ONLINE APPENDIX

Appendix A: Summary of notations used in the paper

Table 7 Notation 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

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

A.2. Data-filtering procedure

The original Compustat dataset comprised 1,376,856 quarterly observations spanning from 1990 to 2020, for 34,019 different firms (totaling a file size of 10.98 GB). This included around 415,000 observations from manufacturing, 58,000 from retail, 35,000 from wholesale, and 867,000 from other sectors. The second step of the filtering process involved narrowing the focus solely to Retail, Wholesale, and Manufacturing firms, reducing the observations to 508,240 across 11,483 firms, with the same sector-specific observation counts as before. The final step further refined the data by selecting only those firms that had at least 120 data points and less than 5 NA values. This resulted in a significantly reduced dataset of 69,554 observations, with manufacturing reduced to around 59,000 observations, retail to around 5,800, and wholesale to around 4,000.

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:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t \tag{18}$$
$$H_0 : \gamma = 0$$

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 fourth 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 11 presents the reduced form estimates for Macy's endogenous variables. We note that the R^2 of the reduced form regressions is usually high. Also, we note that sales has a high correlation with 2nd lag inventory. The estimated B matrix is shown below:

Table 10 Estimated B matrix for Macy's

	Internal Firm Variables					Macro Variables		
	Sales	Inv	Cash	AP	SGA	GDP	CCI	CEO
Sales	1.0000000	0.0924788	0.129811	0.0905234	0.1106809	0.1255154	0.1220526	0.110048
Inv	0.0728184	1.0000000	0.101901	0.0569131	0.0957521	0.1205583	0.0000000	0.136518
Cash	0.1082107	0.1138357	1.000000	0.1011800	0.0847788	0.1045308	0.0000000	0.000000
AP	0.0897532	-0.0014297	0.117951	1.0000000	0.1015468	0.0930723	0.0000000	0.122582
SGA	0.0943611	0.0604414	0.121327	0.0975563	1.0000000	0.1060621	0.0660291	0.113653
GDP	0.0000000	0.0000000	0.000000	0.0000000	0.0000000	1.0000000	0.0000000	0.000000
CCI	0.0000000	0.0000000	0.000000	0.0000000	0.0000000	0.1098808	1.0000000	0.109712
CEO	0.0000000	0.0000000	0.000000	0.0000000	0.0000000	0.1080036	0.1097048	1.000000

Table 11 Reduced form estimates from the model. .1, .2 etc represent the numbered lag of the variable (*,,*** indicates statistical significance at 10%,5%,1% level)**

	sales	invt	cash	ap	sga
	(1)	(2)	(3)	(4)	(5)
sales_1	0.864*** (0.224)	0.272 (0.319)	0.206 (0.182)	0.235 (0.286)	0.104 (0.082)
invt_1	0.433** (0.164)	0.536** (0.234)	0.424*** (0.133)	0.056 (0.210)	0.154** (0.060)
cash_1	−0.066 (0.145)	0.040 (0.207)	0.146 (0.118)	0.218 (0.186)	−0.037 (0.053)
ap_1	0.266 (0.163)	0.002 (0.231)	0.305** (0.132)	0.278 (0.208)	0.020 (0.059)
sga_1	0.042 (0.517)	−0.160 (0.737)	−0.497 (0.420)	0.226 (0.661)	0.576*** (0.189)
gdp_1	80.785 (60.430)	−1.424 (86.044)	41.831 (49.060)	−23.504 (77.245)	28.388 (22.064)
cci_1	8.569* (4.544)	4.304 (6.470)	0.288 (3.689)	2.282 (5.808)	4.644*** (1.659)
ceo_1	−5.186 (6.266)	−1.497 (8.922)	−3.408 (5.087)	7.799 (8.009)	−3.035 (2.288)
sales_2	−0.142 (0.235)	0.154 (0.335)	−0.256 (0.191)	0.218 (0.301)	−0.111 (0.086)
invt_2	−0.647*** (0.197)	−0.228 (0.280)	−0.204 (0.160)	−0.092 (0.252)	−0.152** (0.072)
cash_2	0.074 (0.128)	−0.064 (0.183)	0.129 (0.104)	0.113 (0.164)	0.034 (0.047)
ap_2	−0.105 (0.185)	0.024 (0.263)	−0.295* (0.150)	0.071 (0.236)	−0.011 (0.067)
sga_2	0.086 (0.569)	0.021 (0.810)	0.014 (0.462)	−0.670 (0.727)	0.241 (0.208)
gdp_2	−73.317 (55.063)	−73.656 (78.402)	24.590 (44.703)	−38.117 (70.385)	−27.181 (20.105)
cci_2	−3.991 (5.108)	−0.141 (7.273)	−1.366 (4.147)	0.689 (6.530)	−3.533* (1.865)
ceo_2	1.479 (6.612)	−0.584 (9.415)	2.061 (5.368)	−1.619 (8.452)	1.285 (2.414)
sales_3	−0.063 (0.223)	0.173 (0.318)	0.008 (0.181)	−0.001 (0.286)	−0.096 (0.082)
invt_3	0.242 (0.210)	0.141 (0.299)	0.176 (0.170)	0.235 (0.268)	0.123 (0.077)
cash_3	−0.115 (0.130)	0.167 (0.185)	−0.223** (0.105)	0.148 (0.166)	0.018 (0.047)
ap_3	−0.047 (0.194)	−0.145 (0.276)	0.123 (0.158)	−0.320 (0.248)	−0.068 (0.071)
sga_3	0.280 (0.580)	−0.162 (0.826)	0.900* (0.471)	0.160 (0.742)	0.241 (0.212)
gdp_3	74.241 (56.559)	39.109 (80.533)	−6.950 (45.918)	50.364 (72.298)	46.306** (20.651)
cci_3	0.740 (5.048)	7.758 (7.187)	0.585 (4.098)	2.979 (6.452)	−0.202 (1.843)
ceo_3	−11.075 (6.673)	−10.259 (9.502)	−3.790 (5.418)	−2.775 (8.531)	−2.746 (2.437)
sales_4	−0.088 (0.208)	−0.031 (0.296)	−0.199 (0.169)	−0.199 (0.266)	−0.001 (0.076)
invt_4	0.040 (0.221)	−0.665** (0.314)	0.003 (0.179)	−0.608** (0.282)	−0.140* (0.081)
cash_4	0.271** (0.122)	0.144 (0.174)	−0.120 (0.099)	0.164 (0.156)	0.045 (0.045)
ap_4	−0.236 (0.223)	−0.032 (0.317)	0.016 (0.181)	0.191 (0.285)	0.085 (0.081)
sga_4	−0.405 (0.576)	−0.241 (0.821)	−0.817* (0.468)	0.319 (0.737)	−0.428** (0.210)
gdp_4	−19.302 (58.881)	−40.505 (83.839)	45.324 (47.803)	−64.803 (75.265)	−6.133 (21.499)
cci_4	−3.636 (4.855)	−10.180 (6.913)	−2.449 (3.942)	−9.589 (6.206)	−0.973 (1.773)
ceo_4	13.344** (6.413)	11.738 (9.131)	−1.084 (5.206)	10.609 (8.197)	3.218 (2.341)
sales_5	0.324* (0.174)	0.288 (0.248)	−0.079 (0.142)	0.346 (0.223)	0.083 (0.064)
invt_5	−0.098 (0.209)	0.321 (0.298)	0.153 (0.170)	−0.023 (0.267)	0.012 (0.076)
cash_5	−0.207* (0.120)	−0.304* (0.170)	−0.076 (0.097)	−0.163 (0.153)	−0.124*** (0.044)
ap_5	0.068 (0.234)	−0.178 (0.333)	−0.187 (0.190)	0.211 (0.299)	0.042 (0.085)
sga_5	−0.149 (0.516)	0.131 (0.734)	0.621 (0.419)	−0.447 (0.659)	0.066 (0.188)
gdp_5	−51.317 (65.465)	−107.532 (93.214)	2.630 (53.148)	−57.614 (83.682)	−10.788 (23.903)
cci_5	3.061 (4.087)	6.495 (5.820)	−0.374 (3.318)	10.790** (5.224)	1.500 (1.492)
ceo_5	−4.034 (5.135)	−1.217 (7.312)	1.884 (4.169)	−3.690 (6.564)	−3.278* (1.875)
Observations	115	115	115	115	115
R^2	0.868	0.640	0.786	0.610	0.854
Adjusted R^2	0.797	0.448	0.672	0.401	0.776

*p<0.1; **p<0.05; ***p<0.01

B.3. Evolution of data due to a shock by a competitor's entry

Table 12 Evolution of Data for a shock in sales at time $t = 1$.

Time	Sales Shock	Invt Shock	Sales (log diff.)	Invt (log diff.)	Sales (Level)	Invt (Level)
-1					1,000,000	250,000
0			0.00000	0.00000	**1,000,000**	**250,000**
1	-0.1	0	-0.09412	0.02353	910,176	255,952
2	0	0	-0.00637	-0.00194	904,399	255,457
3	0	0	-0.00313	-0.00034	901,573	255,370
4	0	0	-0.00107	-0.00013	900,610	255,337
5	0	0	-0.00038	-0.00005	900,271	255,325
6	0	0	-0.00013	-0.00002	900,152	255,321
7	0	0	-0.00005	-0.00001	900,110	255,320
8	0	0	-0.00002	0.00000	900,096	255,319
9	0	0	-0.00001	0.00000	900,090	255,319
10	0	0	0.00000	0.00000	900,089	255,319
11	0	0	0.00000	0.00000	**900,088**	**255,319**
12	0	0	0.00000	0.00000	900,088	255,319
13	0	0	0.00000	0.00000	900,088	255,319

Table 13 Mitigating actions using inventory boosting and the corresponding data evolution

Time	Sales Shock	Invt Shock	Sales (log diff.)	Invt (log diff.)	Sales (Level)	Invt (Level)
-1					1,000,000	250,000
0			0.00000	0.00000	**1,000,000**	**250,000**
1	-0.1	0	-0.09412	0.02353	910,176	255,952
2	0	0.1	0.01716	0.09218	925,932	280,667
3	0	0	0.07493	0.00603	997,980	282,364
4	0	0	0.02400	0.00300	1,022,225	283,213
5	0	0	0.00851	0.00102	1,030,958	283,503
6	0	0	0.00298	0.00036	1,034,039	283,605
7	0	0	0.00105	0.00013	1,035,123	283,641
8	0	0	0.00037	0.00004	1,035,504	283,653
9	0	0	0.00013	0.00002	1,035,638	283,658
10	0	0	0.00005	0.00001	1,035,685	283,659
11	0	0	0.00002	0.00000	1,035,702	283,660
12	0	0	0.00001	0.00000	1,035,707	283,660
13	0	0	0.00000	0.00000	1,035,709	283,660
14	0	0	0.00000	0.00000	1,035,710	283,660
15	0	0	0.00000	0.00000	1,035,710	283,660
16	0	0	0.00000	0.00000	**1,035,711**	**283,660**
17	0	0	0.00000	0.00000	1,035,711	283,660
18	0	0	0.00000	0.00000	1,035,711	283,660

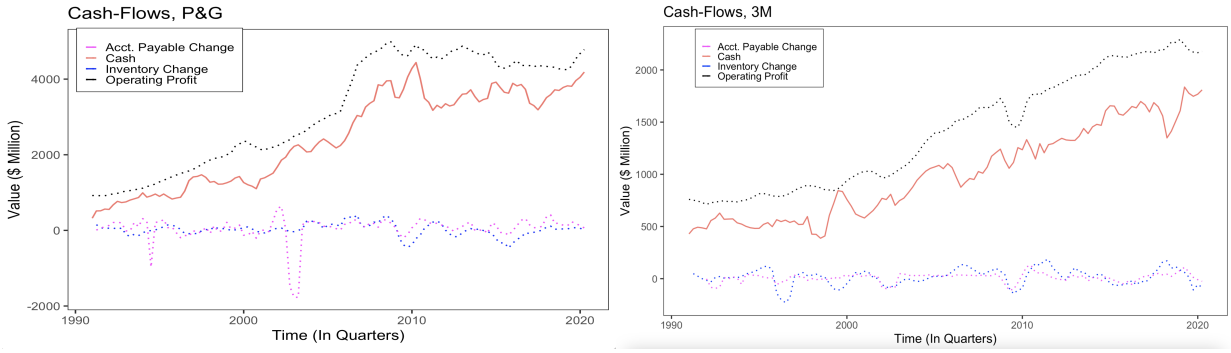
Appendix C: Additional Analysis on Manufacturing firms

We add additional analyses with a focus on manufacturing firms and present the results from these analyses below. Specifically, we (i) provide model-free evidence of relationships for manufacturing firms, (ii) provide model-based evidence of relationships and examine impulse responses to manufacturing firms using our model, as you suggested, and (iii) include forecasting case studies for top manufacturers.

1. Evidence for relationships between the variables: We provide a model-free evidence of the relationships between different operational variables for manufacturing firms. Figure 12 shows the four-quarter moving average of operating cash flow for P&G and 3M. We find that cash flows for both firms are strongly correlated with operating profit (= sales − cost of goods sold − SGA) and with changes in inventory and accounts payables. The corresponding correlation coefficients of cash flow with operating profit, change

in inventory, and accounts payables are 0.97, -0.13, and 0.13 for P&G and 0.97, 0.08, and 0.17 for 3M, respectively. The correlation provides model-free evidence of contemporaneous relationships.

Figure 12 Four-quarter moving average of cash flows for P&G and 3M (red lines). The dotted black lines in black, blue and magenta represent operating profits, change in inventory and accounts payables.

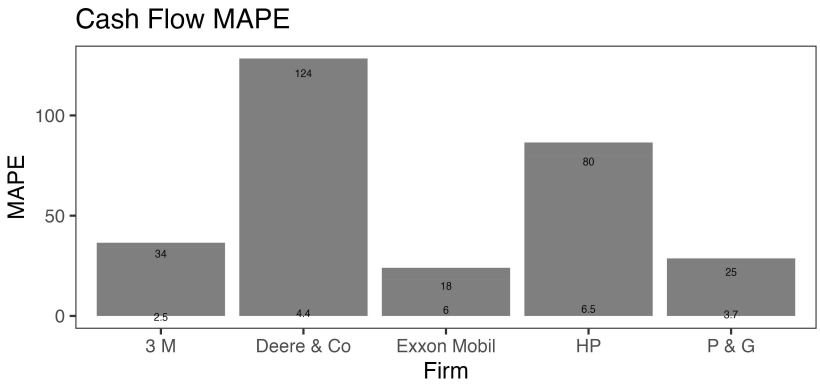


(a) Four-quarter moving average of cash flows for P&G (b) Four-quarter moving average of cash flows for Walmart

2. Model-based evidence and impulse responses from the model: We also provide model-based evidence of contemporaneous relationships among the variables by estimating a panel-data version of the model (please see §4.3). This estimation includes all the firms (including manufacturing firms) in the sample. We find that all our proposed linkages are statistically significant, providing direct evidence for the variables' endogeneity. We also estimate our model for manufacturing firms and examine the impulse responses post-estimation. We include examples of impulse response functions for manufacturers like P&G and 3M in Figures 16 and 15. Overall, we find that our qualitative insights for the impulse response functions remain largely the same, with small variations in some of the estimates.

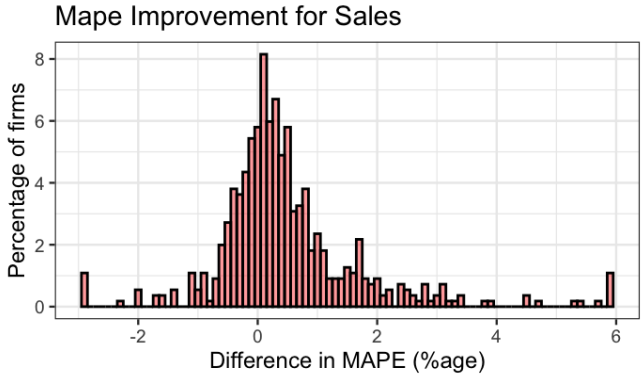
3. Forecasting case studies for manufacturers: We also use our model to forecast cash flows for manufacturing firms. First, we highlight that our finding that cash flow forecasting is hard also applies to manufacturing firms. Figure 13 shows the MAPE for cash flow forecasts for large U.S. manufacturers like 3M, Exxon Mobil, HP, etc. We also depict the corresponding MAPE number for sales forecasts at the bottom of the bars. We find that the MAPE numbers for cash flow forecasts are an order of magnitude higher than the corresponding MAPE for sales forecasts.

Figure 13 MAPE for cash flow forecasts for different manufacturers. The number at the bottom of the bars is the MAPE for sales forecasts.

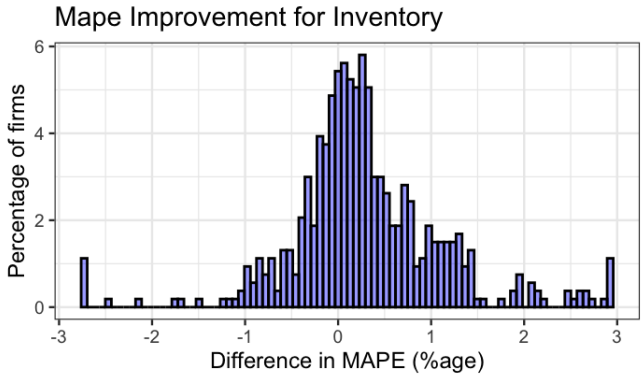


In addition to these styled case-studies, we ran our forecasting exercise for the entire set of firms in our sample (which also includes the manufacturers). Figure 14 shows the distribution of $\text{MAPE}_{\text{AR}} - \text{MAPE}_{\text{VAR}}$ for sales, inventory, and cash forecasts across all firms in the sample. We find that the VAR model achieves an improvement in sales, inventory, and cash flow forecasts for about 68%, 66%, and 63% of firms in the sample, respectively. The average improvement in sales MAPE is 0.49%, and the average improvements in inventory and cash flow forecasts are about 0.29% and 19.13%, respectively.

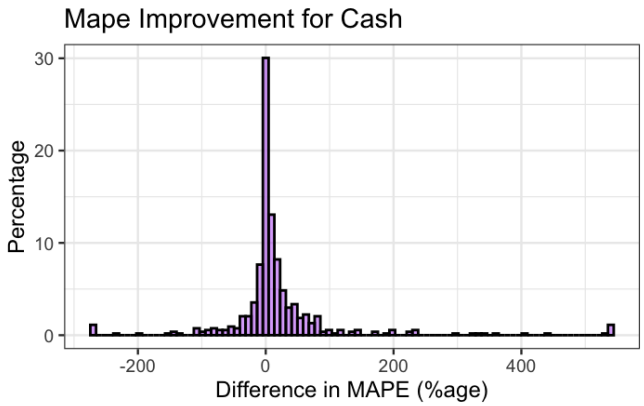
Figure 14 These figures show the distribution of improvement in MAPE for (a) Sales, (b) Inventory, and (c) Cash flow forecasts for the firms in our sample. X-axis represents the difference in MAPE calculated as AR MAPE



(a) Improvement in Sales MAPE



(b) Improvement in inventory MAPE



(c) Improvement in Cash MAPE

Figure 15 IRF estimates for self- and cross-variable impacts for 3M. X-axes represent time in quarters and Y-axes represent the point impacts.

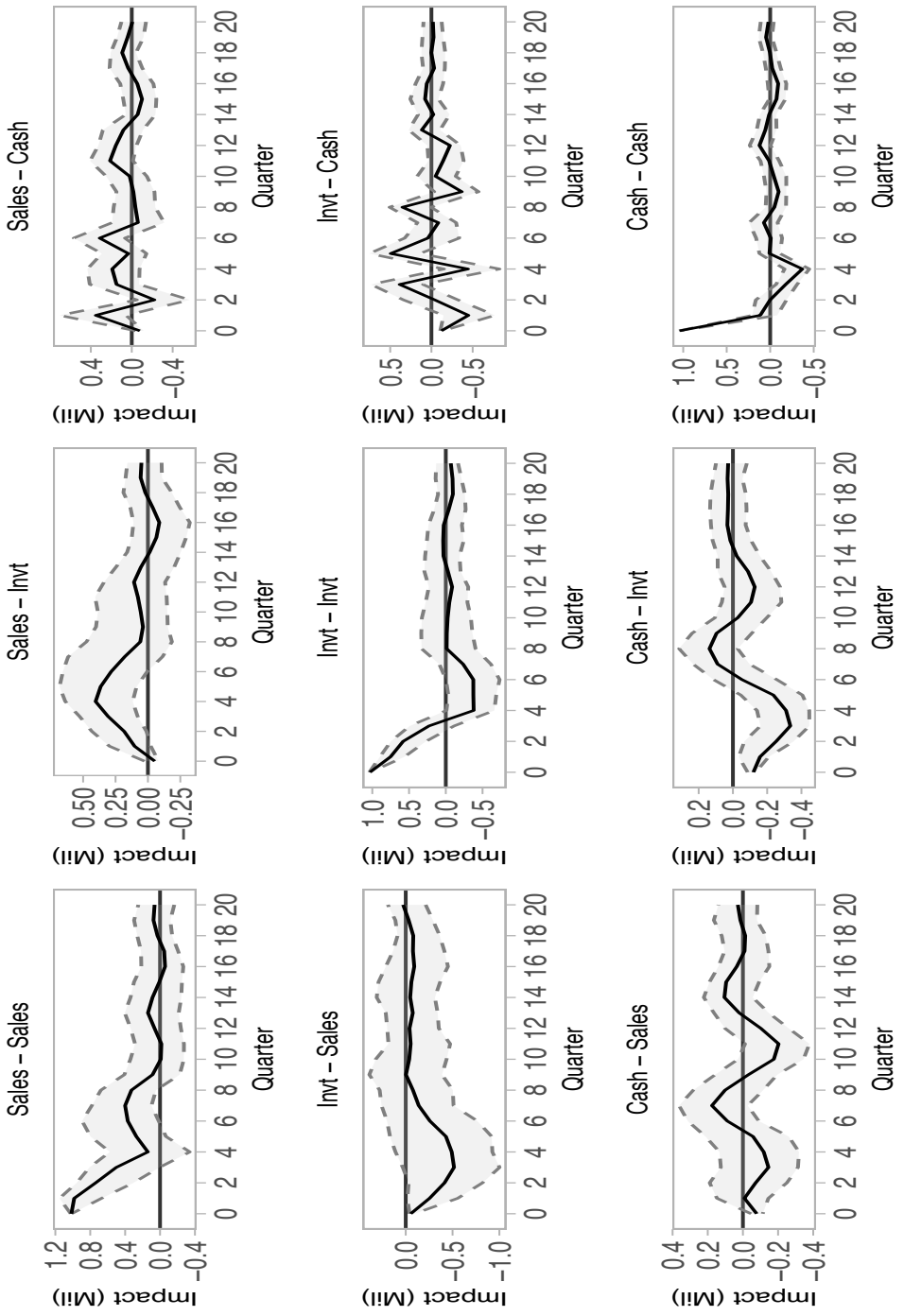
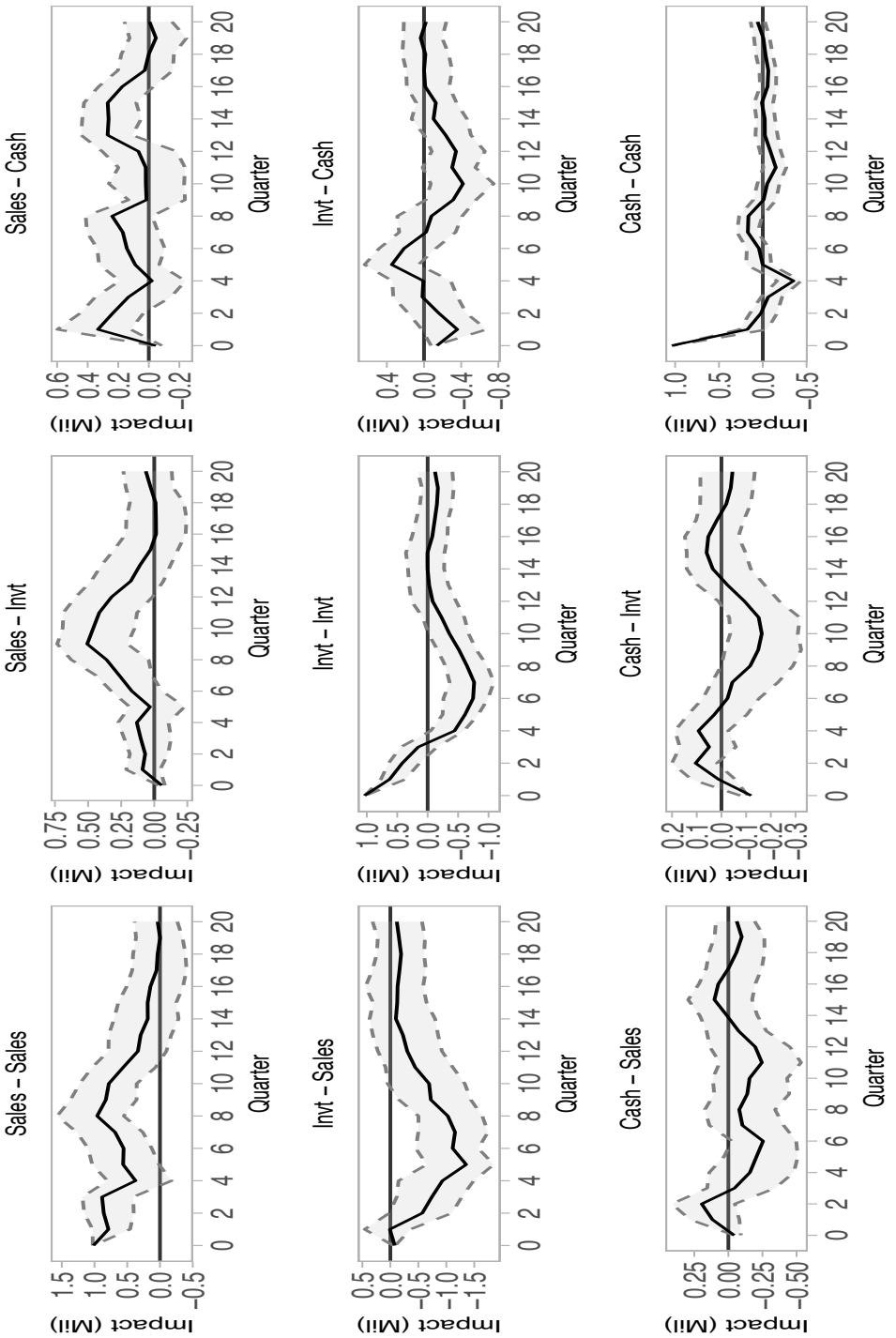


Figure 16 IRF estimates for self- and cross-variable impacts for P&G. X-axes represent time in quarters and Y-axes represent the point impacts.



Appendix D: Additional Forecasting Results

Figure 17 **MAPE values for endogenous variables across different model specifications for PNG. Blue(red) denotes the MAPE for the VAR(autoregressive) model.**

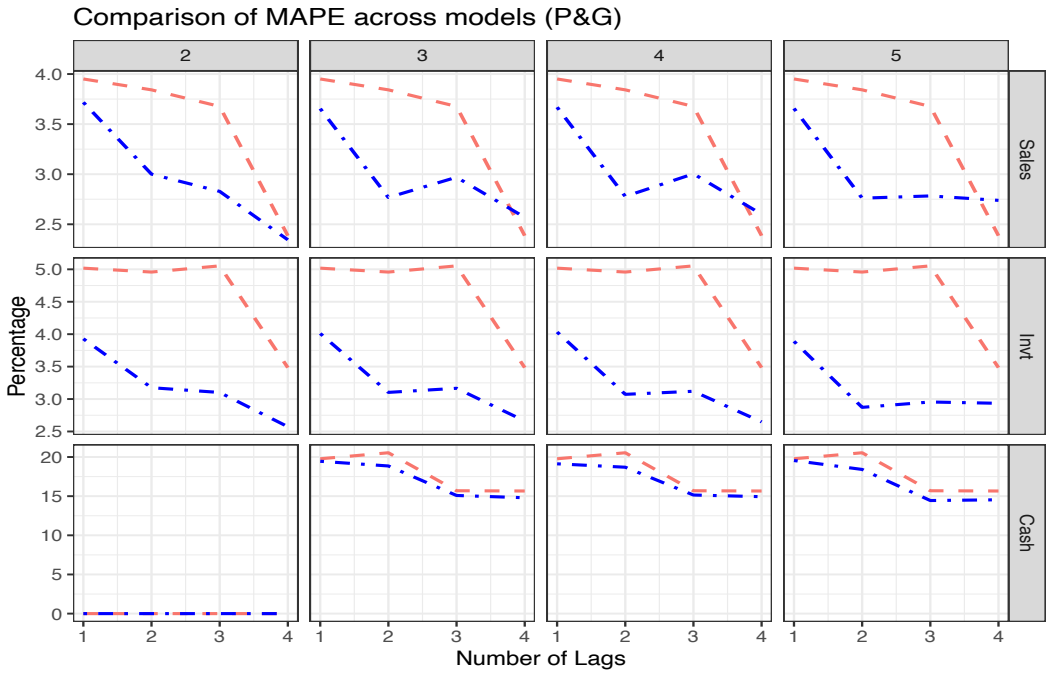


Figure 18 **MAPE values for endogenous variables across different model specifications for GAP. Blue(red) denotes the MAPE for the VAR(autoregressive) model.**

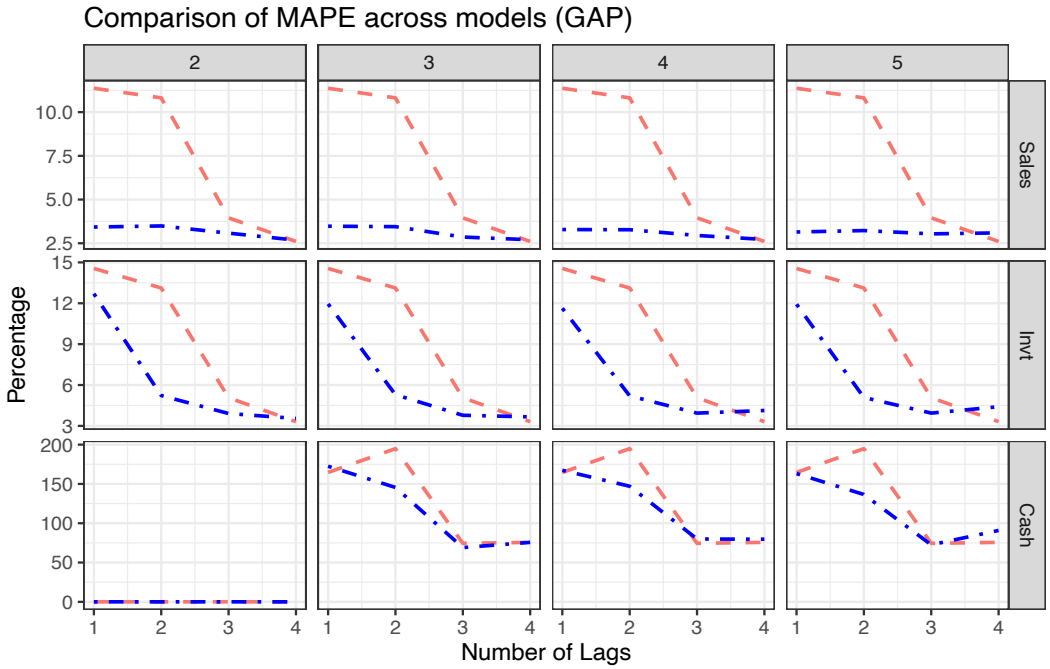


Figure 19 **MAPE values for endogenous variables across different model specifications for Penny (JC). Blue(red) denotes the MAPE for the VAR(autoregressive) model.**

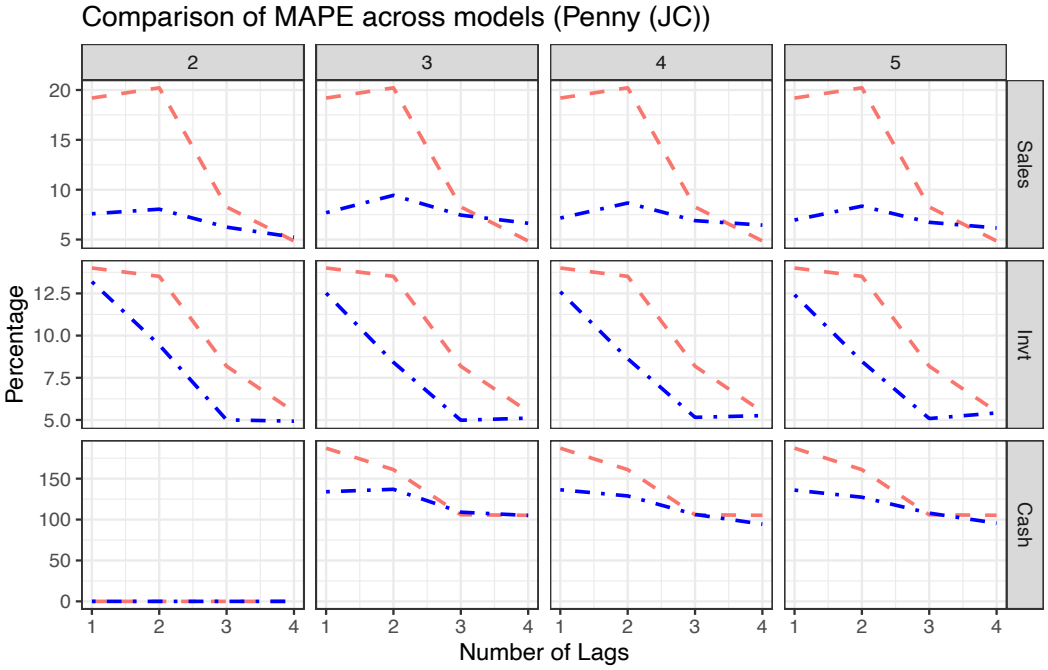


Figure 20 **MAPE values for endogenous variables across different model specifications for Lowe's. Blue(red) denotes the MAPE for the VAR(autoregressive) model.**