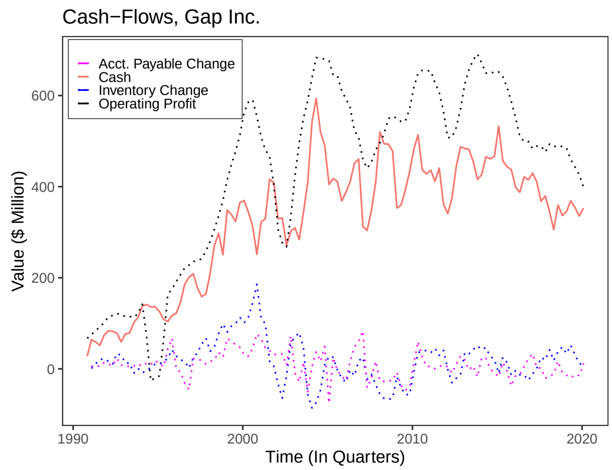
[2] Since our work is focused on *cash flow from operations*, we abbreviate this term to *cash flow* in this paper.

[3] We show moving averages for ease of presentation. The quarterly cash flows are even more volatile than the moving average.

Figure 1 Left panel shows the MAPE for Cash flow (grey) and Sales (black) forecasts for different retailers. Right panel shows four-quarter moving average of cash flows for Gap Inc. (red line). The dotted black lines in black, blue and magenta represent operating profits, change in inventory and accounts payables.

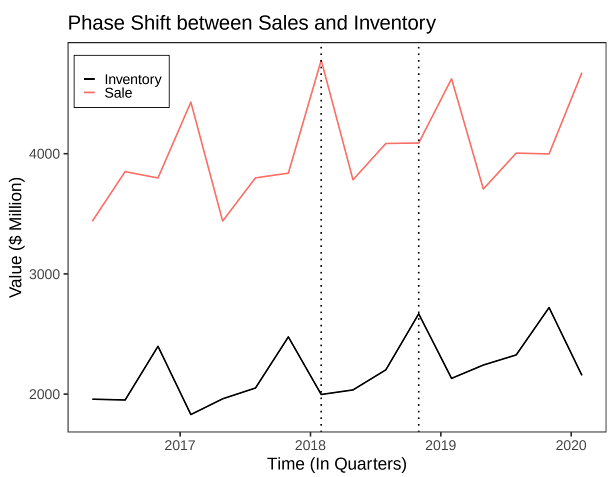


(a) Cash flow and Sales MAPE for different retailers

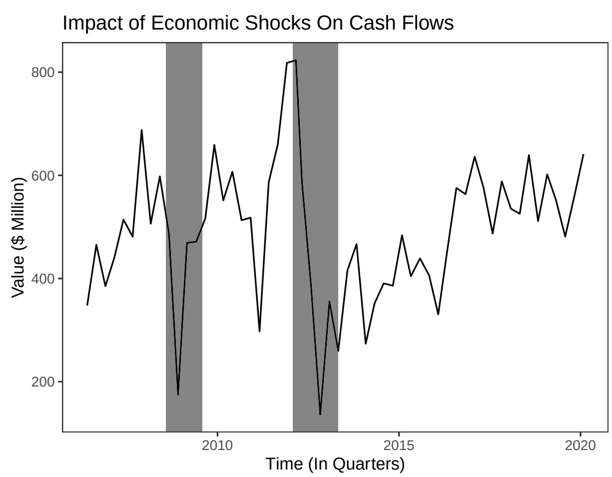


(b) Four-quarter moving average of cash flows for Gap

Figure 2 Left panel shows the phase difference between inventory (black) and sales (red) time series of Gap Inc. from 2015-2020. Right Panel shows the impact of economic shocks on the moving average of Macy's cash flows. Vertical dotted lines correspond to the periods of contraction.



(a) Phase Difference between Inventory and Sales



(b) Impacts of Periods of Recession on Cash Flows

all of these operational and financial variables are expected to be affected by external shocks such as GDP growth rate, interest rates, or other demand and supply disruptions. As an illustration, Figure 2 (right panel) shows that the 2008 and 2012 economic contractions impacted the cash flows of Macy's, Inc. significantly in the subsequent years. Thus, including the effects of such shocks on current and future cash flows becomes an important component of the cash flow forecasting problem.

In this paper, we construct a dynamic structural model of the operational variables and cash flow of a firm and causally estimate this model to answer several questions: (i) Is cash flow merely an outcome of the computation in (1), or does it have an endogenous and dynamic relationship with operational variables? (ii) What is the value of utilizing operational variables for improving the prediction accuracy of cash flow forecasts? (iii) More fundamentally, what are the short- and long-term implications of economic shocks on a firm's cash flow and operational variables?

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 cash flow, other operational and financial variables, and economic shocks. We comprehensively allow for both endogenous and cross-sectional lagged relationships among variables, e.g., cash flow $\rightarrow$ sales, accounts payables $\rightarrow$ 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—a multivariate time series of six variables spanning 30 years from 1990-2020 at quarterly frequency. In the second step, we demonstrate the managerial applications of our model by: (1) quantifying the impacts of shocks in different variables on the evolution of the entire system; (2) predicting the impacts of periods of economic downturn on the firms' future cash flow and operations; and (3) jointly forecasting cash flows and operational variables to improve forecast accuracy as compared to traditional univariate time series models.

The estimation challenge in our setting boils down to appropriate identification. A common identification strategy used in the macroeconomics literature is Cholesky factorization. This strategy imposes a causal hierarchy among the variables of interest. Such an approach does not work in our setting because the prior literature in operations management documents evidence against causal ordering. The operations management literature has studied mechanisms such as scarcity effect, service-level effect, and variety effect due to which inventory and sales both affect each other contemporaneously (Olivares and Cachon 2009, Cachon et al. 2019). Therefore, our identification strategy relies on using exogenous variables and incorporating them in the model by imposing structural restrictions on the matrix of endogenous relationships. These structural restrictions allow us to estimate the model parameters without imposing any causal hierarchy on the variables.

The results of our analysis are as follows. First, we demonstrate improved accuracy in cash flow forecasts. By generating forecasts from our model on an expanding horizon basis and comparing our forecasts' accuracy to those generated from a corresponding autoregressive model with the same number of lags, we find that our model generates better forecasts for one-quarter-ahead sales, inventory, and cash flow than the autoregressive model consistently across firms and different model specifications. For example, for Walmart, the joint model with lag one yields MAPEs of (1.23%, 6.43%, 41%) for sales, inventory and cash flow as compared to the corresponding MAPEs of (5.32%, 6.83%, 54%) as achieved by the AR(1) model. The average improvements in sales, inventory and cash flow MAPE across the retailers in our sample are 4.64%, 1.98%, and 21.11%, respectively. We then vary the number of variables and the number of time lags in the model, and report further insights into the drivers of forecasting accuracy.

Second, our model yields impulse response functions (IRFs) that show the effects of structural "shocks" or innovations in any variable on the evolution of the entire system. These estimates can be useful in assessing the short- and medium-term implications of operational planning decisions. For instance, we find that a shock in sales or inventory has a greater impact on a firm's future cash flows than a shock in cash flow itself. Additionally, the IRFs for GDP growth rate in our model are useful in quantifying the impact of economic shocks on sales, inventory, and cash flow. To demonstrate this application, we train the model on the early 2000s recession caused by the *Dot-Com Bubble* and show accurate forecasting performance for the impact of the *Great Recession of 2008-09* on individual firms multiple quarters ahead. Our analysis of these applications includes both *contemporaneous* and *dynamic* impacts; the contemporaneous impact fully characterizes the system's response to unexpected changes in the current quarter, whereas the dynamic impact projects the response in future quarters.

From a theoretical perspective, our paper yields evidence that cash flow is endogenous with sales and inventory. This implies that firms can benefit through coordination between 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 forecasting and ties together relationships across a wide range of operational and financial variables.

The rest of the paper is organized as follows. In §3, we introduce a structural model of the relationships between the firm's variables. In §4, we introduce the problem of identification in our model and describe our identification strategy and estimation procedure. In §5, we describe the summary statistics of our data. We describe our results in §6-8, showing the parameters' estimates and the IRFs in §6, the forecasting procedure and results in §7, and robustness checks in §8. Section 9 concludes the paper.

2. Literature Review

The problem of cash flow forecasting is relevant to the academic literature in both accounting and operations management. In this section, we summarize relevant papers from these streams of literature and describe the contributions of our paper.

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 to 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 standard, 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); (v) has the relative prediction accuracy of different models changed over time due to shifts in accounting and operating environments, such as trends towards lower profitability, higher growth rates, increasing using of intangibles, and shrinking operating cycles (Nallareddy et al. 2020). Since 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 operations management 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 increase in current inventory can boost current sales and earnings. Thus, it induces both endogenous and lagged effects that are included in our model.

Borrowing from the accounting literature, our paper constructs firm-wise time-series forecasting models and evaluates their performance against a benchmarking model based on lagged cash flows. Our study differs from the accounting literature because we not only improve cash flow forecasting, but also 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. Therefore, we examine contemporaneous relationships in addition to forecasting and show evidence of endogeneity. Our key focus is on operational variables instead of earnings and accruals, so that some of the variables in our study differ from those in accounting. Finally, since we seek to improve forecasting performance, we utilize data for multiple time lags and show that this indeed provides incremental benefit. We use balance-sheet data to measure cash flow in this paper, but our methodology also applies to the use of *actual* cash flows.

The theoretical research on the operations-finance interface provides a strong rationale to study cash flow in conjunction with operational variables. Several research papers consider cash as a constraint in inventory replenishment in models of debt financing (Chod 2017), asset-based lending (Buzacott and Zhang 2004, Alan et al. 2014), 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 multi-divisional 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 multi-echelon inventory system. Even when firms are not financially constrained, timing of payments can affect inventory stocking decisions (Tong et al. 2020). Thus, cash is an integral part of operational decisions, and is linked to sales and inventory. Further, Aviv (2003) make provisions for including exogenous economic variables such as GDP growth rate in demand forecasting in a theoretical inventory planning model.

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 with sales. Gaur et al. (2005) show that inventory turnover is correlated with gross margin and sales surprise; Kesavan et al. (2010) show that inventory, sales revenue, and gross margin are endogenous and that a joint consideration of these variables improves sales forecasts; Olivares and Cachon (2009) present evidence for two mechanisms, a sales effect, and a service-level effect, why sales affects inventory levels; and in the reverse direction, Cachon et al. (2019) show that inventory affects sales through a scarcity effect and a variety effect. Lagged relationships between sales and inventory have been studied by Bray and Mendelson (2012), who use a structural model to 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. While 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 JIT by U.S. manufacturers is associated with a decrease in inventory and an increase in cash hoardings.

Our paper contributes to the empirical OM literature in two ways. First, we provide evidence of contemporaneous and lagged relationships that are relevant for both operational planning and forecasting. In particular, the empirical modeling of cash flows and use of GDP in the model result in novel insights. Second, we develop a generalized model of the joint evolution of different financial and exogenous variables used in firms' operational decisions. The main benefit of our SVAR approach is to allow the inclusion of a larger set of variables, provide causal identification, and permit both contemporaneous and lagged effects. Our model yields managerially useful firm-level coefficients of short- and long-run causal impacts. These impact calculations are simple and can be directly input into a firm's decision support toolbox. We test this model with respect to different structural restrictions and alternative estimation methods such as causal directed graphs used in economics and computer science.

3. Structural Model of Variable Relationships

We model the relationships among a firm's operational and financial variables using a structural vector autoregressive (SVAR) framework. Our structural framework is built on the following axioms. First, our key variables of interest—sales, inventory, and cash flow—are endogenously determined. Second, additional exogenous operational 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. Thus, 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 large number of time periods so that its variables can be modeled using the SVAR structure at the time of data collection.

Let $\mathcal{K} = \{y_1, ..., y_n\}$ denote the set of endogenous variables and $\mathcal{Z} = \{y_{n+1}, ..., y_{n+m}\}$ denote the exogenous variables. Each $z \in \mathcal{Z}$ impacts all $k \in \mathcal{K}$, but not vice-versa. As such, the system consists of $n + m$ variables, with n^2 causal relationships among the endogenous variables and mn one-sided causal effects from the exogenous to the endogenous variables.

Each variable $i \in \mathcal{K} \cup \mathcal{Z}$ evolves over time in a stochastic manner, and the corresponding sequence $y_{i1}, y_{i2}, y_{i3}, y_{i4}, ...$ 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 $(n+m)$ dimensional 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: $\mathcal{B}(L) = \mathbf{B} - \mathbf{\Gamma_1}L - \mathbf{\Gamma_2}L^2 - \mathbf{\Gamma_3}L^3 - ... - \mathbf{\Gamma_k}L^k$. We convert the model from its lag-operator-notation to its matrix-form and write it as:

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

Here, $(\mathbf{\Gamma_1}, \mathbf{\Gamma_2}, ..., \mathbf{\Gamma_k})$ are $(n+m) \times (n+m)$ coefficient matrices that capture 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 $n^2 + mn$ contemporaneous relationships among the variables in $\mathbf{y_t}$; and $\mathbf{u_t}$ is an $n+m$ dimensional vector of structural shocks. We describe the structure of matrix $\mathbf{B}$ in §4 when we discuss identification strategy.

The structural shocks are assumed to be mean zero and independent, with a diagonal covariance matrix, $\mathbf{\Sigma} = \mathbb{E}(\mathbf{u_t}.\mathbf{u_t}')$. For example, an extreme weather event provides a structural shock to a firm's sales. Similarly, disruptions due to unforeseen supply chain disruptions can provide a structural shock to a firm's inventory. 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 due to its structural equation. Second, it has contemporaneous *indirect impact* on other endogenous variables due to the system of relationships modeled in $\mathbf{B}$. Finally, a shock in period t affects variables in future periods due to lagged relationships. For example, an extreme weather event has a direct impact on sales in the same period, an indirect impact on same period inventory and cash flow due to reduced sales, and indirect impact on all three variables in future periods because reduced sales may lead to higher inventory and cash flow depletion in the future. Thus, structural shocks govern the dynamics of the entire system. We are interested in estimating these impacts, $\mathbb{E}[\mathbf{y_{t+s}}|u_{it}=1]-\mathbb{E}[\mathbf{y_{t+s}}|u_{it}=0]$ $\forall s=\{0,1,..\}$, when variable i is subjected to a unit structural shock.

Our goal is to estimate the unknown matrices $(\mathbf{B},\mathbf{\Gamma_1},...,\mathbf{\Gamma_k})$ in (2) and then use these matrices to compute the impacts 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. Moreover, the reduced form of the model gives us a forecasting methodology.

3.1. Reduced Form

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

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

where the coefficient matrices $\mathbf{\Pi_k}\in\mathbb{R}^{(n+m)^2}$ are the reduced-form counterparts of matrices $\mathbf{\Gamma_k}$, and $\boldsymbol{\xi}_\mathbf{t}\in\mathbb{R}^{n+m}$ is a vector of 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. We emphasize here that $\mathbf{\Omega}$ is not a diagonal matrix. Hence, a variable in the reduced-form model cannot be perturbed via shocks (to forecast errors) without having delivered correlated shocks to other variables. Therefore, we cannot separately identify the effect of a shock to a variable on the system, $\mathbb{E}[\mathbf{y_{t+s}}|\xi_{it}=1]-\mathbb{E}[\mathbf{y_{t+s}}|\xi_{it}=0]$, in the reduced-form model. As such, the reduced-form estimates must first be utilized to estimate the structural form before application in any downstream operational decision-making. We describe the identification strategy and estimation procedure required for this purpose in §4.

3.2. Variables

Our model consists of three endogenous variables in set $\mathcal{K}$, Sales, Inventory, and Cash flow from operations, and three exogenous variables in set $\mathcal{Z}$, Accounts Payable, Gross Domestic Product, and Selling, General, and Administrative (SGA) Expenses. Formally, $\mathbf{y_t} = \left(\mathfrak{s}_\mathfrak{t}\mathfrak{i}_\mathfrak{t}\mathfrak{c}_\mathfrak{t}\mathfrak{ap}_\mathfrak{t}\mathfrak{gdp}_\mathfrak{t}\mathfrak{sga}_\mathfrak{t}\right)^T$, respectively.

Previous evidence in the OM literature discussed in §2 shows that sales and inventory are endogenous. Moreover, the endogeneity of cash flow can be expected from the theoretical OM literature. SGA expenses ($\mathfrak{sga}$) are commercial expenses incurred in the regular course of business operations pertaining to the securing of operating income. They include advertising, engineering and marketing expenses, freight-out and strike expense, directors' fees and remuneration, etc. We expect SGA expenses to directly impact sales and cash flow. Accounts payable consists of a firm's short-term outstanding payments to suppliers and creditors, and can be expected to directly impact inventory and cash flow. Thus, consistent with the previous OM (e.g., Kesavan et al. (2016)), we treat these variables as exogenous. Specifically, we assume that a firm optimizes sales, inventory and cash flows using SGA and AP as operational levers, and not the other way around. Finally, we use GDP as an exogenous variable as changes in GDP can be expected to impact sales.

4. Estimation

We describe the estimation of the model in this section. We start by first outlining our identification strategy. Then, we discuss the model primitives and the choice of certain hyper-parameters in our model. Finally, we describe our estimation procedure.

4.1. Identification

Identification in our setting refers to the ability to recover the structural matrices $(\mathbf{B}, \mathbf{\Gamma_1}, ..., \mathbf{\Gamma_k})$ in (2) from the reduced-from 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 $\mathbf{\Pi_k}$. Then, we recover $\mathbf{B}$ from the reduced form covariance matrix using maximum likelihood estimation. Finally, we derive matrices $\mathbf{\Gamma_k}$ from $\mathbf{B}$ and $\mathbf{\Pi_k}$ using $\mathbf{\Gamma_k} = \mathbf{B\Pi_k}$. Thus, the identification problem reduces to estimating the matrix $\mathbf{B}$.

Identification of $\mathbf{B}$ is the most important structural problem in the model because the total number of free parameters in the structural model is $(n+m)^2 \times (k+1)$, but the total number of parameters estimated in the reduced-form of the model is only $(n+m)^2 \times k + (n+m) \times (n+m+1)/2$, where the latter term in the sum corresponds to the estimated covariance matrix $\mathbf{\Omega}$. Thus, we require atleast $(n+m) \times (n+m-1))/2$ additional restrictions in the structural form to estimate the model. This is achieved by restricting matrix $\mathbf{B}$ (Sims 1980). To see why, consider the relationship between $\mathbf{B}$ and the reduced-form covariance matrix $\mathbf{\Omega}$: $\mathbf{B^{-1}B^{-1'}} = \mathbf{\Omega}$. Since $\mathbf{\Omega}$ is symmetric, it only houses

$(n+m)\times(n+m+1)/2$ 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:

$$\begin{bmatrix}\mathbf{B_{11}} & \mathbf{B_{12}}\\ \mathbf{B_{21}} & \mathbf{B_{22}}\end{bmatrix}\begin{pmatrix}\mathbf{x_t}\\ \mathbf{z_t}\end{pmatrix} = \mathbf{\Gamma_1 y_{t-1}} + \mathbf{\Gamma_2 y_{t-2}} + ... + \mathbf{\Gamma_k y_{t-k}} + \mathbf{u_t}, \tag{4}$$

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.

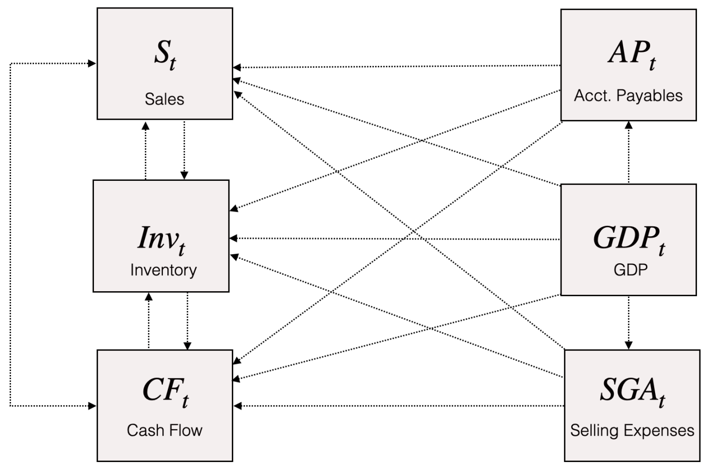
Since variables $z\in\mathcal{Z}$ are exogenous, we set $\mathbf{B_{21}}$ to $\mathbf{0}$. Further, we set $\mathbf{B_{22}}$ to $\mathbf{I}$, so that no exogenous variable $z\in\mathcal{Z}$ affects any other $z_-\in\mathcal{Z}\setminus z$ contemporaneously. This gives us mn restrictions achieved through the inclusion of exogenous variables in the model. The operational researcher can further restrict $\mathbf{B_{11}}$ or $\mathbf{B_{12}}$ to get to atleast $(n+m)\times(n+m-1))/2$ restrictions. In our procedure, normalizing the diagonal of $\mathbf{B_{11}}$ to 1 and setting $\mathbf{B_{21}}=0$ and $\mathbf{B_{22}}=\mathbf{I}$ identifies $\mathbf{B}$. Thus, the exogeneity of $\mathfrak{sga}, \mathfrak{ap}, \mathfrak{gdp}$ provides sufficient identifying restrictions for estimation. The resulting matrix $\mathbf{B}$ is depicted below and the corresponding directed graph is shown in Figure 3.

$$\mathbf{B} = \left[\begin{array}{ccc:ccc} 1 & b_{12} & b_{13} & b'_{11} & b'_{12} & b'_{13}\\ b_{21} & 1 & b_{23} & b'_{21} & b'_{22} & b'_{23}\\ b_{31} & b_{32} & 1 & b'_{31} & b'_{32} & b'_{33}\\ \hdashline 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\end{array}\right]$$

We now highlight alternative methods for identifying $\mathbf{B}$. 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)). There is a unique lower triangular $\mathbf{B}^{\triangle}$ that satisfies $\mathbf{B^{-1}B^{-1'}}=\mathbf{\Omega}$. $\mathbf{B}^{\triangle}$ can be recovered by the Cholesky decomposition of $\mathbf{\Omega}$, and hence, this identification method is also called *Cholesky* identification. Our paper departs from prior literature in this respect because setting $\mathbf{B}$ to be lower diagonal imposes a specific casual ordering on the operational variables in the system that is not supported by the OM literature. As an example, consider a system of two variables, sales ($\mathfrak{s}$), and inventory ($\mathfrak{i}$). This system has $\mathbf{B}^{\triangle}\in\mathbb{R}^{2\times2}$ and the following structural form:

$$\begin{bmatrix}b_{11} & 0\\ b_{21} & b_{22}\end{bmatrix}\begin{pmatrix}\mathfrak{s}_\mathbf{t}\\ \mathfrak{i}_\mathbf{t}\end{pmatrix} = \mathbf{\Gamma_1}\begin{pmatrix}\mathfrak{s}_\mathbf{t-1}\\ \mathfrak{i}_\mathbf{t-1}\end{pmatrix} + \mathbf{\Gamma_2}\begin{pmatrix}\mathfrak{s}_\mathbf{t-2}\\ \mathfrak{i}_\mathbf{t-2}\end{pmatrix} + ... + \mathbf{\Gamma_k}\begin{pmatrix}\mathfrak{s}_\mathbf{t-k}\\ \mathfrak{i}_\mathbf{t-k}\end{pmatrix} + \begin{pmatrix}\mathbf{u}_{\mathfrak{s}_\mathbf{t}}\\ \mathbf{u}_{\mathfrak{i}_\mathbf{t}}\end{pmatrix} \tag{5}$$

Figure 3 Directed graph obtained by structural restrictions on B. The arrows represent the direction of the contemporaneous relationships among the variables.



The first relationship in (5) models the evolution of sales over time. Sales at any given time t depends 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 prior operations management literature has shown that the contemporaneous casual impact of inventory on sales is non-zero (Kesavan et al. (2010)). The same problem still arises if the ordering of the variables for triangulation were 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-run impacts of shocks to certain variables to be 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 the above identification methods rely on apriori theory and evidence from the operations management literature. An alternate approach for identification that has been used in causal inference and computer science literature aims to infer causal relationships using the statistical properties of data directly (see, for example - Pearl (2009)). One example of this approach is constraint-based algorithms for graphical models, such as PC, SGS, and FCI algorithms (Spirtes et al. 1993). On a high level, the 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 end result is a set of graphs that are *Markov equivalent*. Though this is not the focus of the current study, such approaches may be plugged into our setup to estimate **B** in a data-driven manner. We leave the validation of such analysis for future work as an extension to the current work.

4.2. Estimation Procedure

The following steps outline our estimation procedure. We run the procedure for each firm in the sample, and obtain its IRF matrices for all variables, which trace the path of the system when subjected to shocks.

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 $\hat{\boldsymbol{\xi}}_{\boldsymbol{t}}$ as the vector of the residuals of these regressions.

Step 4: Construct $\hat{\mathbf{\Omega}} = \mathbb{E}[\hat{\boldsymbol{\xi}}_{\boldsymbol{t}}' \hat{\boldsymbol{\xi}}_{\boldsymbol{t}}]$, the estimator of covariance matrix of the residuals.

Step 5: Set $\mathbf{B_{11}}$ to 1, $\mathbf{B_{21}} = 0$ and $\mathbf{B_{22}} = \mathbf{I}$, our identifying restrictions.

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

Step 7: Estimate structural matrices $\mathbf{\Gamma_k} = \mathbf{B\Pi_k}$, and compute the Impulse Response Functions (IRF) from matrices $\mathbf{\Gamma_k}$.

In Step 1, we fourth-difference all our series to make them stationary. This is necessary because our individual quarterly time-series are not stationary due to trends. Fourth-differencing eliminates both trends and seasonal differences in our time-series. We run Augmented Dickey-Fuller (ADF) 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 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 $k \in \{1, ..., 10\}$, 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 shocks dissipate over time. Formally, stability is defined as:

Definition 1 (Stability). The reduced form VAR is covariance stationary, or *stable*, if:

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

or, the reverse characteristic polynomial has no roots in the unit circle (Lutkepohl 2007).

The above definition is equivalent to saying that the eigen values of the companion matrix $\mathbf{\Xi}$ have modulus less than 1, where the companion matrix has the following form:

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

Thus, we compute eigen values to verify stability. For example, with 5 lags, there are 35 eigen values of the companion matrix and the largest eigen value for Macy's has modulus = 0.97.

In Step 4, we construct the sample estimator of the covariance matrix, $\hat{\Omega} = \mathbb{E}[\boldsymbol{\xi_t}\boldsymbol{\xi_t'}]$, after recovering $\boldsymbol{\xi_t}$. In Step 5, we impose identifying restrictions using exogenous variables. Then, 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 $\boldsymbol{\xi_t}$ follow a multivariate normal distribution $\mathbf{N}(\mathbf{0}, \mathbf{\Omega})$. The log likelihood function is given by:

$$\mathcal{LL}(\Omega) = -\frac{(n+m)T}{2} log(2\pi) - \frac{T}{2}\log|\mathbf{\Omega}| - \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{\xi_t}'\mathbf{\Omega}^{-1}\boldsymbol{\xi_t} \tag{6}$$

$$= -\frac{(n+m)T}{2}\log(2\pi) - \frac{T}{2}\log|\mathbf{\Omega}| - \frac{1}{2}\sum_{t=1}^{T} Tr(\mathbf{\Omega}^{-1}\boldsymbol{\xi_t}\boldsymbol{\xi_t}') \tag{7}$$

$$= -\frac{(n+m)T}{2}\log(2\pi) - \frac{T}{2}\log|\mathbf{\Omega}| - \frac{T}{2} Tr(\mathbf{\Omega}^{-1}\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\xi_t}\boldsymbol{\xi_t}')) \tag{8}$$

$$= -\frac{(n+m)T}{2}\log(2\pi) - \frac{T}{2}\log|\mathbf{\Omega}| - \frac{T}{2} Tr(\mathbf{\Omega}^{-1}\hat{\mathbf{\Omega}}) \tag{9}$$

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

$$\mathcal{LL}(\mathbf{B}) = -\frac{(n+m)T}{2}\log(2\pi) + T\log|\mathbf{B}| - \frac{T}{2} tr(\mathbf{B'B}\hat{\mathbf{\Omega}}) \tag{10}$$

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

Finally, in Step 7, we construct impulse response functions (IRFs) from our estimated model. For this, we use the *Wold representation* of the system. Consider the VAR representation in (3). By repeated substitution, we arrive at the following representation of the stable reduced form VAR:

$$\mathbf{y_t} = \mu + \mathbf{\Psi_0}\boldsymbol{\xi_t} + \mathbf{\Psi_1}\boldsymbol{\xi_{t-1}} + \mathbf{\Psi_2}\boldsymbol{\xi_{t-2}} + \dots \tag{11}$$

$$= \mu + \mathbf{\Psi}(L)\boldsymbol{\xi_t}$$

where $\boldsymbol{\Psi}(L) = \sum_{m=0}^{\infty} \boldsymbol{\Psi}_m L^m$, $\boldsymbol{\Psi}_\mathbf{0} = \mathbf{I}$, and $\boldsymbol{\Psi}_\mathbf{m} = \sum_{j=1}^{m} \boldsymbol{\Psi}_\mathbf{m-j} \boldsymbol{\Pi}_\mathbf{j}$. $\boldsymbol{\Pi}_\mathbf{j}$ are the estimated lagged reduced form matrices. $\boldsymbol{\Pi}_\mathbf{j}$ is set to 0 for $j > k$, where k is the number of lags included. $\boldsymbol{\Psi}_\mathbf{m}$ are the dynamic multipliers of the VAR system. The corresponding moving average representation of the structural form can be obtained by substituting $\boldsymbol{\xi}_\mathbf{t} = \mathbf{B}^{-1}\mathbf{u}_\mathbf{t}$ in (11), by which we get:

$$\begin{aligned} \mathbf{y}_\mathbf{t} &= \mu + \boldsymbol{\Psi}(L)\mathbf{B}^{-1}\mathbf{u}_\mathbf{t} \\ &= \mu + \boldsymbol{\Theta}(L)\mathbf{u}_\mathbf{t} \end{aligned} \tag{12}$$

where $\boldsymbol{\Theta}(L) = \sum_{m=0}^{\infty} \boldsymbol{\Theta}_m L^m$, $\boldsymbol{\Theta}_\mathbf{0} = \mathbf{B}^{-1}$, and $\boldsymbol{\Theta}_\mathbf{m} = \sum_{j=1}^{m} \boldsymbol{\Theta}_\mathbf{m-j} \boldsymbol{\Gamma}_\mathbf{j}$. It is helpful to look at this moving average representation of the structural form to arrive at the impulse response of the system. Consider the following moving average representation of $\mathbf{y}_\mathbf{t+s}$ for a two variable system of sales and inventory:

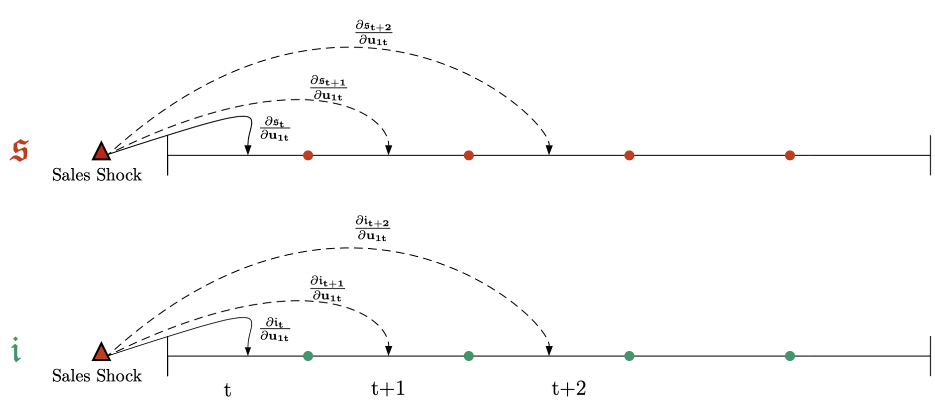
$$\begin{pmatrix} \mathfrak{s}_\mathbf{t+s} \\ \mathfrak{i}_\mathbf{t+s} \end{pmatrix} = \underbrace{\begin{bmatrix} \theta_{11}^{s} & \theta_{12}^{s} \\ \theta_{21}^{0} & \theta_{22}^{0} \end{bmatrix}}_{\text{Contemporaneous}} \begin{pmatrix} \mathbf{u}_\mathbf{1t+s} \\ \mathbf{u}_\mathbf{2t+s} \end{pmatrix} + \ldots + \underbrace{\begin{bmatrix} \theta_{11}^{s} & \theta_{12}^{s} \\ \theta_{21}^{s} & \theta_{22}^{s} \end{bmatrix}}_{\text{Dynamic}} \begin{pmatrix} \mathbf{u}_\mathbf{1t} \\ \mathbf{u}_\mathbf{2t} \end{pmatrix} + \ldots$$

where the element θ_{ij}^{s} 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 in $\boldsymbol{\Theta}_\mathbf{m}$ fully characterize how a system evolves in response to a shock in any variable over time. For example, the changes in sales and inventory, due to unit shocks in these variables, s periods in the future are:

$$\begin{aligned} \frac{\partial \mathfrak{s}_\mathbf{t+s}}{\partial \mathbf{u}_\mathbf{1t}} = \theta_{11}^{s}, \quad \frac{\partial \mathfrak{s}_\mathbf{t+s}}{\partial \mathbf{u}_\mathbf{2t}} = \theta_{12}^{s}, \ldots \\ \frac{\partial \mathfrak{i}_\mathbf{t+s}}{\partial \mathbf{u}_\mathbf{1t}} = \theta_{21}^{s}, \quad \frac{\partial \mathfrak{i}_\mathbf{t+s}}{\partial \mathbf{u}_\mathbf{2t}} = \theta_{22}^{s}, \ldots \end{aligned} \tag{13}$$

Figure 4 demonstrates the impact of a univariate shock in sales in period t, on both sales and inventory in the same period and the future. The solid arrows in the figure are the contemporaneous impacts, whereas the dashed arrows represent the impacts in the future periods.

Figure 4 Impacts of a unit \$ shock in sales, on sales and inventory. The solid arrows represent contemporaneous impacts. The dashed arrows represent impacts in the future periods.



5. Data

We use publicly available firm-level financial and operational data from S&P's Compustat database. The data span 30 years from 1990-2020, with quarterly frequency. In addition, we use historical quarterly U.S. GDP data for the same period from the Bureau of Economic Analysis (BEA) website.[4] Our sample consists of all retail, manufacturing, and wholesale firms (NAICS codes starting with 42, 44-45, and 48-89 respectively). We only keep firms that were active over the entire 30 years. 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 five missing values (out of 120).[5] Finally, we discard data with negative sales or inventory values. The resulting sample includes 50,714 observations across 419 firms. Table 1 reports the summary statistics.

Table 1 Sample and business-segment summary statistics. All financial and operational variables are measured in millions of dollars. Note that quarterly cash flow can be negative.

	Full Sample	Retail	Wholesale	Manufacturing
# Observations	50,714	4,855	3,394	42,465
# Firms	419	40	28	351
# Industries (naics)	205	21	20	164
Median Sales	249	984	268	214
Median Inventory	137	750	167	116
Median Cash Flow	12	31	4	12
Median Accounts Payable	65	259	79	54
Median SGA	45	251	36	38

6. Estimates

6.1. Evidence for endogeneity and interpretation

In this section, we provide evidence for the presence of contemporaneous causal relationships among the variables. We estimate our model separately for every firm using the approach described in §4. 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 $\mathbf{H} \in \mathbb{R}^{15\times 15}$ is the hessian matrix of the log-likelihood function at convergence. We determine whether the parameters' estimates for each firm are significantly different from zero using a $p = 0.01$ wald test. Table 2 shows the results obtained including the distributions of the parameters across all the firms in the sample and the percentage of relationships that are estimated to be statistically significant at p=0.01. We find that all our proposed linkages are statistically significant, which provides evidence for endogeneity among the variables.

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

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

Table 2 Estimates of B for the sample. All the linkages are significantly different than 0 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 percentage of firms with positive values of the parameters is depicted in column 3.

Matrix	Parameter	Linkage	% +ve	% significant***	25%ile	Median	75%ile	Mean	S.D.
B	b_{12}	I $\rightarrow$ S	68.4	100	-0.06	0.097	0.198	0.036	0.32
	b_{13}	C $\rightarrow$ S	60.1	100	-0.08	0.053	0.143	0.025	0.25
	b_{21}	S $\rightarrow$ I	51.7	100	-0.11	0.006	0.099	-0.01	0.17
	b_{23}	C $\rightarrow$ I	90.9	100	0.09	0.190	0.300	0.211	0.18
	b_{31}	S $\rightarrow$ C	50.1	100	-0.15	0.000	0.100	-0.04	0.20
	b_{32}	I $\rightarrow$ C	93.1	100	0.13	0.245	0.437	0.292	0.23

We also present the estimates of contemporaneous impacts by two-digit-SIC categories to allow an interpretation of coefficients. Note that the contemporaneous impacts are equal to $\mathbf{B^{-1}}$. We estimate the model for each two-digit-SIC separately. For each SIC, the series are a sum across individual firms constituting the SIC. Table 3 presents our estimates. For brevity, we only focus on the interactions between the endogenous variables.

The I $\rightarrow$ S estimates are positive for 20/24 SIC groups, indicating the positive contemporaneous reinforcement of inventory on sales. Our result is consistent with 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 $\rightarrow$ I estimates are positive for 21/24 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 $\rightarrow$ S and C $\rightarrow$ I estimates are positive (negative) for 15/24 and 20/24 SIC groups, respectively. The positive effect on sales may stem from firms' ability to hire more sales personnel and adjust staffing levels when cash flow is higher. Both Hall and Porteus (2000) and Bray and Mendelson (2015) allude to a firm's ability to change staffing levels quickly in line with standard labor contracts that are on a week-to-week basis. The negative effect on inventory may occur because the firms in our sample are large firms, that not cash constraint. The firm is able to invest more in inventory management and reduction policies through the use of extra cash flows.

The S$\rightarrow$ C estimates are positive for 17/24 SIC groups, and the I $\rightarrow$ C estimates are negative for 21/24 groups. These effects are 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, none of these effects are equal to one, due to the contemporaneous feedbacks between the variables.

Table 3 $\mathbf{B}^{-1'}s$ **estimates for various industries in the sample. Industries are defined by two-digit-SIC. SIC's time series are a sum of individual firms'. Columns on the right contain the estimated parameters for endogenous relationships. Horizontal lines separate retail, wholesale and manufacturing SICs.**

Industry	SIC	R/W/M	I → S	C → S	S → I	C → I	S → C	I → C
Hardware and Garden	52	R	-0.33	0.722	0.061	-0.15	0.208	-0.88
General Merchandise	53	R	0.344	-0.08	0.167	-0.02	-0.21	-0.54
Food Stores	54	R	0.622	0.013	0.105	-0.13	-0.02	-0.50
Automotives	55	R	-0.25	0.295	0.103	-0.21	0.096	-0.99
Apparel	56	R	0.133	0.795	-0.00	-0.92	0.513	-1.21
Homefurnishings	57	R	0.136	0.547	0.208	0.111	0.413	-0.38
Durables	50	W	3.641	-2.05	0.939	-2.45	-0.59	-2.36
NonDurables	51	W	0.642	0.033	0.148	0.003	-0.14	-0.30
Tobacco	21	M	-0.40	0.101	-0.15	-0.08	0.046	-0.26
Textile Mill	22	M	1.075	0.255	0.745	0.083	0.190	-0.37
Apparel	23	M	0.287	-0.38	0.142	-0.51	-0.01	-0.43
Wood	24	M	0.722	0.175	0.593	-0.09	0.398	0.140
Furniture	25	M	0.777	-0.41	0.153	-0.63	0.110	-0.71
Paper	26	M	0.921	0.297	0.210	-0.10	0.290	-0.50
Printing and publishing	27	M	0.562	-0.05	0.147	-0.31	-0.08	-1.03
Chemicals	28	M	0.140	0.282	0.276	0.054	0.315	0.171
Rubber and Plastics	30	M	0.021	0.110	0.099	-0.16	0.116	-0.90
Leather goods	31	M	0.532	-0.27	0.162	-1.57	0.188	-0.96
Stone and Glass	32	M	-0.50	0.143	-0.07	-0.12	0.478	-0.90
Primary Metals	33	M	2.781	-2.34	0.045	-2.24	0.091	-2.11
Fabricated Metals	34	M	0.816	-0.09	0.325	-0.02	-0.08	-0.33
Industrial Machinery	35	M	0.161	0.385	0.243	-0.24	0.080	-1.02
Electronics	36	M	0.948	0.903	0.030	-0.05	0.612	0.296
Transport	37	M	1.210	-0.38	0.139	-0.32	0.250	-0.51

Overall, the estimates from our model show that cash flow is endogenous with sales and inventory. Since 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.

6.2. IRF example for a single firm

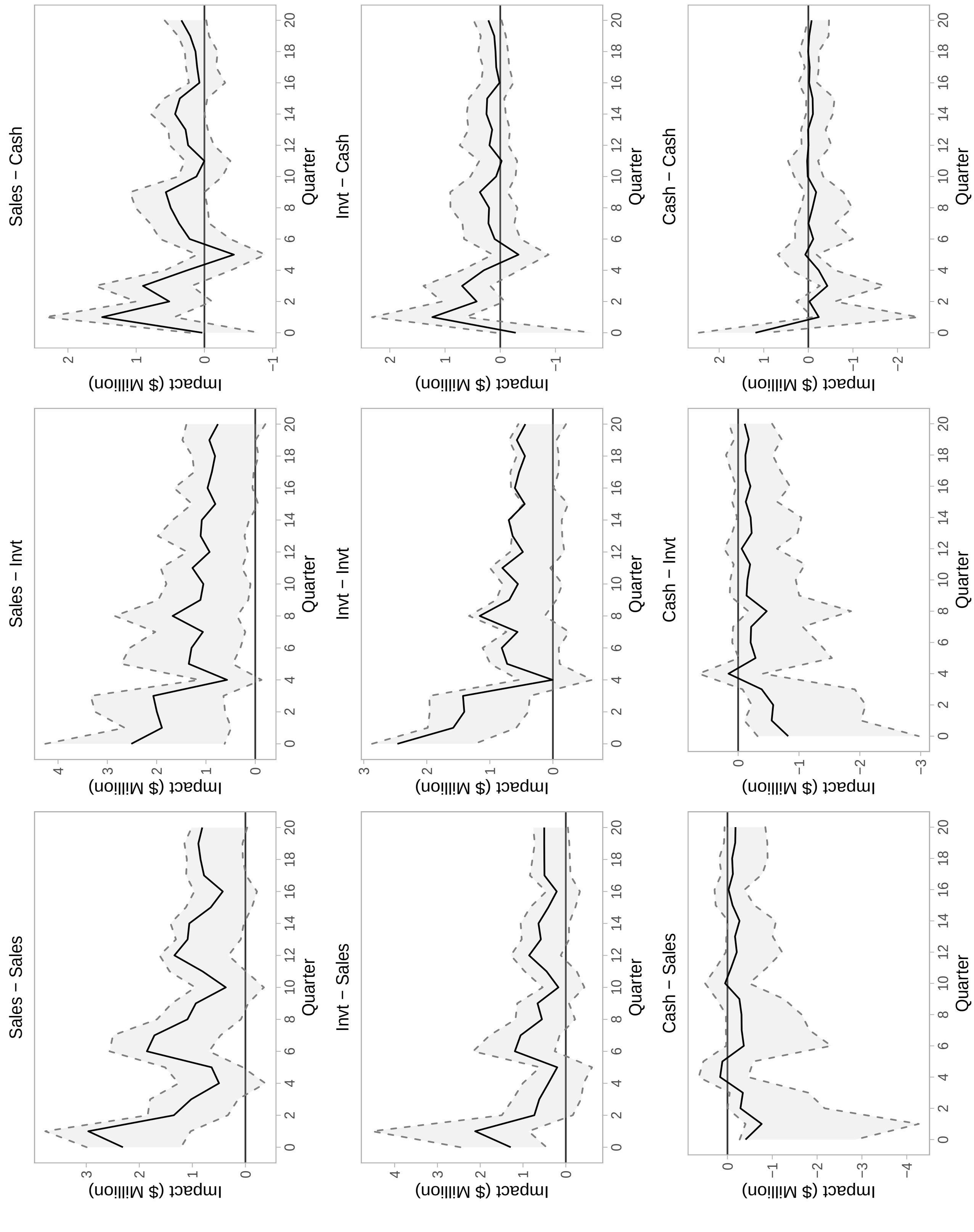
Since our model includes both contemporaneous and lagged covariates, it enables us to estimate the effect of a unit shock in one variable in period t on itself and on other variables in periods $t, t+1, \ldots$. The impulse response functions are useful to predict the effects of current uncertainties on future performance. Moreover, since the model yields a joint probability distribution of the endogenous variables, it can be utilized in decision support. In this section, we illustrate the IRFs using the data for one firm, Macy's Inc., then show summary results by industry. We estimate IRF matrices for the system of variables using the procedure outlined in §4.

We first give some examples of structural shocks in the model variables. A unit shock in sales can occur due to causes such as an unexpected weather event. Our model allows such a shock to affect other variables contemporaneously and into the future. A shock in inventory can occur when shipments are delayed or arrive sooner than average. A shock in cash flow can be due to random changes in yields on its short term investments or in banks' lending policies. Examples of shocks in the other variables are: changes in credit terms by suppliers (accounts payable), labor shortages, absenteeism, or changes in marketing plans (SGA expenses), and macroeconomic events such as recessions (GDP growth rate). Our goal is to utilize the estimated model to compute the effects of such shocks on the system. A manager can use these impacts to make operational and financial decisions, such as procuring more inventory, making cash infusions, and renegotiating terms of accounts payable to offset the effect of a shock.

In Figure 5, we plot the salient impulse responses for Macy's. The title of each plot displays the impulse variable followed by the response variable. For example, Sales-Cash on the top-right plot denotes the impact of a unit shock in sales on cash in the current and the future periods. X-axes represent the time in quarters. We extrapolate the response functions to 5 years in the future. We observe that even a small shock to a variable has a statistically significant long-term impact on the entire system of variables. Most 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 correlations among the variables, and not due to mere seasonality. We note the following observations from Figure 5.

Shock to Sales. A unit shock in sales results in an immediate increase in sales of around 2 units. The shock has a statistically significant effect that lasts till 20 quarters. The long-run impact measured at 20 quarters is around 0.8 units. The shock also has a positive impact on inventory, both in the

Figure 5 **IRF estimates for self- and cross-variable impacts for Macy's Inc. X-axes represent time in quarters and Y-axes represent the point impacts. Error bands are 95% confidence intervals.**



short- and long-run. The impact is statistically significant till 14 quarters. Finally, the impacts on cash in the short- and long-run are mostly less than one. This suggests that to increase future cash flows, a large sales push is required in any period.

Shock to Inventory. A unit shock to inventory results in an increase in sales contemporaneously. This demonstrates the positive reinforcement of inventory on sales. The impact persists for about 2 periods in the future. The impact on cash flow is negative contemporaneously. But there is a positive and significant impact in the next period. Hence, inventory affects cash flows with a lag of order one. Overall we find that positive shocks to inventory increase cash flows in the long run.

Shock to Cash. Finally, a shock to cash flow dissipates very quickly. There is a significant impact contemporaneously, suggesting that large cash infusions are helpful in the short-run. However, in the long-run, cash infusions do not have a significant impact. This runs counter to conventional wisdom, and suggests that operational improvement in sales and inventory are more beneficial to a firm's long-run cash flows as compared to significant cash infusions themselves.

In addition to firm-wise impulse response functions, we also estimate impulse responses for industry segments. We cluster firms into different industry segments using the Standard Industry Classification (SIC) codes. Then, we construct the IRF estimates by estimating our model on the summed up series. Table 6.2 shows the IRF estimates for the retail segment - variety stores. The segment consists of eight firms and has the SIC code = 5331. $\mathcal{S} = \{$Walmart, Macy's, Penny JC, Target, Dollar General, Big Lots, Tuesday Morning, Dillards$\}$. The table includes four matrices of impulse responses - one contemporaneous, and three corresponding to the future periods.

6.3. Case Study: Estimating the Impact of Periods of Economic Downturn

Global economic recessions severely impact a firm's operational and financial performance. Historically, macroeconomic crises and shocks have led to disruption of supply-chain operations, plummeting of sales, depletion of cash flows and reserves, and in the worst case, caused businesses to go bankrupt. A case in point is the COVID 19 pandemic, where many firms witnessed some of the consequences mentioned above with 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. Of those who survived, a large proportion of firms had to close down temporarily, lay off employees, or sell their assets.

In this section, we use our model to quantify the impact of two historical macroeconomic shocks on firms' operational and financial variables. The first macroeconomic shock we include in our analysis is the economic downturn following the *Dot-com Bust of 2001*. The shock lasted eight months from March-Nov 2001. The second shock is due to the *Great recession of 2007*, that lasted for 18

Table 4 IRF estimates for the first 4 periods for the segment - Variety Stores. The first matrix is the contemporaneous impact matrix. The next 3 correspond to the future periods. * indicates statistical significant at 5% level.

	$\mathfrak{s}_t$	$\mathfrak{i}_t$	$\mathfrak{c}_t$	$\mathfrak{ap}_t$	$\mathfrak{gdp}_t$	$\mathfrak{sga}_t$
$\mathfrak{s}_t$	1.064*	0.344*	−0.089*	−0.057	−0.448*	0.061
$\mathfrak{i}_t$	0.167*	1.068*	−0.044*	−0.028	0.796	0.030*
$\mathfrak{c}_t$	−0.219*	−0.545*	1.032*	0.667*	−0.745	−0.708*
$\mathfrak{ap}_t$	0	0	0	1	0	0
$\mathfrak{gdp}_t$	0	0	0	0	1	0
$\mathfrak{sga}_t$	0	0	0	0	0	1
Period 1						
$\mathfrak{s}_t$	0.765*	0.617*	0.040*	0.367*	−0.878	0.439
$\mathfrak{i}_t$	0.323*	0.772*	0.252*	−0.081	1.621	−0.882*
$\mathfrak{c}_t$	0.371*	1.407*	−0.002	−0.732*	−1.371	0.299
$\mathfrak{ap}_t$	0	0	0	0.558*	2.360*	−0.056
$\mathfrak{gdp}_t$	0	0	0	0	1.276*	0
$\mathfrak{sga}_t$	0	0	0	0.096*	−0.310*	0.727*
Period 2						
$\mathfrak{s}_t$	0.796*	0.615*	−0.008*	0.365	−3.280*	0.767
$\mathfrak{i}_t$	0.231*	0.661*	0.022*	−0.038	4.152*	−0.440
$\mathfrak{c}_t$	0.374*	−0.159*	0.497	−0.156	−1.890	−0.551
$\mathfrak{ap}_t$	0	0	0	0.502*	4.099*	0.417*
$\mathfrak{gdp}_t$	0	0	0	0	1.434*	0
$\mathfrak{sga}_t$	0	0	0	0.079*	−0.455	0.718*
Period 3						
$\mathfrak{s}_t$	0.506*	0.533*	0.094*	0.500*	−0.859	0.374
$\mathfrak{i}_t$	0.133	0.519*	−0.006	0.052	6.071*	−0.024
$\mathfrak{c}_t$	−0.073	−0.165	−0.040*	0.477	−2.884	0.692
$\mathfrak{ap}_t$	0	0	0	0.529*	3.795*	0.444
$\mathfrak{gdp}_t$	0	0	0	0	1.602*	0
$\mathfrak{sga}_t$	0	0	0	0.132*	0.230	0.531*
Period 4						
$\mathfrak{s}_t$	0.313	0.806*	0.309	0.508	1.566	0.721
$\mathfrak{i}_t$	−0.014	0.109*	0.085	0.155	5.843*	−0.079
$\mathfrak{c}_t$	−0.296*	−0.312*	−0.432*	0.226	−0.408	1.846
$\mathfrak{ap}_t$	0	0	0	0.236	2.739*	0.482
$\mathfrak{gdp}_t$	0	0	0	0	1.296*	0
$\mathfrak{sga}_t$	0	0	0	0.187*	0.457	0.090

months from December 2007-June 2009.[6] We quantify these impacts by including a dummy variable $\mathfrak{recession}_t$ in our model. $\mathfrak{recession}_t = 1$ for the two periods of recession, and 0 otherwise. With the inclusion of the dummy variable, the dynamic system of variables consists of seven variables, and the corresponding matrix $\mathbf{B} \in \mathbb{R}^{7\times7}$. We re-estimate the structural model for this system of seven variables as in §4. Then, we compute the corresponding impulse-response functions to estimate the impacts. We are interested in the impulse response function generated by the dummy variable. Since we set $\mathfrak{recession}_t = 1$ for both the shocks, our estimated impacts have the same magnitude for each of the recession periods.

[6] We obtain the exact months corresponding to these periods of contraction from the NBER website

Figure 6 **Impacts of Recession on Sales and Inventory - Macy's**

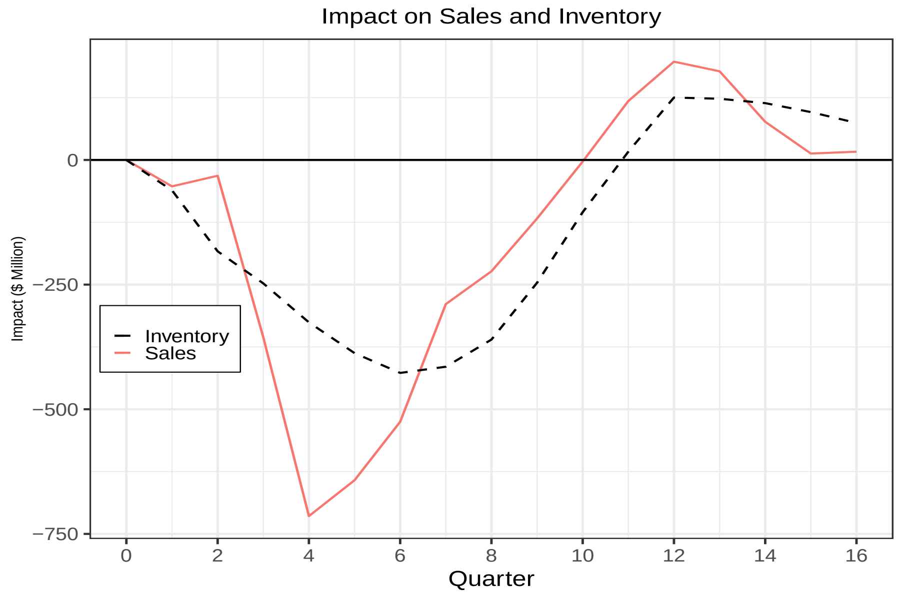
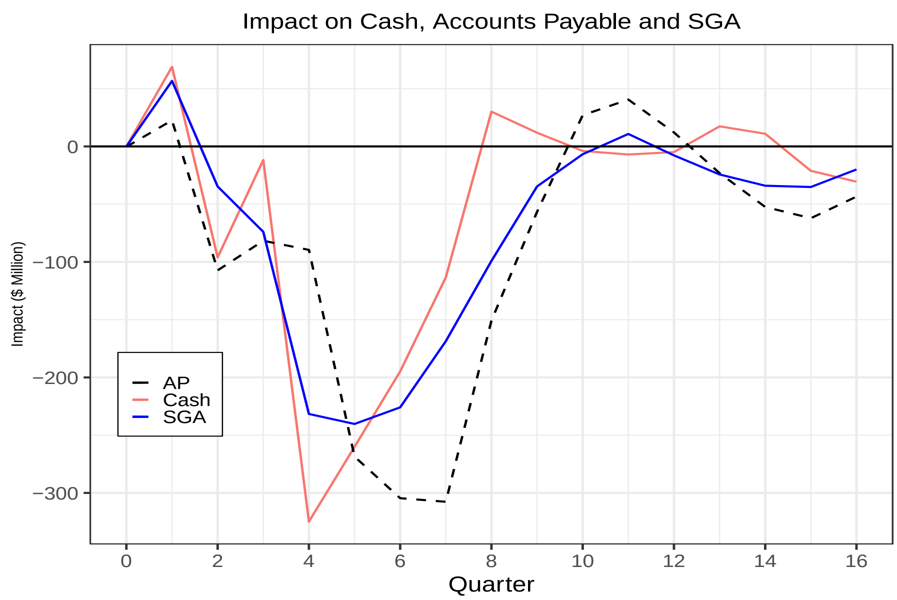


Figure 7 **Impacts of Recession on Cash Flow, Accounts Payable and SGA - Macy's**



Figures 6 and 7 depict the impacts of the shock due to recession on Macy's operational and financial variables. We make the following observations about the impacts:

- The recession significantly impacted all variables.
- Macy's sales and inventories plummeted immediately following the recession. The decline in sales was sharp and had the highest magnitude in the third and the fourth quarter.

- It took a year for Macy's sales to recover from its lowest slump and a year and a half for inventories to improve. Both show gradual recovery that only got completed after two and a half years. This highlights that the impacts of recessions were severe and long-lasting.
- Cash flows improved immediately. This may be due to significant markdowns by the firm to get rid of excess inventory at hand. However, soon after, sales declined faster than inventory, and as a result, cash flows also dropped, only recovering after a year.
- Both sga and accounts payables were impacted like cash flow. It took around 9 to 10 quarters for these variables to recover again.

6.3.1. Out of sample prediction with the seven variable system To provide support to our results in this section, we run an out of sample fit exercise. We run two models, one with a six variable system, and other with a seven variable system that includes the recession dummy. The recession dummy for Macy's is a 120×1 vector with ones in the elements corresponding to the periods of economic downturn, and zeros elsewhere. We run the two models for a *training period* comprised of data till the start of second recession. Using the estimates, we forecast sales, inventory and cash for a *test period* comprised of the next four quarters. Specifically, $\mathfrak{recession}_t = [\underbrace{0, 0, , \ldots, 1, 1, 1, 1, 0, 0, 0, \ldots}_{\text{Training}}, \underbrace{1, 1, 1, 1}_{\text{Test}}, \ldots]$. We compare the MAPE for the three variables across the two models. We find that $\text{MAPE}_6 = \{2.15\%, 5.82\%, 16.4\%\}$, whereas $\text{MAPE}_7 = \{1.69\%, 6.76\%, 14.9\%\}$.

7. Joint Forecasting of Variables

Forecasting operational and financial time series data is of high importance to firms. Sales forecast for a firm is a necessary input for its valuation and is used to project earnings and growth rate Kesavan et al. (2010). Thus, firms and equity analysts exert a significant amount of time and effort in forecasting sales. Similarly, accurately forecasting inventory is crucial for managing a firm's operations. It reduces the risk of stock-outs and helps to reduce product waste and inventory holding costs. Finally, accurately forecasting cash flow helps avoid having shortfalls in the cash balances, reduces the risk of insolvency, and ensures that the firm honors its payment commitments to the suppliers.

7.1. Forecasting Procedure

We generate joint forecasts for the variables in our model. Specifically, we use the reduced form VAR to generate one-step-ahead forecasts for the multivariate time series.[7] We calculate one-step-ahead

[7] 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 due to the inclusion of additional explanatory information from other variables.

forecast values of $\mathbf{y_t}$ based on the available information till time t. Let the one-step ahead forecast be $\mathbf{y^*_{t+1}} = \mathbf{y_{t+1|t}}$. The best linear predictor of $\mathbf{y^*_{t+1}}$ is given by

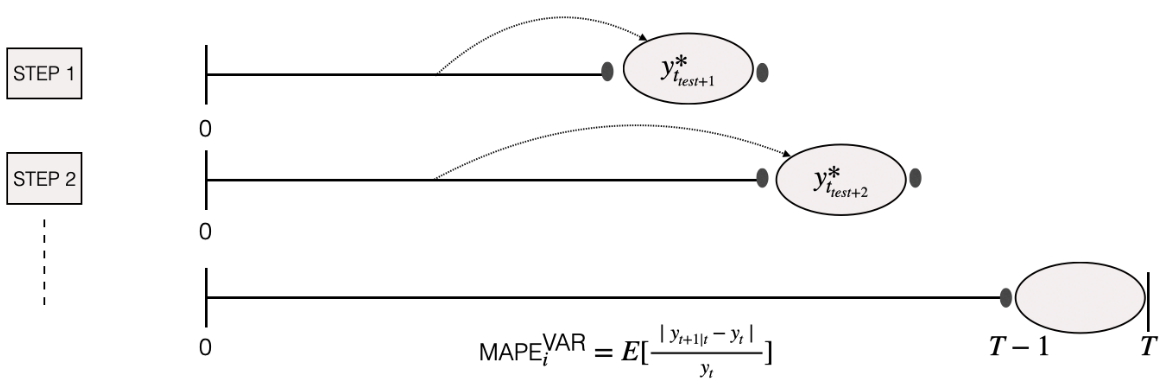
$$\mathbf{y^*_{t+1}} = \mathbf{\Pi_1 y_t} + \mathbf{\Pi_2 y_t} + ... + \mathbf{\Pi_k y_{t-k+1}} \tag{14}$$

where, $\mathbf{\Pi} = \{\mathbf{\Pi_1}, \ldots, \mathbf{\Pi_k}\}$ are the estimated reduced-form matrices. $\mathbf{y^*_{t+1}}$ 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 (MAPE) of the *predicted* values with respect to the *actual values*. Specifically, for variable $i \in \mathcal{K} \cup \mathcal{Z}$, the MAPE measure is defined as:

$$\text{MAPE}_i^{\text{VAR}} = \mathbb{E}[\frac{|\, y_{t+1|t} - y_t \,|}{y_t}] \tag{15}$$

We generate forecasts on an expanding-window basis. First, we estimate reduced-form matrices using the first t_{test} quarters of data. Then, we generate the one-step-ahead forecasts $\mathbf{y_{t_{test}+1|t_{test}}}$. Subsequently, we use the data from from the first $t_{test}+1$ periods, and generate the forecasts, and so on. For our main results, we set $t_{test} = 80$ (1990-2010). We check the robustness of our results to this specification and report the corresponding results for $t_{test} = 40$ in the Appendix.[8] Figure 8 shows the schema for the expanding-window forecasting approach.

Figure 8 Expanding-Window method of forecasting. t_{test} **is the size of the initial test set**



We compare the MAPE of the predicted values of the endogenous variables — sales, inventory, and cash-flow with the corresponding MAPE predicted by an autoregressive model. The autoregressive model is applied in a expanding-window manner as well. To make the model predictions comparable, we generate autoregressive forecasts using the same lags as the VAR model. We evaluate the MAPE over the last 40 quarters.

We run this forecasting procedure for different multivariate systems composed of many variables and lag values. Specifically, we start with the binary system of lag one, comprised of the variables

[8] For $k = 5$, we find that setting $t_{test} = 40$ provides just enough degrees of freedom to estimate the reduced form.

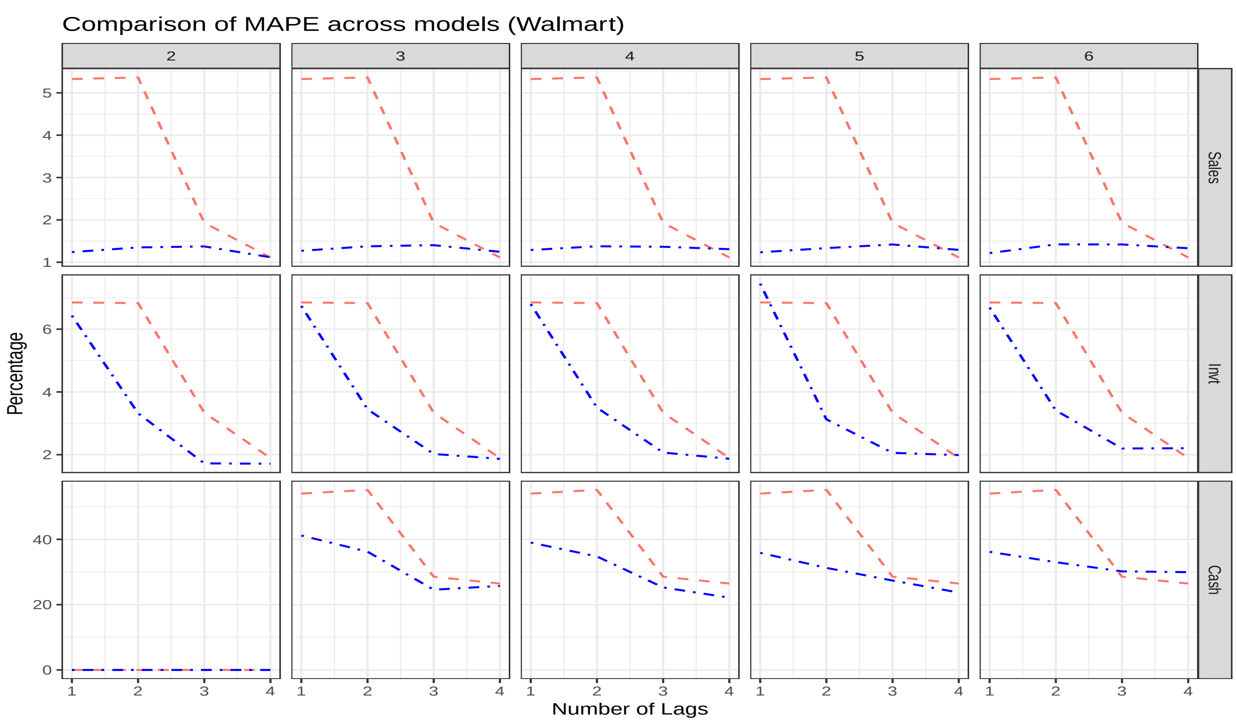
sales and inventory, and compute the MAPE from the two models. We then repeat the procedure for higher lags of the binary system. Then, we keep adding a variable in the design and compute the corresponding MAPE from both models. The purpose of this procedure is to test the robustness of our results and generate comparative statics across a wide range of model specifications.

7.2. Results and Discussion

We collect the results of our forecasting exercise for Walmart and Home Depot in Figures 9 and 10. The columns in Figures 9 and 10 denote the number of variables in the system, starting from the two-variable system of sales and inventory and then adding cash, accounts payable, GDP, and SGA.[9] The y- and x-axes in the plots denote MAPE percentage and the number of lags. In total, we evaluate 5×4 model specifications.

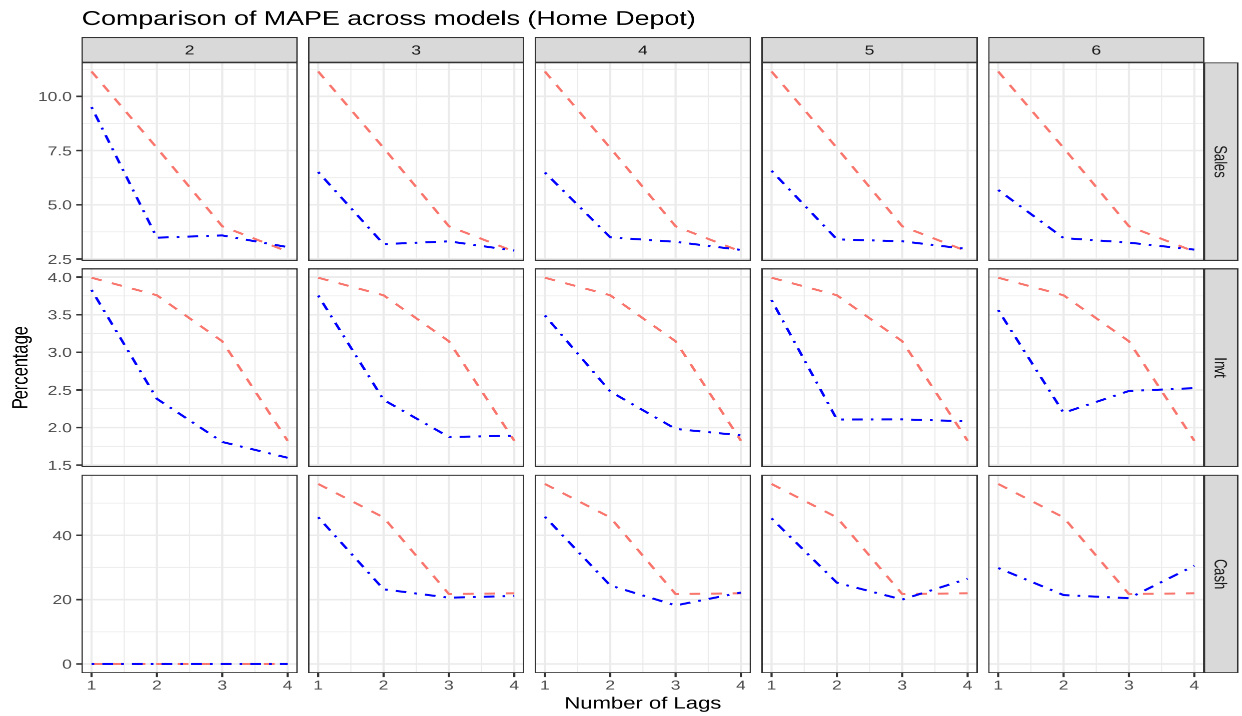
First, for sales forecasts, we find that the VAR model outperforms the autoregressive model across all specifications for lags 1,2, and 3. For the lag 4 model, both models perform nearly similarly. The top left square shows the immediate improvement in sales forecast in the two-variable model of sales and inventory. The improvement is in line with previous literature (Kesavan et al. (2010)), which documents that adding inventory information improves sales forecasts.

Figure 9 MAPE values for endogenous variables across different model specifications for Walmart. Blue(red) denotes the MAPE for the VAR(autoregressive) model.



[9] We demonstrate our results for Walmart and Home Depot in this section. We include the results for some other retailers in the Appendix.

Figure 10 **MAPE values for endogenous variables across different model specifications for Home Depot. Blue(red) denotes the MAPE for the VAR(autoregressive) model.**

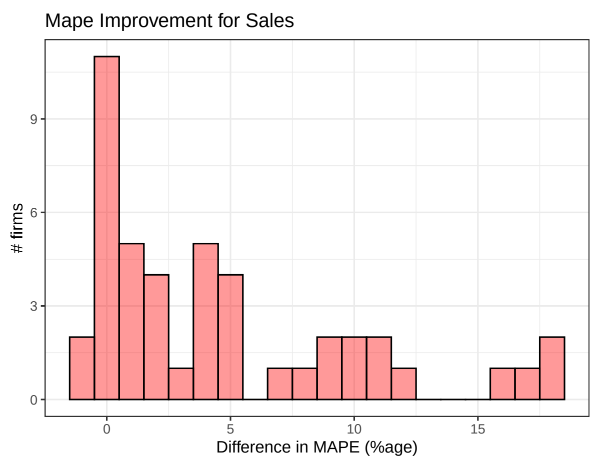


Second, Inventory forecasts are also better with joint forecasting across a wide range of models. The most significant improvement in inventory forecasts occurs with the inclusion of lagged values of sales. In the basic two-variable setup, adding lagged sales values increases the accuracy of the forecasts drastically. The finding is also in line with previous literature, demonstrating how sales information with different lead time signals impacts inventory.

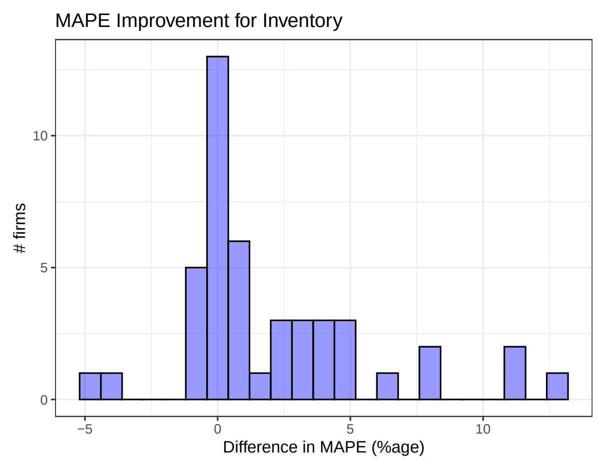
Third, we also document a significant improvement in cash flow forecasts across the models. Quarterly cash flows are highly variable, and the corresponding forecasting problem is hard. The lowest MAPE achieved across different models for cash-flow forecasting is of the order of 40%, which is significantly higher than the MAPE achieved in sales and inventory forecasts. Nevertheless, the forecasts are improved when additional explanatory information is included. A new finding from our analysis is that cash flow forecast accuracy also increases with lagged information from sales and inventory. This is the first empirical evidence for improved accuracy in cash flows by utilizing operational information. The intuition for this improvement is as follows. The per-period cash flow is determined from the accounting identity in Equation 1. The value of cash flow directly depends on all the other variables in the Equation. VAR models' prediction accuracy is higher than autoregressive models since the former contains additional predictive information from other variables and their lags.

We find a consistent improvement in MAPE for the endogenous variables across firms. In Figure 11, we plot the frequency distributions of the improvement in MAPE for sales, inventory, and cash flow for various retailers in our sample. The improvement is calculated as AR MAPE - VAR MAPE. The plot shows that VAR leads to improved forecast accuracy for the endogenous variables for most retailers in our sample. Specifically, the average improvement in sales, inventory and cash flow MAPE across the retailers in our sample is around 4.64%, 1.98%, and 21.11% respectively. Finally, the predictive accuracy for sales is also higher than analyst forecasts, which have a MAPE of around 4 % (see, Kesavan et al. (2010)). All the above observations help make a strong case for using the VAR model for forecasting.

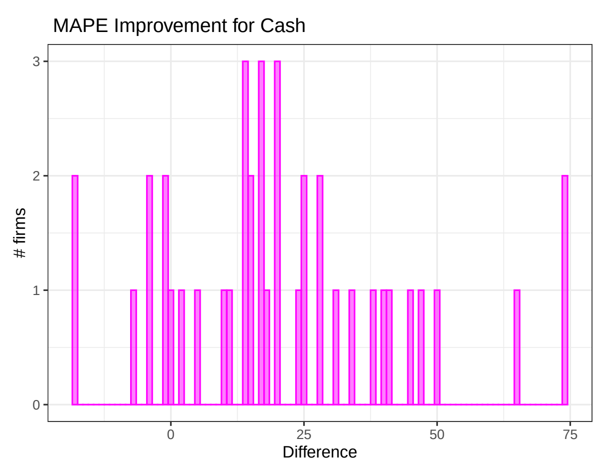
Figure 11 These figures show the frequency distribution of improvement in MAPE for (a) Sales, (b) Inventory, and (c) Cash flow forecasts for the retailers in our sample. X-axis represents the difference in MAPE calculated as AR MAPE - VAR MAPE. The VAR MAPE is calculated for a three variable system of lag order two.



(a) Improvement in Sales MAPE



(b) Improvement in inventory MAPE



(c) Improvement in Cash MAPE

8. Robustness Analysis

Throughout this paper, we assume that a firm's operational and financial variables are determined simultaneously and are econometrically endogenous. The assumption also implies that the variables affect each other, and the interactions are significant. We rigorously test this central assumption of simultaneity in this section in a couple of ways. First, we decompose the variance of the forecast errors in the variables as a function of other system variables. We show that the variables' interactions are strong, thereby rejecting the idea that the variables are exogenously determined. Second, we perform Granger and Instantaneous Causality tests to test that our system's variables provide statistically significant information about the other variables contemporaneously and in the future.

8.1. Decomposing the Variance in Forecast Errors

In this section, we decompose the variance of forecast error in the system's variables into various components. The one-step-ahead forecast error of a variable $i \in \mathcal{K} \cup \mathcal{Z}$ at time t is the difference of the one-step-ahead forecast at $t-1$ and the realized value of the variable at t, i.e., the error is the change in the variable that one is unable to forecast at $t-1$. The error can be decomposed as a linear function of different structural shocks to the system, as described in the text.

Forecast error variance decomposition quantifies the proportion of the variance of forecast error in a variable due to different structural shocks in all the variables. It is an indicator of the amount of information that a variable contributes to others in the future in our model's framework. Specifically, if a large proportion of the variance of a focal variable's forecast error is due to the variable itself's structural shocks, then the interaction between other variables and the focal variable would be weak. The interaction is strong if the structural shocks in other variables contribute to the focal variable's forecast error variance.

Expression for Forecast Error Variance Decomposition Using wold's representation (as in §4), the h-step ahead forecast error vector at any time t can be written as a function of structural errors:

$$\mathbf{y_{t+h}} - \mathbf{y_{t+h|t}} = \sum_{\mathbf{s=0}}^{\mathbf{h-1}} \mathbf{\Theta_s u_{t+h-s}} \tag{16}$$

For a particular series in $\mathbf{y_t}$, the forecast error is given by:

$$y_{i,t+h} - y_{i,t+h|t} = \sum_{s=0}^{h-1} \theta_{i1}^s u_{1,t+h-s} + \ldots + \sum_{s=0}^{h-1} \theta_{in+m}^s u_{n+m,t+h-s} \tag{17}$$

Since the structural errors are orthogonal, the variance of the above can be written as:

$$\text{Var}(y_{i,t+h} - y_{i,t+h|t}) = \sigma_{u_1}^2 \sum_{s=0}^{h-1} (\theta_{i1}^s)^2 + \ldots + \sigma_{u_{n+m}}^2 \sum_{s=0}^{h-1} (\theta_{in+m}^s)^2 \tag{18}$$

where $\sigma_{nj}^2 = \mathrm{Var}(u_{njt})$ and Var is the variance operator. The contribution of the shock u_j to $\mathrm{Var}(y_{i,t+h} - y_{i,t+h|t})$ is then:

$$\mathrm{FEVD}_{i,j}(h) = \frac{\sigma_{u_j}^2 \sum_{s=0}^{h-1} (\theta_{ij}^s)^2}{\sigma_{u_1}^2 \sum_{s=0}^{h-1} (\theta_{i1}^s)^2 + \ldots + \sigma_{u_{n+m}}^2 \sum_{s=0}^{h-1} (\theta_{in+m}^s)^2}, i,j = 1,\ldots n+m \tag{19}$$

For our system, there are $(n+m) \times n$ values for $\mathrm{FEVD}_{i,j}(h)$ for the endogenous variables. Table 5 shows the corresponding FEVD values for variables for Macy's for $h \in \{1,\ldots,5\}$.

Table 5 These figures show the frequency distribution of improvement in MAPE for (a) Sales, (b) Inventory, and (c) Cash flow forecasts for the retailers in our sample. X-axis represents the difference in MAPE calculated as AR MAPE - VAR MAPE. The VAR MAPE is calculated for a three variable system of lag order two.

Forecast error in:	h	Shock to: $\mathfrak{s}_t$	$\mathfrak{i}_t$	$\mathfrak{c}_t$	$\mathfrak{ap}_t$	$\mathfrak{gdp}_t$	$\mathfrak{sga}_t$
$\mathfrak{s}_t$	1	0.399	0.118	0.053	0.054	0.355	0.020
	2	0.317	0.171	0.057	0.067	0.344	0.044
	3	0.231	0.227	0.074	0.060	0.291	0.116
	4	0.225	0.204	0.072	0.058	0.312	0.129
	5	0.236	0.174	0.058	0.045	0.243	0.244
$\mathfrak{i}_t$	1	0.047	0.406	0.077	0.087	0.340	0.043
	2	0.043	0.370	0.070	0.078	0.401	0.037
	3	0.047	0.339	0.064	0.073	0.438	0.040
	4	0.098	0.279	0.053	0.056	0.350	0.165
	5	0.107	0.215	0.041	0.045	0.442	0.150
$\mathfrak{c}_t$	1	0.070	0.088	0.474	0.015	0.260	0.092
	2	0.088	0.157	0.302	0.063	0.215	0.175
	3	0.088	0.141	0.244	0.083	0.284	0.160
	4	0.067	0.111	0.193	0.063	0.306	0.260
	5	0.053	0.084	0.151	0.046	0.248	0.417
$\mathfrak{ap}_t$	1	0	0	0	1	0	0
	2	0	0	0	0.970	0.010	0.020
	3	0	0	0	0.410	0.005	0.584
	4	0	0	0	0.348	0.077	0.576
	5	0	0	0	0.373	0.088	0.539
$\mathfrak{gdp}_t$	1	0	0	0	0	1	0
	2	0	0	0	0	1	0
	3	0	0	0	0	1	0
	4	0	0	0	0	1	0
	5	0	0	0	0	1	0
$\mathfrak{sga}_t$	1	0	0	0	0	0	1
	2	0	0	0	0.021	0.022	0.957
	3	0	0	0	0.024	0.022	0.954
	4	0	0	0	0.025	0.036	0.938
	5	0	0	0	0.030	0.056	0.914

If a variable in Table 5 were strictly exogenous, the column elements corresponding to the self shock when explaining its forecast error would be 1 for all periods, and all other columns would have zeroes. In other words, the structural shocks in a variable itself account for all the variance in its forecast error. The above is seen for the $\mathfrak{gdp}$ variable, which is truly exogenous.

Except $\mathfrak{gdp}$, none of the variables have all the forecast error variance accounted for by shocks in itself, in all time horizons. The endogenous variables have at least 40 % of their variance accounted for by own shocks contemporaneously. This percentage gets reduced quickly by period 5. $\mathfrak{inv}$t shocks explain a sizeable amount of variance in forecast errors of $\mathfrak{sales}$ and $\mathfrak{cash}$. The $\mathfrak{sga}$ variable has more than 90% variance accounted for by innovations in itself. The above observations reinforce the choice of the dynamic model proposed in the paper.

8.2. Granger and Instantaneous Causality Tests

In this section, we provide additional evidence, via hypotheses testing, that the variables affect each other. Specifically, we test for Granger and Instantaneous Causality. The variable $\mathfrak{invt}$ is said to granger cause $\mathfrak{s}$, if the past values of $\mathfrak{invt}$ provide statistically significant information about $\mathfrak{s}$ in the future. Similarly, variable one is said to instantaneously granger cause the others if the variable's current values provide statistically significant information about the others contemporaneously. Granger and instant causality can be tested for by Wald tests. The granger (and instantaneous) wald test tests for the null hypothesis that variable one does not granger cause the others. Table 6 reports the p-values for the granger and instantaneous causality Wald tests for the different variables.

Table 6 Wald Tests

Series	Granger Causality	Instantaneous Causality	Null-Rejected?
$\mathfrak{s}_t$	9.0×10^{-4}	1.8×10^{-7}	✓
$\mathfrak{i}_t$	1.4×10^{-3}	2.1×10^{-8}	✓
$\mathfrak{c}_t$	7.0×10^{-2}	7.9×10^{-7}	✓
$\mathfrak{ap}_t$	3.8×10^{-2}	2.5×10^{-8}	✓
$\mathfrak{gdp}_t$	8.0×10^{-2}	1.5×10^{-6}	✓
$\mathfrak{sga}_t$	3.3×10^{-3}	1.7×10^{-6}	✓

As is evident from the p-values in Table 6, the null hypotheses corresponding to the granger and instantaneous causality tests are rejected for all the variables except $\mathfrak{gm}_t$. Hence, all variables in our system provide statistically significant information about the others, both contemporaneously and in the future.

9. Conclusion

We study the problem of jointly forecasting a firm's sales, inventory, and cash flows. To this end, we propose a generalizable and data-driven 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 relationships across a wide range of variables. Though we include six variables in our analysis, one can easily add other operational variables to our model. We quantify both the *contemporaneous*, as well as *dynamic* impacts of shocks in a variable on the entire system of variables. Hence, we add to the previous literature that has only considered contemporaneous impacts. Our model is entirely data-driven. We do not impose any prior theory-based restrictions on the data generating process for 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 improvement in sales and inventory is 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 comes in the third quarter a the cumulative impacts accumulate over time. The short-run impact of recession on inventories is positive in the short-run. We find that inventories are most sensitive to the third lag of sales. Hence, the effect of a shock on the sales is experienced by inventory farther out.

In a second managerial application, we jointly 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 pandemic on firms' short- and long-term operational and financial performance. Such estimates can enable managers to make optimal operational policy decisions such as optimal inventory planning, staffing, and cash management in light of various economic disruptions.

References

Alan, Yasin, George P. Gao, Vishal Gaur. 2014. Does inventory productivity predict future stock returns? a retailing industry perspective. *Management Science* **60**(10) 2416–2434. doi:10.1287/mnsc.2014.1897.

Aviv, Yossi. 2003. A time-series framework for supply-chain inventory management. *Operations Research* **51**(2) 210–227. doi:10.1287/opre.51.2.210.12780.

Barth, Mary E., William H. Beaver, Wayne R. Landsman. 2001. The relevance of the value relevance literature for financial accounting standard setting: another view. *Journal of Accounting and Economics* **31**(1-3) 77–104.

Bartik, Alexander W., Marianne Bertrand, Zoe Cullen, Edward L. Glaeser, Michael Luca, Christopher Stanton. 2020. The impact of covid-19 on small business outcomes and expectations. *Proceedings of the National Academy of Sciences* **117**(30) 17656–17666.

Bray, Robert L., Haim Mendelson. 2012. Information transmission and the bullwhip effect: An empirical investigation. *Management Science* **58**(5) 860–875.

Bray, Robert L., Haim Mendelson. 2015. Production smoothing and the bullwhip effect. *Manufacturing & Service Operations Management* **17**(2) 208–220. doi:10.1287/msom.2014.0513.

Buzacott, John A., Rachel Q. Zhang. 2004. Inventory management with asset-based financing. *Management Science* **50**(9) 1274–1292. doi:10.1287/mnsc.1040.0278.

Cachon, Gérard P., Santiago Gallino, Marcelo Olivares. 2019. Does adding inventory increase sales? evidence of a scarcity effect in u.s. automobile dealerships. *Management Science* **65**(4) 1469–1485. doi:10.1287/mnsc.2017.3014.

Chod, Jiri. 2017. Inventory, risk shifting, and trade credit. *Management Science* **63**(10) 3207–3225. doi:10.1287/mnsc.2016.2515.

FASB. 2010. Conceptual framework for financial reporting. statement of financial accounting concepts no. 8 .

Gao, Xiaodan. 2018. Corporate cash hoarding: The role of just-in-time adoption. *Management Science* **64**(10) 4858–4876. doi:10.1287/mnsc.2017.2775.

Gaur, Vishal, Marshall Fisher, Ananth Raman. 2005. An econometric analysis of inventory turnover performance in retail services. *Management Science* **51** 181–194. doi:10.1287/mnsc.1040.0298.

Gupta, Diwakar, Lei Wang. 2009. A stochastic inventory model with trade credit. *Manufacturing & Service Operations Management* **11**(1) 4–18. doi:10.1287/msom.1070.0191.

Hall, Joseph, Evan Porteus. 2000. Customer service competition in capacitated systems. *Manufacturing & Service Operations Management* **2**(2) 144–165.

Hribar, Paul, Daniel W. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* **40**(1) 105–134.

Hu, Qiaohai, Matthew Sobel. 2007. Echelon base-stock policies are financially sub-optimal. *Oper. Res. Lett.* **35** 561–566. doi:10.1016/j.orl.2006.11.001.

Jola-Sanchez, Andres F., Juan Camilo Serpa. 2021. Inventory in times of war. *Management Science* **0**(0) null. doi:10.1287/mnsc.2020.3801.

Kesavan, Saravanan, Vishal Gaur, Ananth Raman. 2010. Do inventory and gross margin data improve sales forecasts for u.s. public retailers? *Management Science* **56**(9) 1519–1533.

Kesavan, Saravanan, Tarun Kushwaha, Vishal Gaur. 2016. Do high and low inventory turnover retailers respond differently to demand shocks? *Manufacturing & Service Operations Management* **18**(2) 198–215.

Kim, Myungsun, William Kross. 2005. The ability of earnings to predict future operating cash flows has been increasing—not decreasing. *Journal of Accounting Research* **43**(5) 753–780.

Li, Lode, Martin Shubik, Matthew J. Sobel. 2013. Control of Dividends, Capital Subscriptions, and Physical Inventories. *Management Science* **59**(5) 1107–1124. doi:10.1287/mnsc.1120.1629.

Lorek, Kenneth, G. Willinger. 2009. New evidence pertaining to the prediction of operating cash flows. *Review of Quantitative Finance and Accounting* **32** 1–15. doi:10.1007/s11156-007-0076-1.

Luo, Wei, Kevin Huei-Min Shang. 2015. Joint inventory and cash management for multidivisional supply chains. *Oper. Res.* **63** 1098–1116.

Lutkepohl, Helmut. 2007. *New Introduction to Multiple Time Series Analysis*. Springer Publishing Company, Incorporated.

Moneta, Alessio, Nadine Chlaß, Doris Entner, Patrik Hoyer. 2011. Causal search in structural vector autoregressive models. *Journal of Machine Learning Research - Proceedings Track* **12** 95–114.

Nallareddy, Suresh, Mani Sethuraman, Mohan Venkatachalam. 2020. Changes in accrual properties and operating environment: Implications for cash flow predictability. *Journal of Accounting and Economics* **69** 101313. doi:10.1016/j.jacceco.2020.101313.

Olivares, Marcelo, Gérard P. Cachon. 2009. Competing retailers and inventory: An empirical investigation of general motors' dealerships in isolated u.s. markets. *Management Science* **55**(9) 1586–1604. doi:10.1287/mnsc.1090.1050.

Pearl, Judea. 2009. *Causality: Models, Reasoning and Inference*. 2nd ed. Cambridge University Press, USA.

Rumyantsev, Sergey, Serguei Netessine. 2007. What can be learned from classical inventory models? a cross-industry exploratory investigation. *Manufacturing & Service Operations Management* **9**(4) 409–429.

Sims, Christopher A. 1980. Macroeconomics and reality. *Econometrica* **48**(1) 1–48.

Spirtes, Peter, Clark Glymour, Richard Scheines. 1993. *Causation, Prediction, and Search*, vol. 81. doi:10.1007/978-1-4612-2748-9.

Tong, Jordan, Gregory DeCroix, Jing-Sheng Song. 2020. Modeling payment timing in multiechelon inventory systems with applications to supply chain coordination. *Manufacturing & Service Operations Management* **22**(2) 346–363. doi:10.1287/msom.2018.0741.

Yang, S. Alex, John R. Birge. 2018. Trade credit, risk sharing, and inventory financing portfolios. *Management Science* **64**(8) 3667–3689. doi:10.1287/mnsc.2017.2799.

Zivot, Eric. 2006. Modeling financial time series with s-plus. *Springer New York* 385–429.

Appendix A: Summary of notations used in the paper

Symbol	Description
i	Operational or Financial variable
$\mathcal{K}$	Set of endogenous variables with size m
$\mathcal{Z}$	Set of exogenous variables with size n
T	Number of time periods in the model
y_{it}	Time series for variable i
$\mathbf{y_t}$	Multi-variate time series
L	Lag operator
$\mathcal{B}(L)$	Matrix-valued polynomial in the positive powers of L
$\mathbf{B}$	An $m \times m$ matrix encoding contemporaneous relationships among the variables in $\mathcal{K}$
$\mathbf{\Gamma_k}$	$(n+m) \times (n+m)$ matrices of lagged coefficients in the structural form
$\mathbf{u_t}$	Vector housing the orthogonal structural shocks,
$\mathbf{\Sigma}$	Covariance matrix of structural shocks, assumed to be diagonal
$\boldsymbol{\xi}_\mathbf{t}$	Vector of forecast errors, estimated by ordinary least squares
$\mathbf{\Pi_k}$	Reduced-form matrices
$\mathbf{\Omega}$	Covariance matrix of forecast errors
$\mathbf{x_t}$, $\mathbf{z_t}$	Multi-variate time series for endogenous and exogenous variables, respectively
$(\mathfrak{s}_\mathbf{t}, \mathfrak{i}_\mathbf{t}, \mathfrak{c}_\mathbf{t})$	Sales, Inventory and Cash Series
$(\mathfrak{ap}_\mathbf{t}, \mathfrak{gdp}_\mathbf{t}, \mathfrak{sga}_\mathbf{t})$	Accounts Payable, GDP and SGA series
$\mathbf{\Theta}_m$	IRF matrices
$\mathbf{B_{11}}$	Leading principal sub matrix of order n, of $\mathbf{B}$
H	Hessian matrix for log-likelihood
$\mathbf{y^*_{t+1}}$	MSE minimizing forecast using VAR
$\text{FEVD}_{i,j}(h)$	Contribution of shock u_j to forecast error variance of variable i, h periods in the future

Table 7 **Notations used in the paper.**

A.1. Compustat Variable Definitions

Table 8 Compustat variable codes and definitions

Variable code	Definition
OANCFY	Year to date net cash flow from operating activities. Increases (decreases) in cash are presented as +ve (-ve) numbers. The item is a sum of Accounts Payable and Accrued Liabilities, Accounts Receivable, Assets and Liabilities, Deferred Taxes, Depreciation, Funds from Operations , Income Before Extraordinary Items, Income Taxes, Inventory, Sale of PP&E etc
OANCFQ	Quarterly cash flows. Equal to OANCFY[i] for i=1 and OANCFY[i]-OANCFY[i-1] for i=2,3,4, where i is the quarter
INVTQ	End of quarter total inventory. Sum of finished goods inventory raw materials inventory, work in progress inventory, and other inventory
SALEQ	quarterly gross sales, the amount of actual billings to customers for regular sales reduced by cash discounts, trade discounts, and returned sales and allowances for which credit is given to customers
APQ	accounts payable at the end of the quarter. Includes Accounts and notes payable, Banks and savings and loans' total deposits, Trade notes checks outstanding, Brokerage houses' payable to brokers, dealers, and clients
XSGAQ	all commercial expenses of operation incurred in the regular course of business pertaining to the securing of operating income. Includes Advertising expense, Bad debt expense, Commissions, Engineering Expense Marketing expense, Freight-out expense, Strike expense, Directors' fees and remuneration Research revenue, Parent company charges for administrative services

Appendix B: Estimation Appendix

B.1. ADF tests

We test the presence of a unit root in all our series using an Augmented Dickey-Fuller (ADF) test. ADF test tests for the presence of a unit root in the time-series. Specifically, for an AR process below, it tests the following null hypothesis:

$$\begin{aligned} \Delta y_t &= \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t \qquad (20) \\ H_0 : \gamma &= 0 \end{aligned}$$

where, y_t is an AR series. $\gamma = 0$ implies that the series has a unit-root. We run the ADF tests for all the five series, and for all firms in our sample. Table 9 shows the percentage of firms that have a unit-root in the original series (column 2). Column 3 shows the percentage of firms that reject the null-hypothesis after first differencing. The ADF test statistics are evaluated for a 95% confidence interval.

Table 9 Augmented Dickey-Fuller Test Results.

Series	Original Series	First-Differenced Series
$\mathfrak{s}_t$	87.5	92.3
$\mathfrak{i}_t$	87.8	90.9
$\mathfrak{c}_t$	47.2	98.8
$\mathfrak{ap}_t$	86.3	96.1
$\mathfrak{sga}_t$	95.7	92.5

B.2. Example - Estimation Steps for a single firm

B.2.1. Reduced form Table 10 presents the reduced form estimates for PG's endogenous variables. We note that the R^2 of the model is really high, indicating that unobservables, if any outside our model, do not explain much of the variance in the variables. Also, we note that inventory has a high correlation with the 3rd lag sales, and sales has a high correlation with 1st lag inventory.

Table 10 Reduced form estimates from the model. .1, .2 etc represent the numbered lag of the variable (*,,*** indicates statistical significance at 10%,5%,1% level)**

	Endogenous Variable		
	Sales	*Invt*	*Cash*
	(1)	(2)	(3)
Sales.1	−0.090 (0.177)	0.140 (0.090)	-0.037 (0.102)
Invt.1	0.978*** (0.368)	0.025 (0.187)	−0.139 (0.212)
Cash.1	0.221 (0.264)	0.034 (0.134)	−1.259*** (0.152)
Ap.1	−0.134 (0.341)	−0.170 (0.173)	−0.020 (0.197)
Gdp.1	1.424 (2.074)	1.857* (1.054)	−0.410 (1.195)
Sga.1	0.639 (0.763)	−0.601 (0.388)	0.383 (0.440)
Sales.2	−0.063 (0.157)	0.005 (0.080)	0.108 (0.091)
Invt.2	−0.633* (0.358)	−0.237 (0.182)	0.222 (0.207)
Cash.2	−0.361 (0.377)	0.091 (0.192)	−1.258*** (0.218)
Gdp.2	1.156 (2.480)	−0.883 (1.260)	−0.024 (1.429)
Sga.2	1.128 (0.704)	−0.545 (0.358)	0.806* (0.406)
Sales.3	−0.123 (0.149)	0.273*** (0.076)	−0.285*** (0.086)
Invt.3	−0.010 (0.349)	−0.056 (0.177)	−0.281 (0.201)
Cash.3	−0.261 (0.403)	0.009 (0.205)	−1.013*** (0.232)
Ap.3	0.043 (0.305)	0.111 (0.155)	−0.154 (0.176)
Gdp.3	−1.588 (2.459)	0.458 (1.250)	1.361 (1.417)
Sga.3	0.100 (0.708)	−0.311 (0.360)	0.939** (0.408)
Sales.4	0.640*** (0.170)	−0.019 (0.086)	0.038 (0.098)
Invt.4	0.410 (0.349)	0.302* (0.178)	0.437** (0.201)
Cash.4	−0.071 (0.332)	0.011 (0.169)	−0.414** (0.191)
Ap.4	−0.264 (0.303)	0.229 (0.154)	−0.522*** (0.175)
Gdp.4	−0.378 (2.506)	−0.746 (1.274)	−0.366 (1.444)
Sga.4	−1.250* (0.721)	−0.168 (0.366)	0.889** (0.416)
Sales.5	−0.190 (0.176)	−0.088 (0.090)	0.034 (0.102)
Invt.5	−0.277 (0.398)	0.297 (0.202)	0.154 (0.229)
Cash.5	−0.134 (0.173)	0.083 (0.088)	−0.209** (0.100)
Ap.5	−0.032 (0.310)	−0.221 (0.158)	−0.353* (0.179)
Gdp.5	2.524 (2.383)	−0.319 (1.211)	−0.203 (1.373)
Sga.5	0.588 (0.807)	0.714* (0.410)	0.292 (0.465)
Observations	116	116	116
Adjusted R^2	0.963	0.949	0.964
F Statistic (df = 30; 86)	102.889***	72.567***	104.780***

The estimated B matrix is shown below:

Table 11 B matrix from the model

	Sales	Invt	Cash	Ap	Gdp	Sga
Sales	1	0.509	-0.542	-0.088	0.052	-0.275
Invt	0.222	1	0.131	-0.257	1.037	0.306
Cash	-0.191	0.304	1	-0.060	0.350	0.173
Ap	0	0	0	1	0	0
Gdp	0	0	0	0	1	0
Sga	0	0	0	0	0	1

Appendix C: Appendix D

Figure 12 MAPE values for endogenous variables across different model specifications for PNG. Blue(red) denotes the MAPE for the VAR(autoregressive) model.

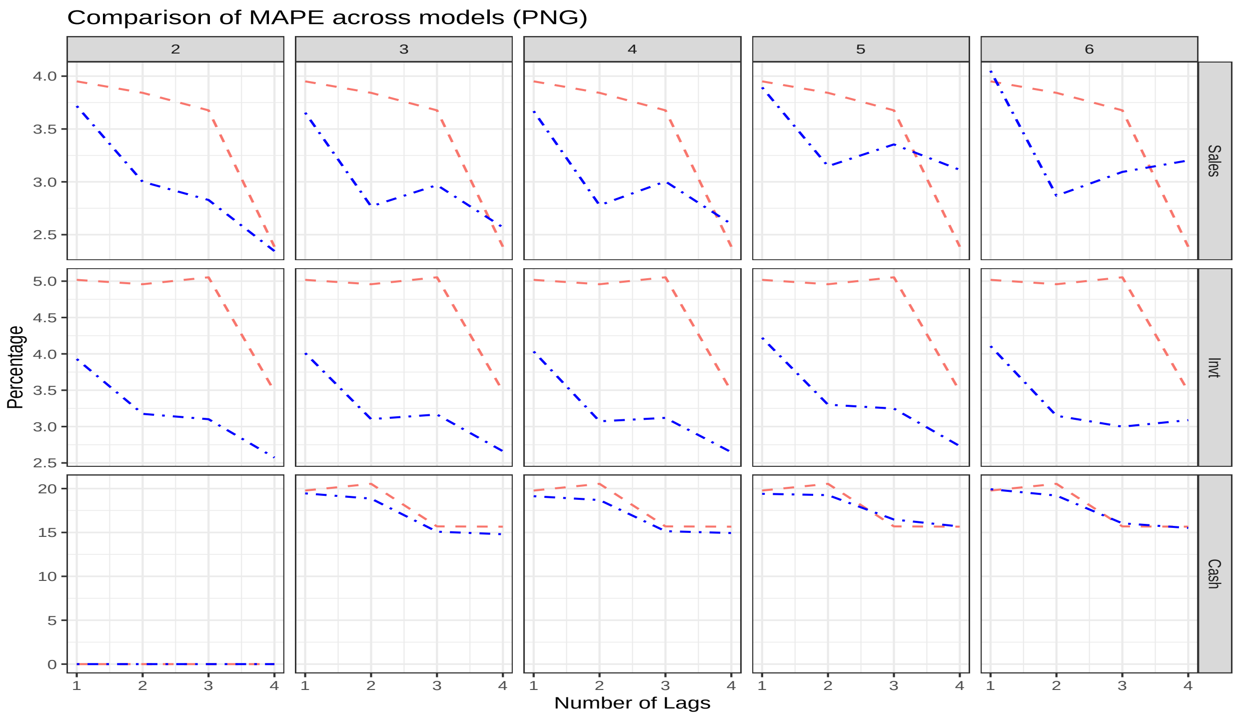


Figure 13 MAPE values for endogenous variables across different model specifications for Gap. Blue(red) denotes the MAPE for the VAR(autoregressive) model.

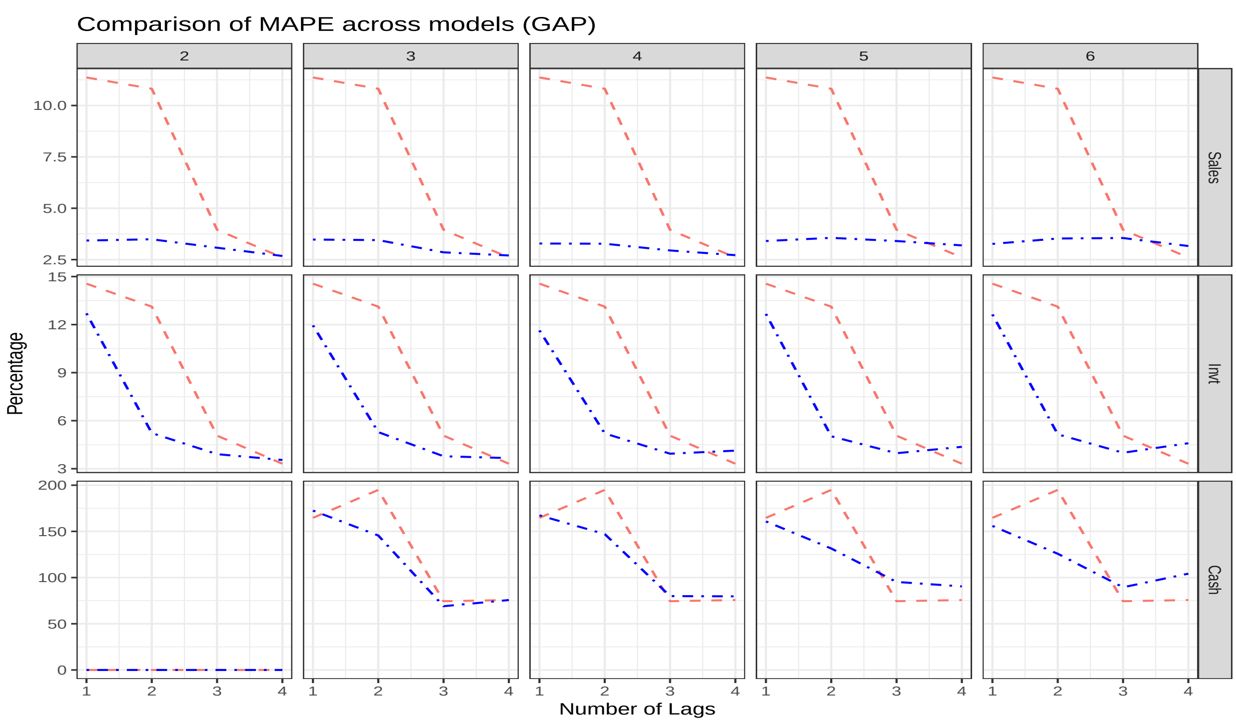


Figure 14 **MAPE values for endogenous variables across different model specifications for Penny (JC). Blue(red) denotes the MAPE for the VAR(autoregressive) model.**

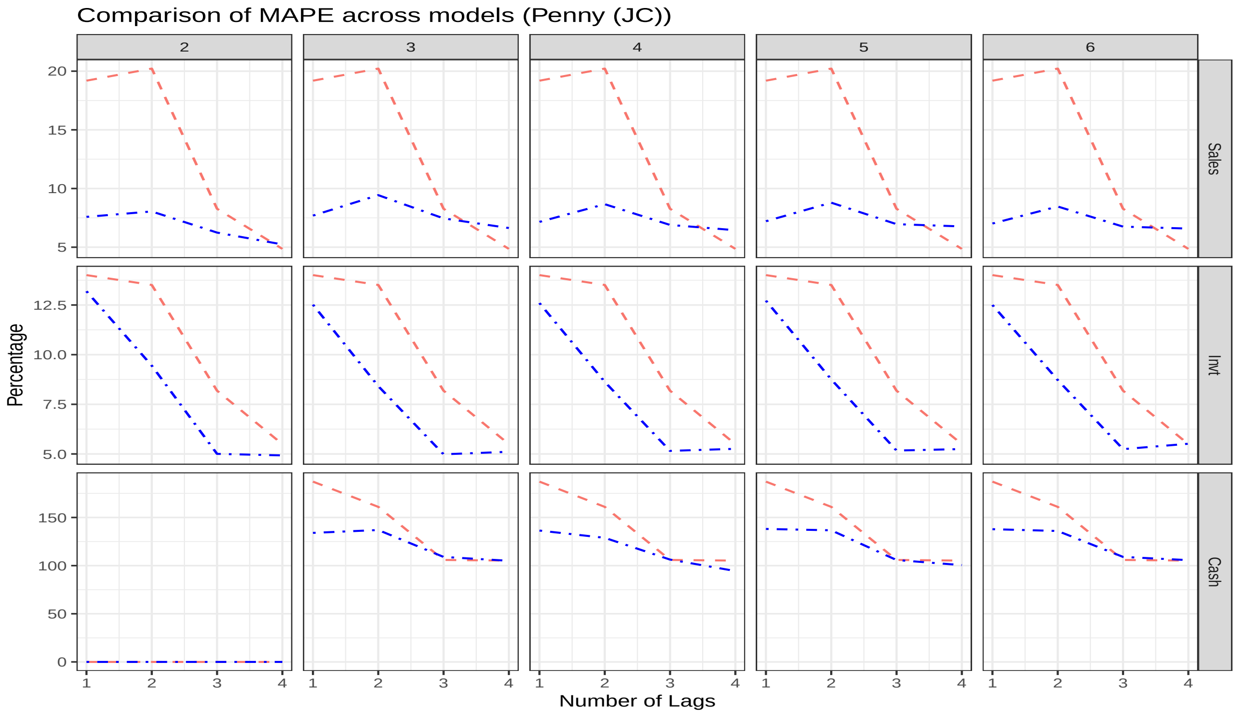


Figure 15 **MAPE values for endogenous variables across different model specifications for Lowe's. Blue(red) denotes the MAPE for the VAR(autoregressive) model.**