Supplementary Appendix: A Time-varying Bayesian Variable Selection for Macroeconomic Forecasting

Hyun Jae Stephen Chu, Jaeho Kim, Kyu Ho Kang

March 2025

1 Simulated Application

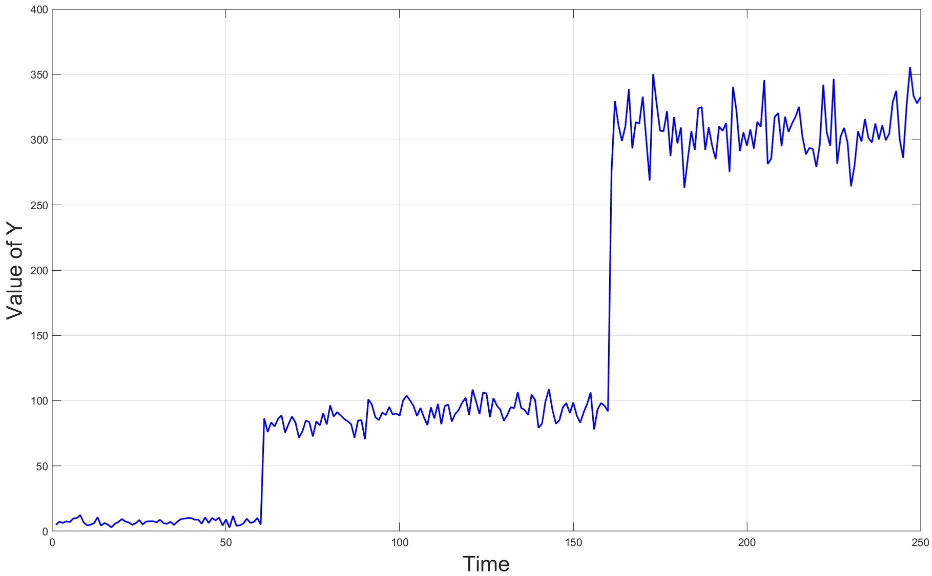


Figure A: Simulated Data of Y

Note. This figure depicts the simulated data for 250 periods using the same data generating process as **Figure 1** in the main text. Y is split into three groups regarding the value of X_1 and X_2.

In this section, we illustrate the in-sample performance of the B-DART model by applying it to a simulated data set corresponding to the tree depicted as **Figure 1** in the main text. Although, as stated in Rossi (2013), there is numerous evidence that a strong

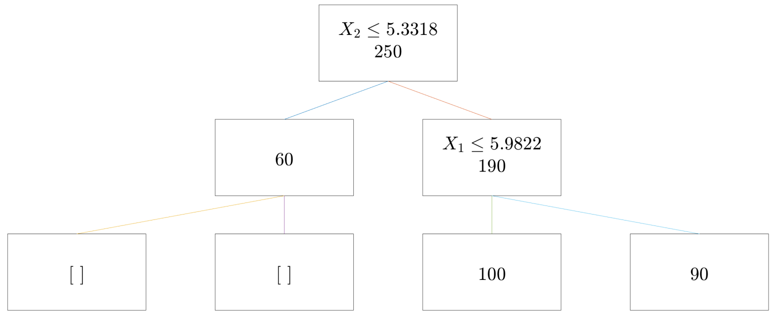


Figure B: Estimated Optimal Tree Structure by the B-DART Model

Note. This figure depicts the optimal tree structure estimated by the B-DART model. One can observe that the structure is well-estimated as the data generating process.

in-sample fit does not necessarily guarantee robust out-of-sample forecast accuracy, it is equaly important to recognize that poor in-sample fits can undermine out-of-sample forecasts. Therefore, it is crucial to examine the in-sample performance of the B-DART model. Moreover, this section is worth showing as it demonstrates how various hyperparameter values in the B-DART model are set, which are later employed in the empirical application of forecasting eight macroeconomic variables in section 6.

The data generating process of the simulated data can be expressed as follows.

$$y_t = \left[\sum_{g=1}^{3} (d_t = g) x_t' \beta_g \right] + \varepsilon_t,\ \varepsilon_t \sim \mathcal{N}(0, \sigma_g^2)$$

where

$$\begin{cases} \{\beta_g = (1,2,0,\cdots,0)',\ \sigma_g^2 = 1\} & \text{if } X_1 \leq 0.5 \text{ and } X_2 \leq 0.5 \\ \{\beta_g = (0,5,6,0,\cdots,0)',\ \sigma_g^2 = 2\} & \text{if } X_1 \leq 0.5 \text{ and } X_2 > 0.5 \\ \{\beta_g = (11,12,13,0,\cdots,0)',\ \sigma_g^2 = 4\} & \text{if } X_1 > 0.5 \end{cases}$$

Note that the sample size is set as $T = 250$, to match the actual quarterly data used in the empirical application. Then the simulated data of y_t can be depicted as the above **Figure A**. The hyperparameter values of each prior distribution were set the same as the empirical application case, except for the case of σ_g^2. In particular, we set $\nu = 5$ and $\lambda = 3$ such that $\mathbb{E}[\sigma_g^2] = (\nu \cdot \lambda/2)/(\nu/2 - 1) = 15/3$.

One characteristic of the MH algorithm that generates the candidate trees is that the acceptance ratio, i.e. the MH ratio, is significantly low. This is different from the convention that the MH ratio ranging from 0.4 to 0.8 is an indicator for using a good proposal distribution. The reason of such phenomenon is because the existence of nu-

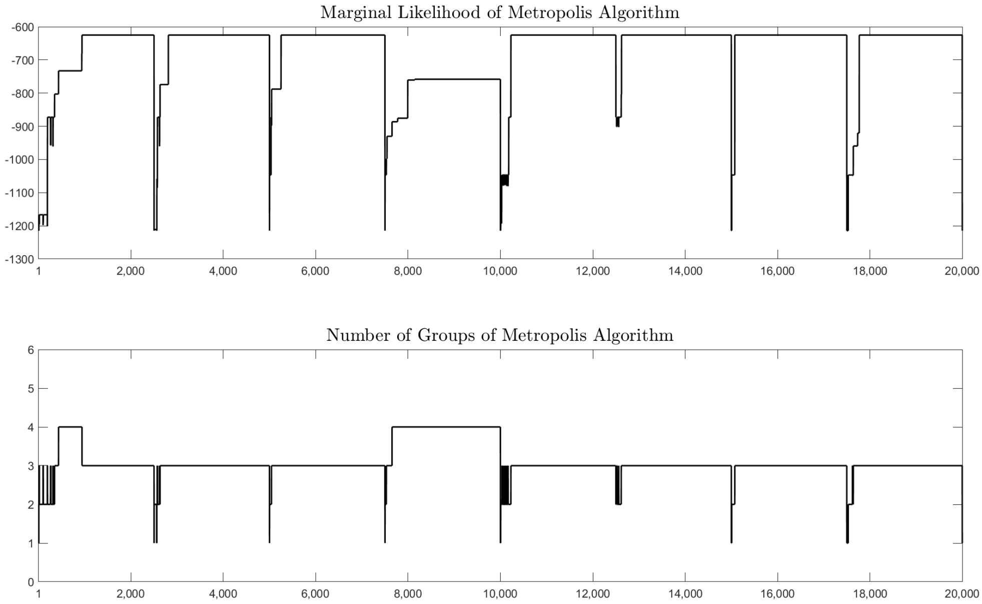


Figure C: Marginal Likelihood and Estimated Number of Groups

Note. This figure depicts marginal likelihood and estimated number of groups of the estimated trees across 20,000 iterations. One can observe a spike every 2,500 times due to resetting the generation of trees.

merous candidate trees, leads to the fact that it is highly likely that the form of a tree with a high conditional marginal likelihood is not persistent. Moreover, the structure of the deep trees, i.e. the trees that are generated afterwards, highly depends on the initial tree. Therefore, we generate $20,000$ trees while resetting it every $2,500$ times.[1] At the terminal node of every generated tree, the sampling process for Dirac variable selection is done $1,000$ iterations after a 10% burn-in.

The estimated optimal tree structure using the B-DART method is as **Figure B**. It can be observed that the tree is well split regarding the values of X_1 and X_2, which is in line with the simulated data. Additionally, it is shown in **Figure C** that the conditional marginal likelihood quickly converges every time the initial tree is reset, as mentioned above. Moreover, the estimated number of groups seem to be consistent for each iteration cycle. The average group is around 3.09 which is related to the value of the stopping rule.

Given the selected tree structure in **Figure B**, one can further estimate the parameters of each group given δ, i.e. using only the selected predictor variable estimated by

[1]Note that decreasing the generation of tree to $10,000$ while resetting it every $2,500$ times does not significantly affect the result. Thus, we later generate $10,000$ trees when working with empirical data for computational ease.

Table A: Comparison of Results

Variable	Dirac	B-CART			B-DART		
		Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
β_1	32.8858	1.0798	0.1681	10.8773	1.0330	0.0009	10.9113
β_2	11.4080	1.9473	5.0407	12.3007	1.9183	5.0450	12.1585
β_3	10.1561	0.2167	6.0618	13.2887	0.0025	6.0771	13.1673
β_4	-0.0007	0.0216	0.0303	-0.1005	0.0001	0.0002	-0.0011
β_5	0.0006	-0.0552	0.0347	-0.1654	-0.0010	0.0001	-0.0197
β_6	0.0031	0.0608	0.0210	0.0092	0.0007	0.0003	0.0001
β_7	-0.0014	0.0165	-0.0123	0.0482	0.0005	0.0001	0.0004
β_8	0.0008	0.0631	-0.0737	-0.0311	0.0008	-0.0013	-0.0001
β_9	0.0034	0.1099	0.0411	-0.0855	0.0025	0.0003	-0.0010
β_{10}	-0.0016	0.0008	-0.0299	0.0299	0.0001	-0.0002	0.0001
σ^2	657.1956	1.1033	2.0332	4.1300	1.2270	1.9814	4.0148
ML	-1202.505		-634.092			-625.053	

Dirac variable selection. In order to evaluate the performance of the B-DART model, we compare the estimated parameter values of each group with those estimated by using the Dirac variable selection and Bayesian CART (B-CART) model alone. The estimated parameter values of each method is stated in **Table A** above. Note that for both models, the same prior values as the B-DART model was used.

First, the Dirac variable selection model led to the most heavily biased estimators of the true parameters. This shows that overlooking the time varying variable selection leads to biased estimators and the possibility of inaccurate forecasting performance. Above all, the B-DART model not only provides parameter estimates comparable to those derived from the B-CART model, but also demonstrates a significant improvement in the conditional marginal likelihood. For these reasons, the B-DART model exhibits a better in-sample fit than the Dirac variable selection model, while also proving to be a sufficient and robust alternative to the B-CART model.

Finally, the posterior distributions obtained by Dirac variable selection to the simulated data is depicted in **Figure D** below. One can observe that X_1, X_2 and X_3 were estimated to be important, i.e. $\hat{\delta}_k \approx 1$. Conversely, the importance of other predictor variables were close to zero so that they are forced to follow the spike distribution with all its mass at zero.

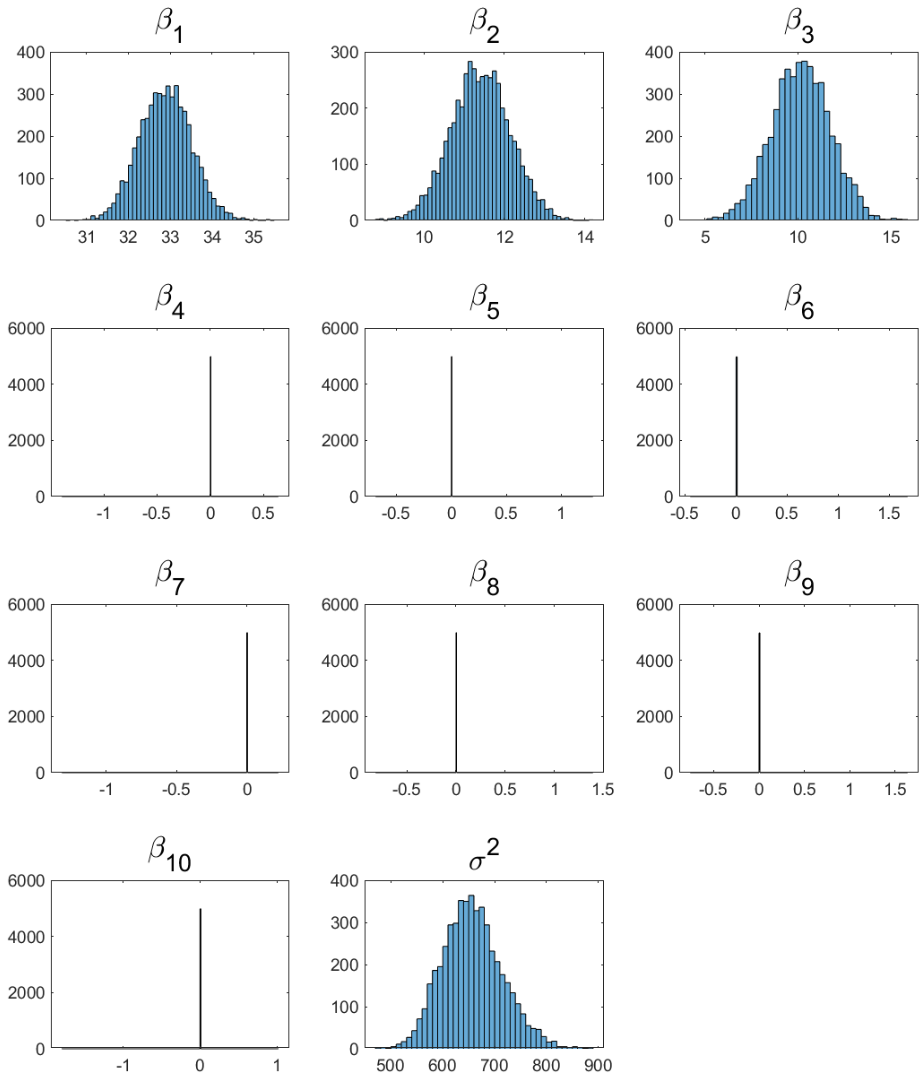


Figure D: Posterior Distributions for Simulated Data

Note. This figure depicts the estimated posterior distributions of each parameter for the simulated data. One can observe that β_1, β_2 and β_3 to follows the slab distribution, while the others to follow the spike distribution of all its mass on 1. However, β_1, β_2 and β_3 are inaccurately estimated as σ^2 is estimated to be extremely large.

2 Hyper-parameter Tuning Process

When generating the trees using the B-DART model, we have found that the tree structure depends greatly in the prior of σ_g^2, i.e.

$$\sigma_g^2 \sim InverseGamma(\frac{\nu}{2}, \frac{\nu\lambda}{2})$$

Therefore, we apply a tuning procedure to the hyper-parameter λ of σ_g^2. Specifically, **?** states that the hyper-parameters of σ_g^2 can be calibrated using the data-based estimate $\hat{\sigma}_g^2$. One method that we apply is the "naive" specification where $\hat{\sigma}_g^2$ is taken to be the sample standard deviation of the response variable Y. We set $\nu = 5$ and tune λ such that

$$\lambda = \left\{ \frac{\hat{\sigma}_g^2}{l} \right\}_{l=1}^{5}$$

Consequently, we tune the hyper-parameter λ five times for each eight forecast horizons of a macroeconomic variable. The selected values of l for each forecast horizon of each macroeconomic variable is shown in **Table B** below. The upper and lower table each illustrates the selected values regarding the RMSE (for point forecasts) and log PPL (for interval forecasts), respectively.

Table B: Selected Value of Hyper-parameter Tuning

(a) Selected Hyper-parameter Value for RMSE Case

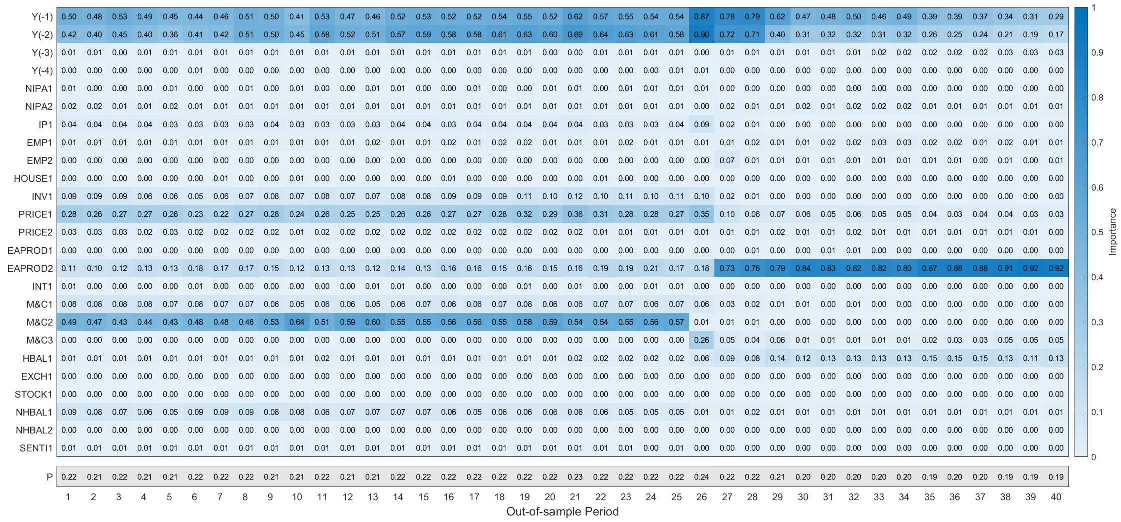
	inf PCE			inf CPI			inf PPI			Crude Oil		
H	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac
1	3	3	5	1	2	4	1	2	2	2	3	1
2	1	2	5	1	5	3	3	4	4	2	2	3
3	5	5	1	5	5	4	4	2	2	4	1	4
4	1	5	5	3	1	5	2	1	2	4	3	2
5	2	5	1	2	5	4	3	4	1	3	1	4
6	4	3	3	1	3	4	1	5	3	2	1	1
7	2	5	4	2	2	5	3	5	1	4	3	4
8	3	3	5	5	1	3	5	2	5	1	3	5

	FFE			Long INT			rGDP			Unemp		
H	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac
1	1	1	3	5	1	1	2	4	4	1	5	4
2	4	3	1	1	1	1	5	3	5	2	1	4
3	2	4	2	5	5	5	3	4	3	5	4	5
4	4	1	1	3	5	2	2	5	2	1	5	3
5	3	4	4	2	4	3	1	4	2	4	2	4
6	2	3	3	1	4	4	2	3	3	4	5	5
7	5	3	4	1	4	1	1	1	4	4	1	3
8	5	3	2	5	2	3	3	3	5	1	4	4

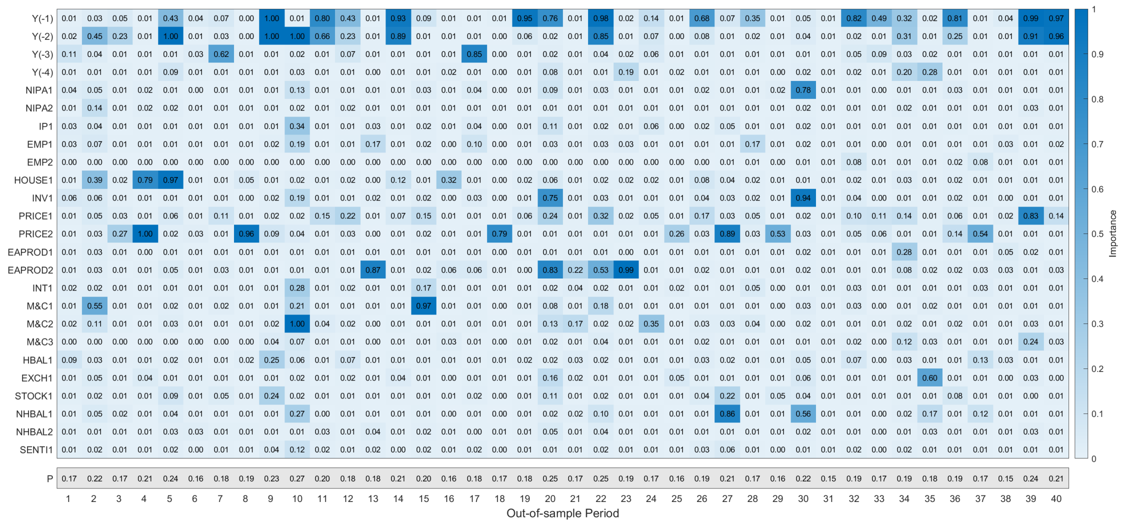
(b) Selected Hyper-parameter Value for log PPL Case

	inf PCE			inf CPI			inf PPI			Crude Oil		
H	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac
1	3	3	5	3	5	1	4	1	4	3	3	1
2	5	2	1	5	5	3	2	2	1	1	2	1
3	4	1	4	5	1	2	1	2	2	4	2	1
4	5	2	5	3	2	2	1	2	1	1	4	1
5	2	5	4	3	1	2	1	4	1	4	5	1
6	2	3	4	4	3	3	5	5	1	1	1	1
7	5	4	5	2	3	4	3	1	1	4	3	1
8	3	2	3	3	1	5	4	5	1	4	1	1

	FFE			Long INT			rGDP			Unemp		
H	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac	AR(1)	B-DART	Dirac
1	3	1	4	5	1	5	3	4	4	2	5	4
2	3	1	1	5	1	4	1	1	5	3	1	2
3	3	4	3	3	2	4	1	5	1	4	3	1
4	4	1	2	2	5	5	1	5	1	1	3	2
5	3	4	4	3	4	3	3	2	1	2	4	1
6	4	1	4	2	4	3	1	2	2	1	3	1
7	2	3	3	5	4	4	1	2	1	2	2	1
8	3	5	4	4	2	2	2	2	1	5	2	1



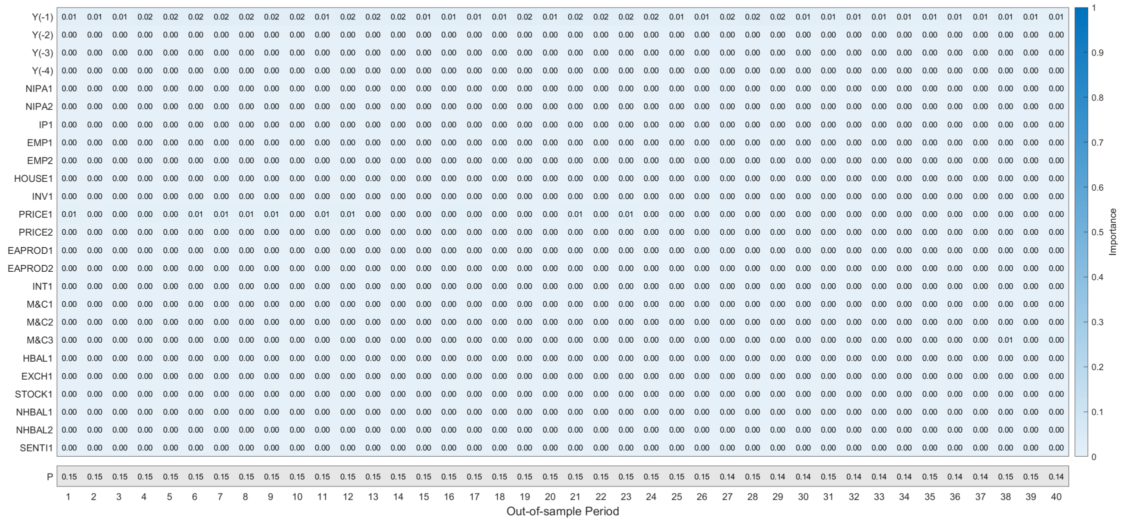
(a) **Dirac Model Case**



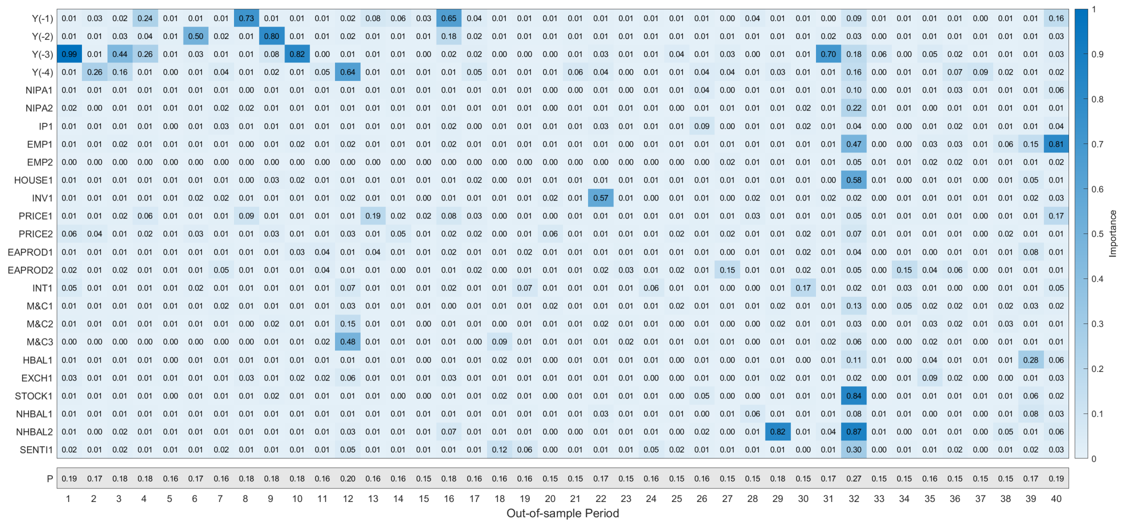
(b) **B-DART Model Case**

Figure E: Importance for Forecasting 1-quarter-ahead PCE Inflation Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 1-quarter-ahead PCE inflation rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



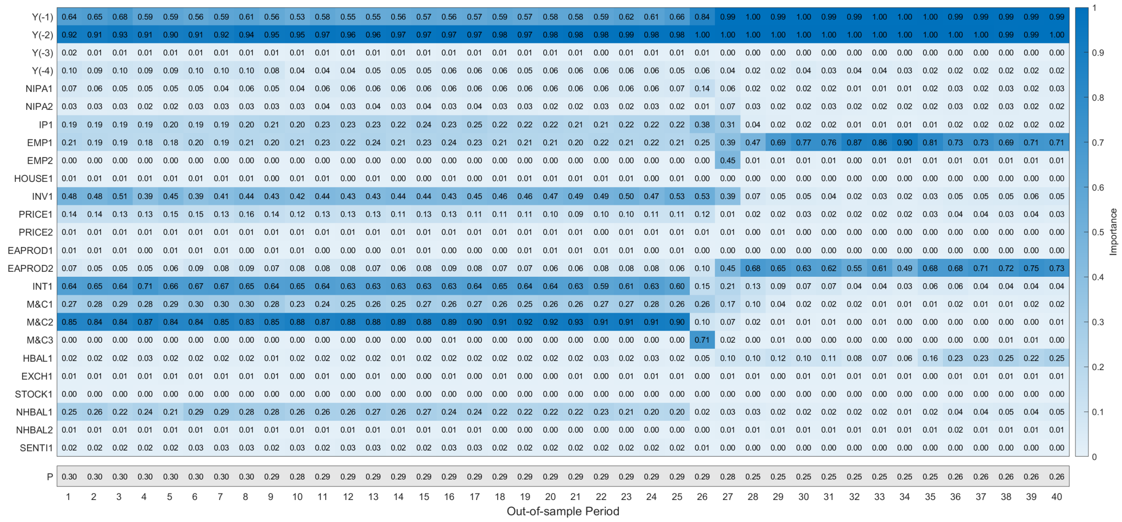
(a) Dirac Model Case



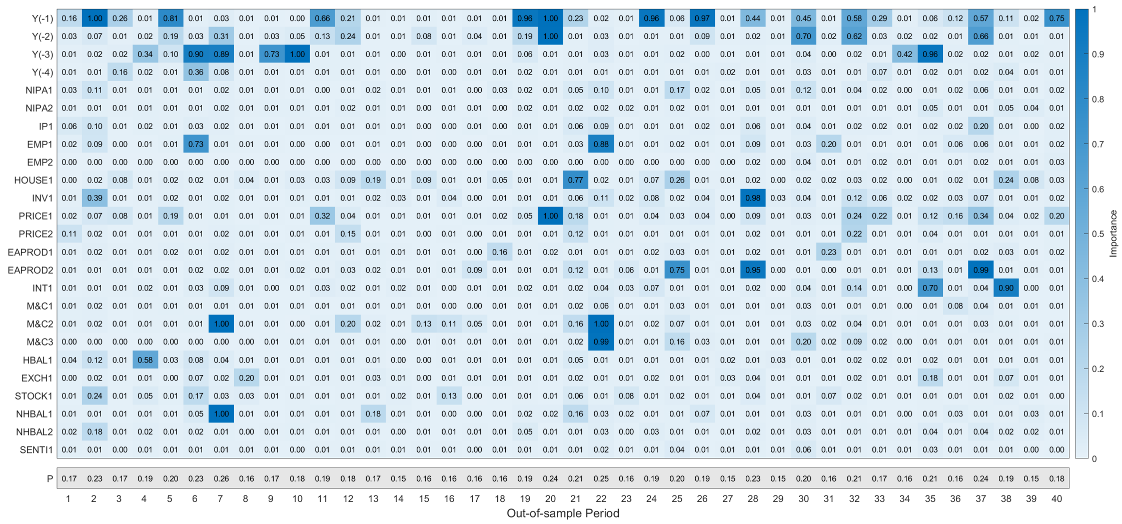
(b) B-DART Model Case

Figure F: Importance for Forecasting 6-quarter-ahead PCE Inflation Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 6-quarter-ahead PCE inflation rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



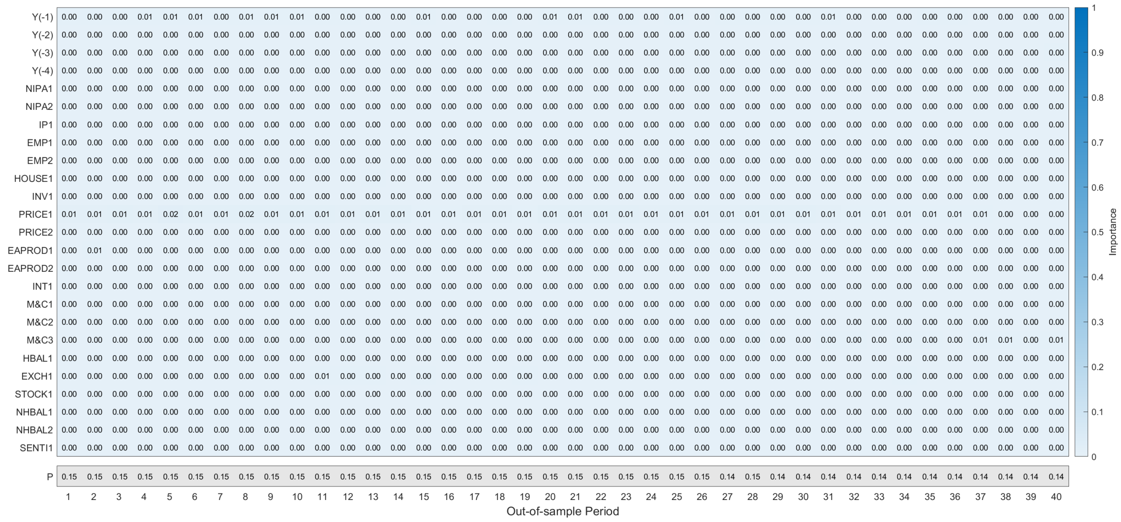
(a) Dirac Model Case



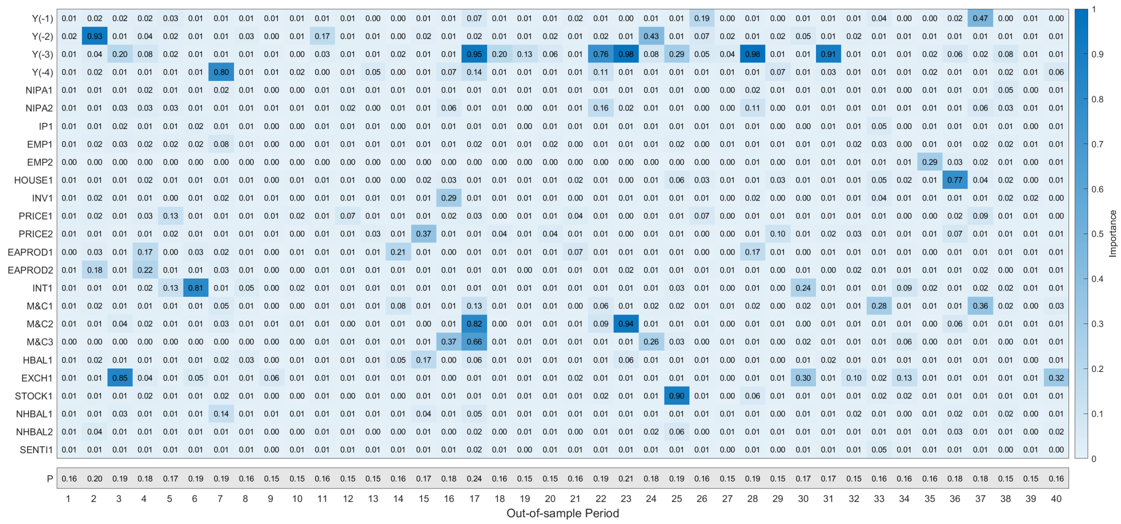
(b) B-DART Model Case

Figure G: Importance for Forecasting 1-quarter-ahead CPI Inflation Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 1-quarter-ahead CPI inflation rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



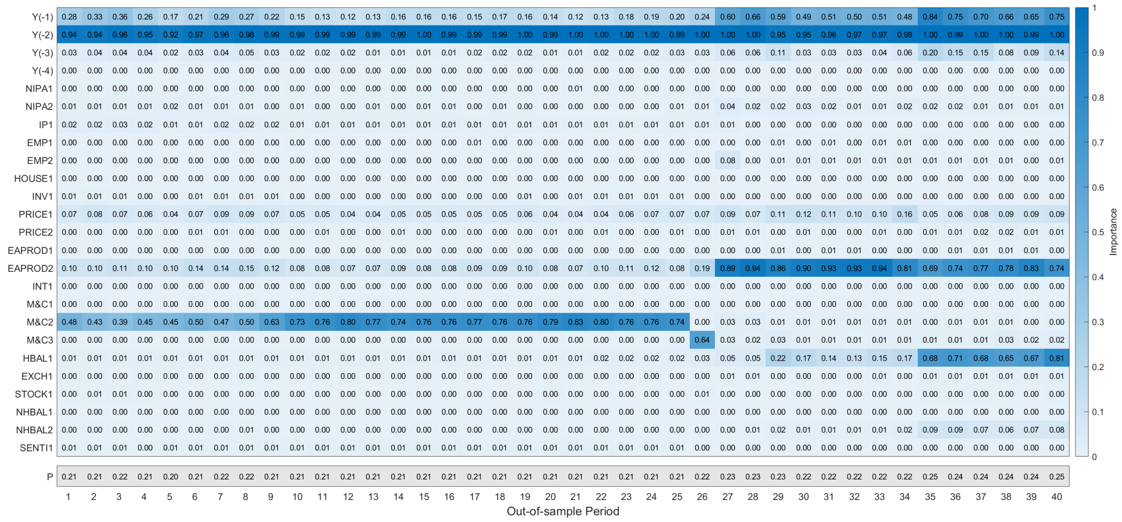
(a) **Dirac Model Case**



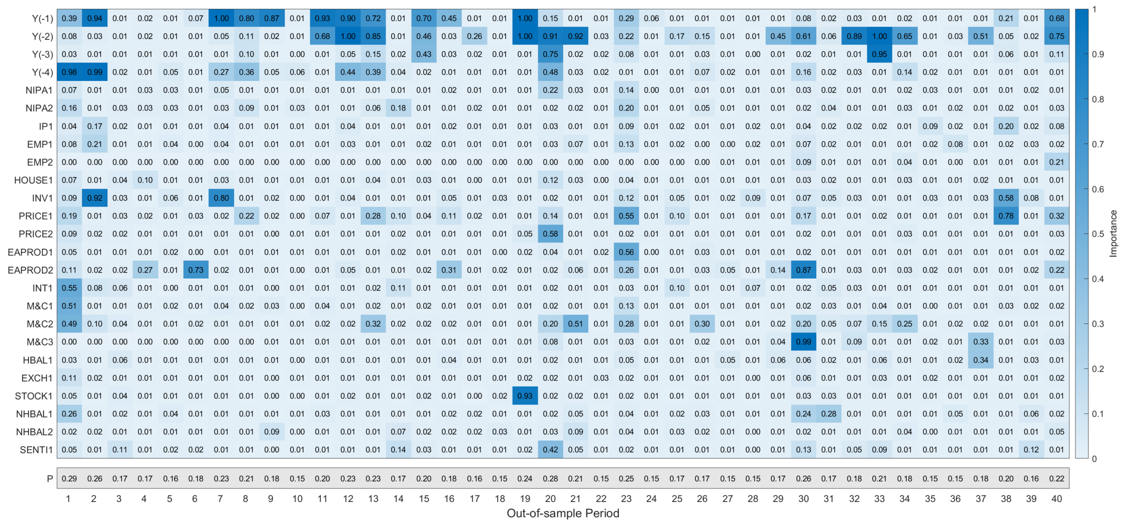
(b) **B-DART Model Case**

Figure H: Importance for Forecasting 6-quarter-ahead CPI Inflation Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 6-quarter-ahead CPI inflation rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



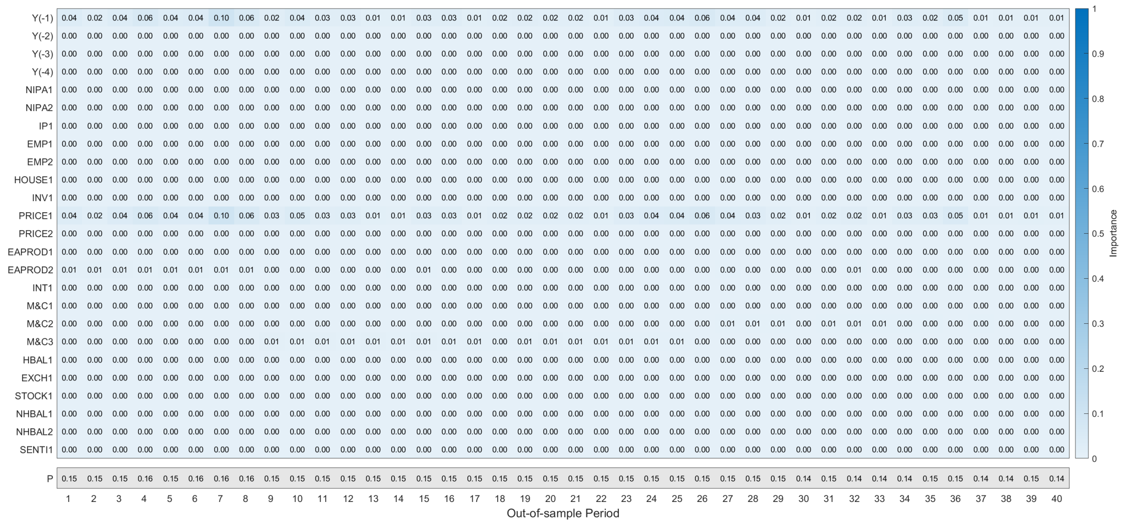
(a) Dirac Model Case



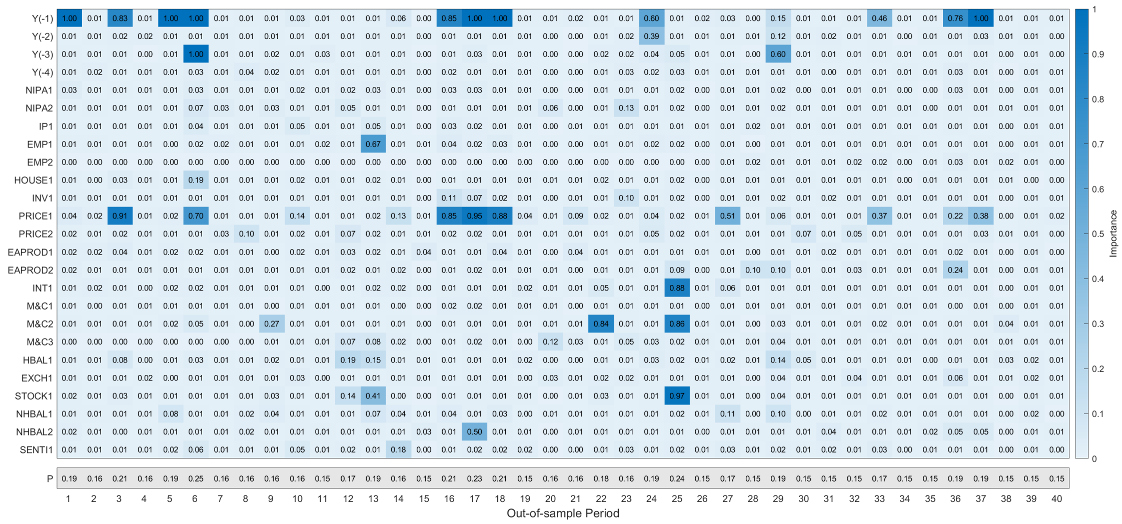
(b) B-DART Model Case

Figure I: Importance for Forecasting 1-quarter-ahead PPI Inflation Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 1-quarter-ahead PPI inflation rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



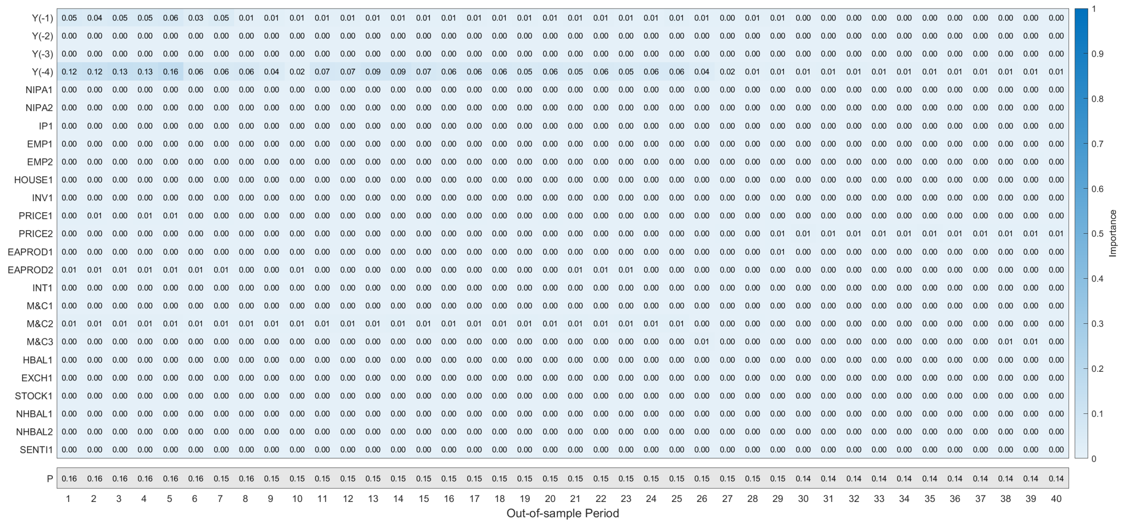
(a) Dirac Model Case



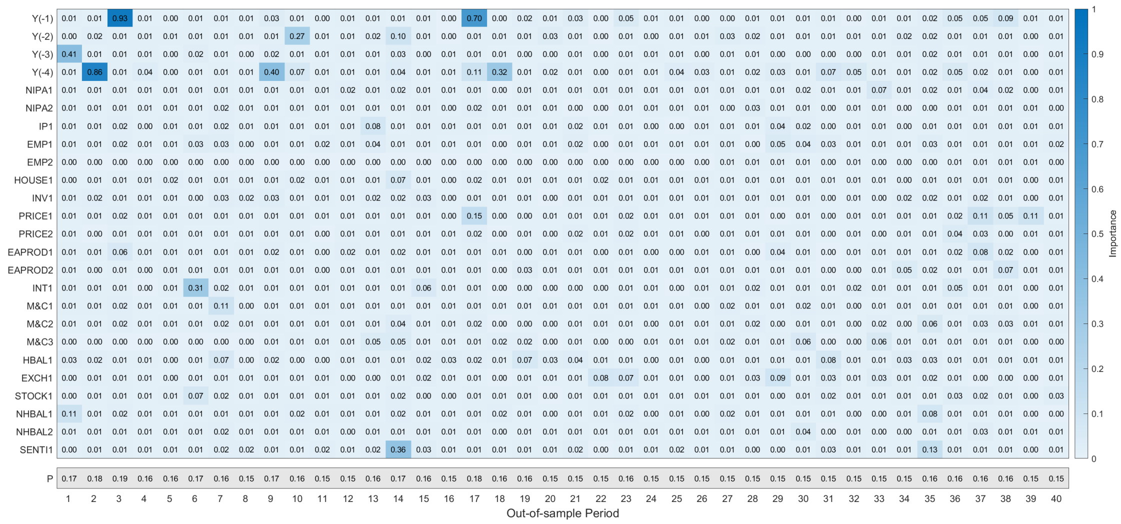
(b) B-DART Model Case

Figure J: Importance for Forecasting 4-quarter-ahead PPI Inflation Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 4-quarter-ahead PPI inflation rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



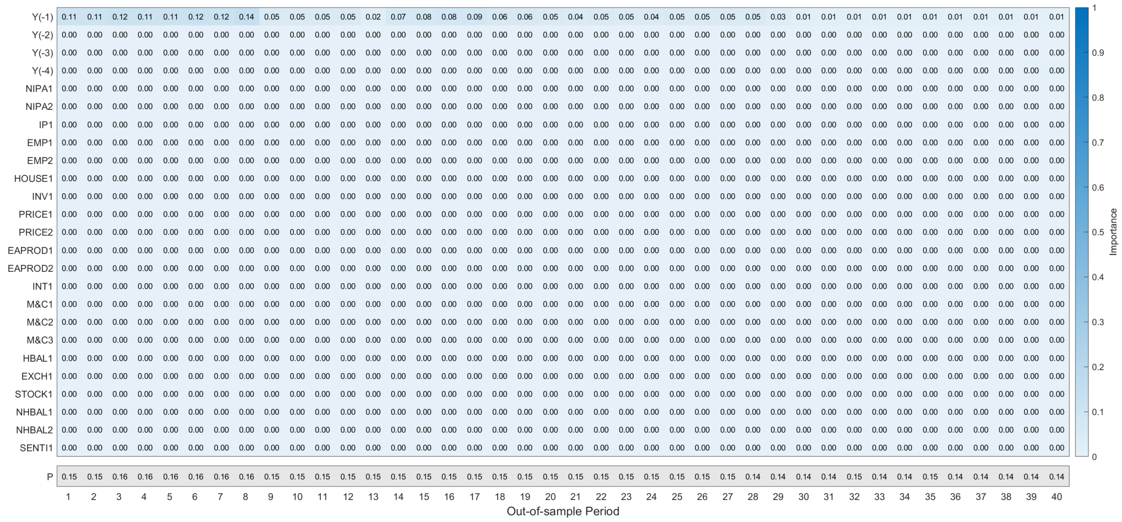
(a) Dirac Model Case



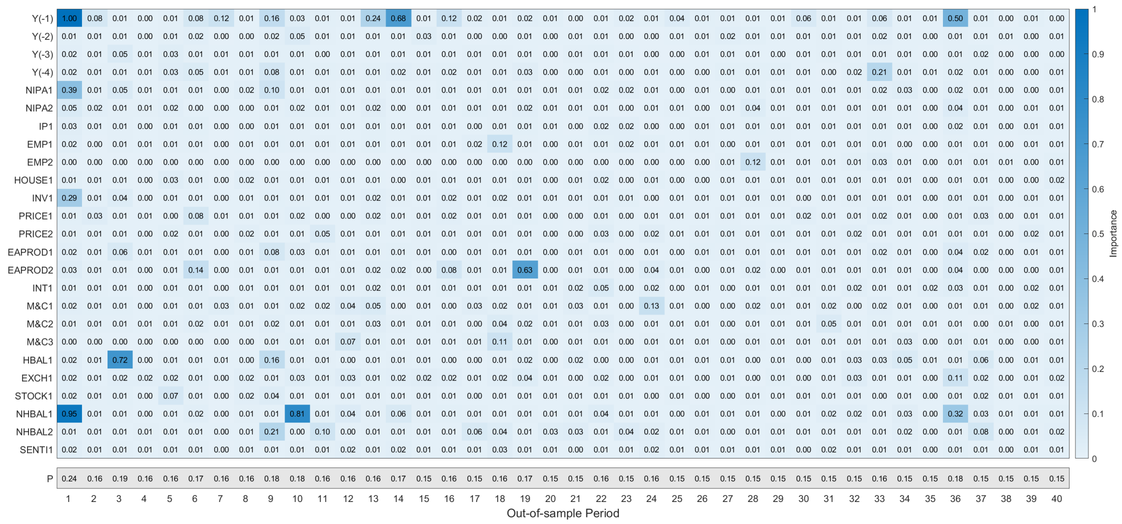
(b) B-DART Model Case

Figure K: Importance for Forecasting 2-quarter-ahead Crude Oil Price

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 2-quarter-ahead Crude Oil Price. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



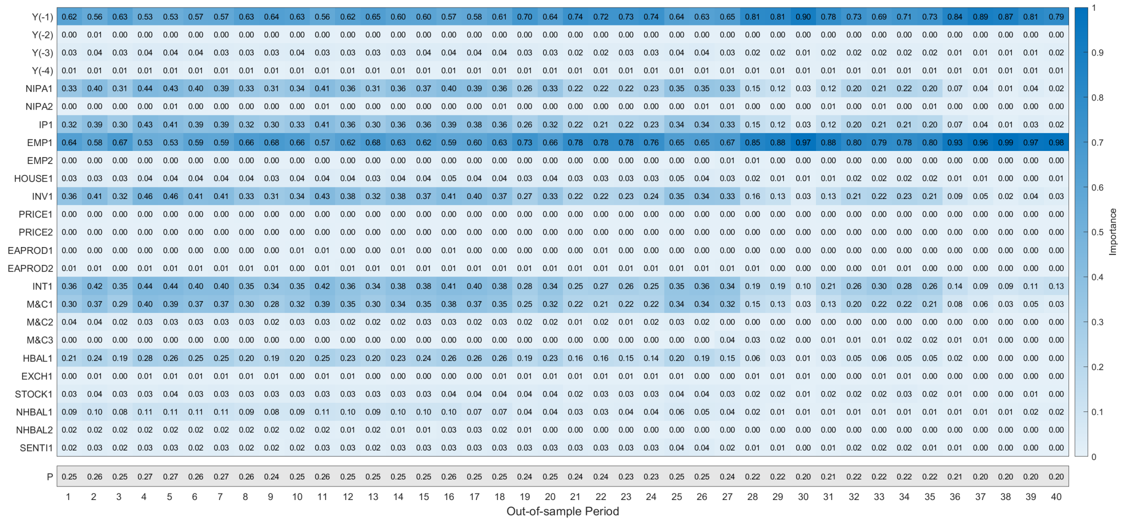
(a) Dirac Model Case



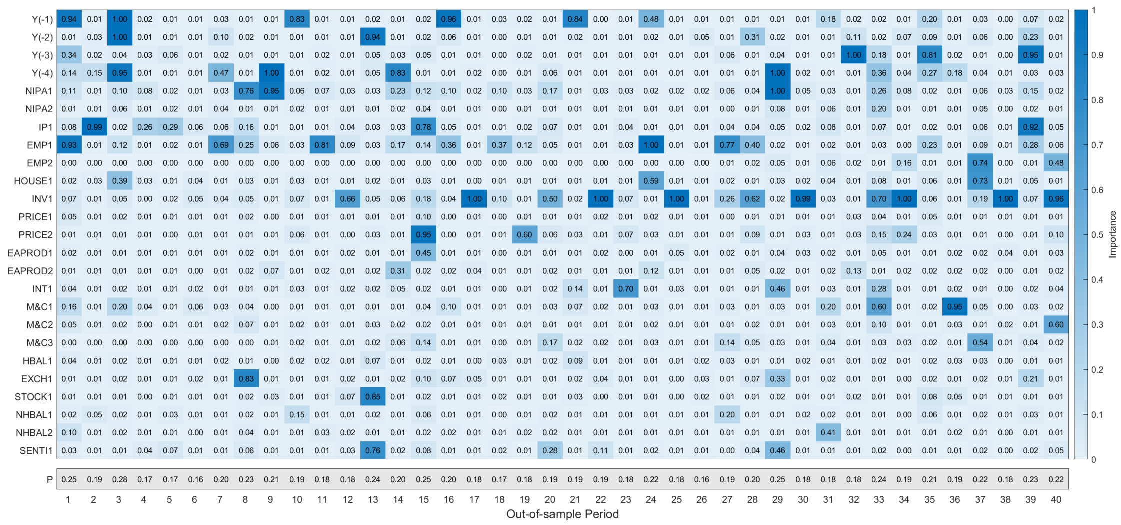
(b) B-DART Model Case

Figure L: Importance for Forecasting 5-quarter-ahead Crude Oil Price

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 5-quarter-ahead Crude Oil Price. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



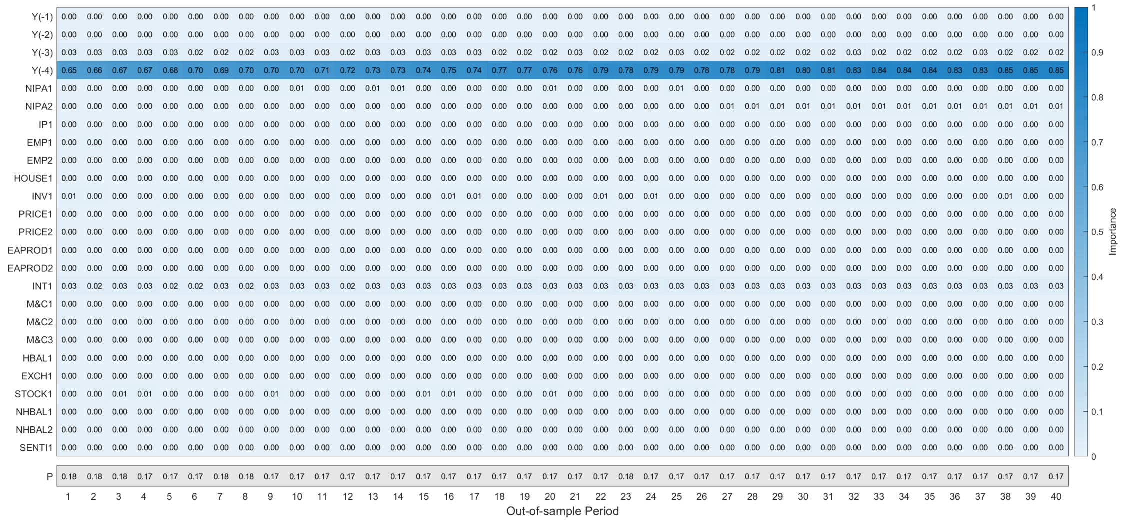
(a) Dirac Model Case



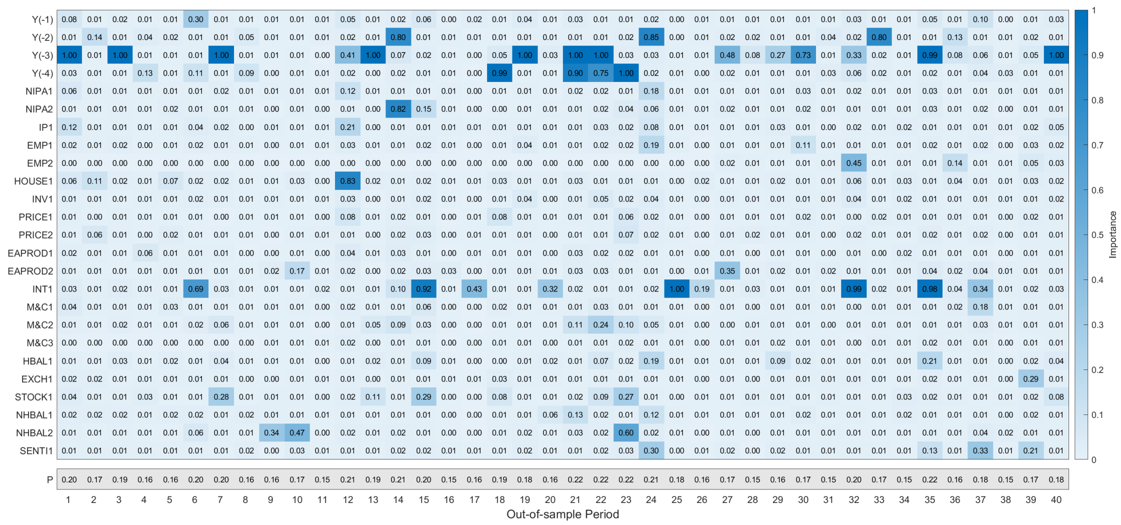
(b) B-DART Model Case

Figure M: Importance for Forecasting 2-quarter-ahead Federal Funds Effective Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 2-quarter-ahead Federal Funds Effective Rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



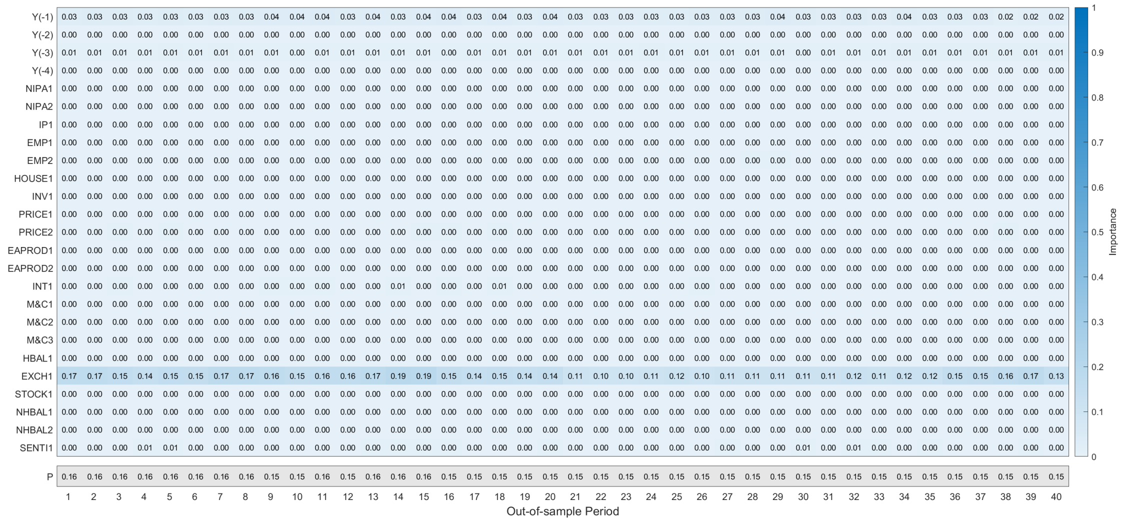
(a) Dirac Model Case



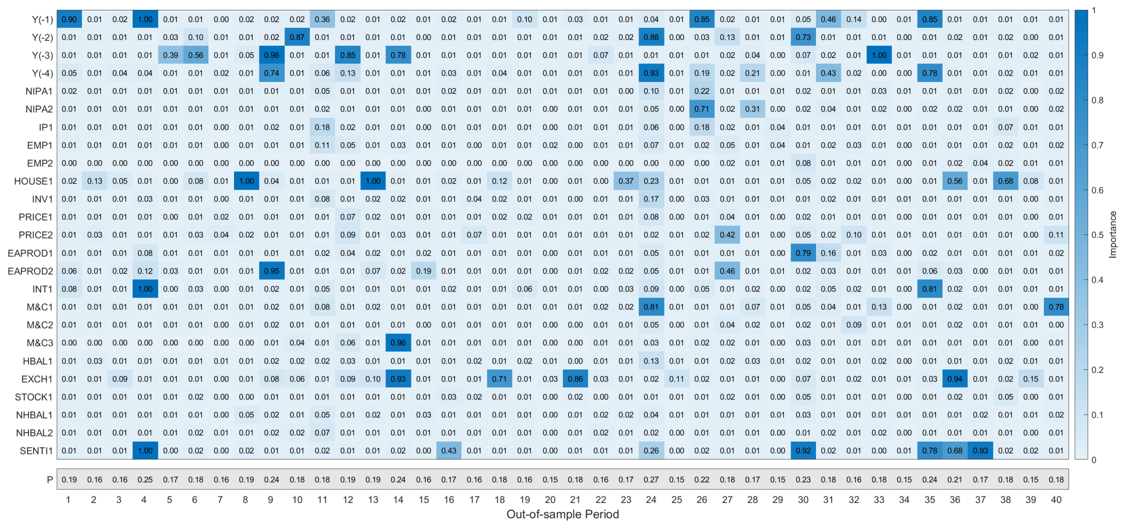
(b) B-DART Model Case

Figure N: Importance for Forecasting 8-quarter-ahead Federal Funds Effective Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 8-quarter-ahead Federal Funds Effective Rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



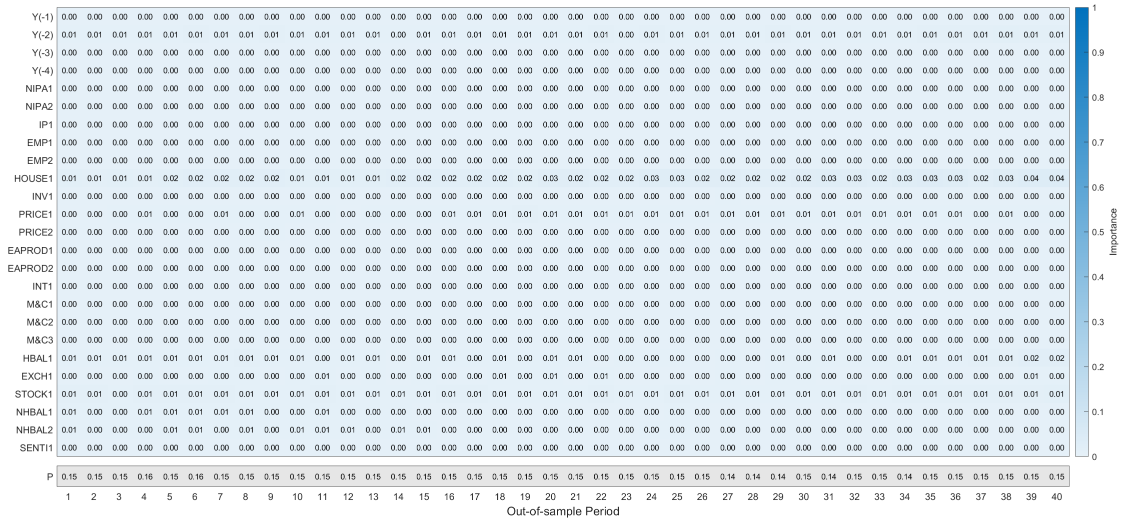
(a) **Dirac Model Case**



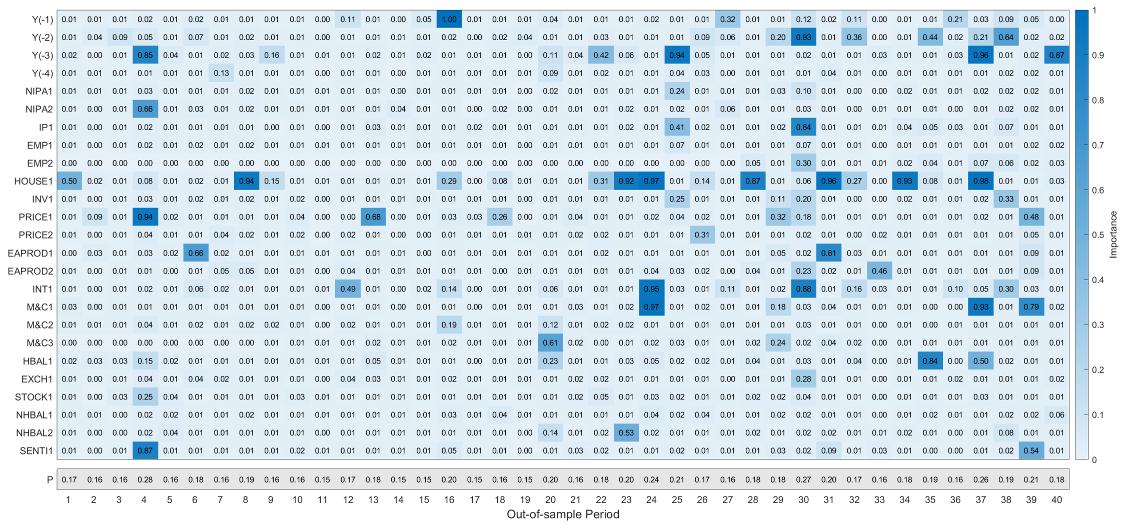
(b) **B-DART Model Case**

Figure O: Importance for Forecasting 5-quarter-ahead 10-year Maturity Interest Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 5-quarter-ahead 10-year Maturity Interest Rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



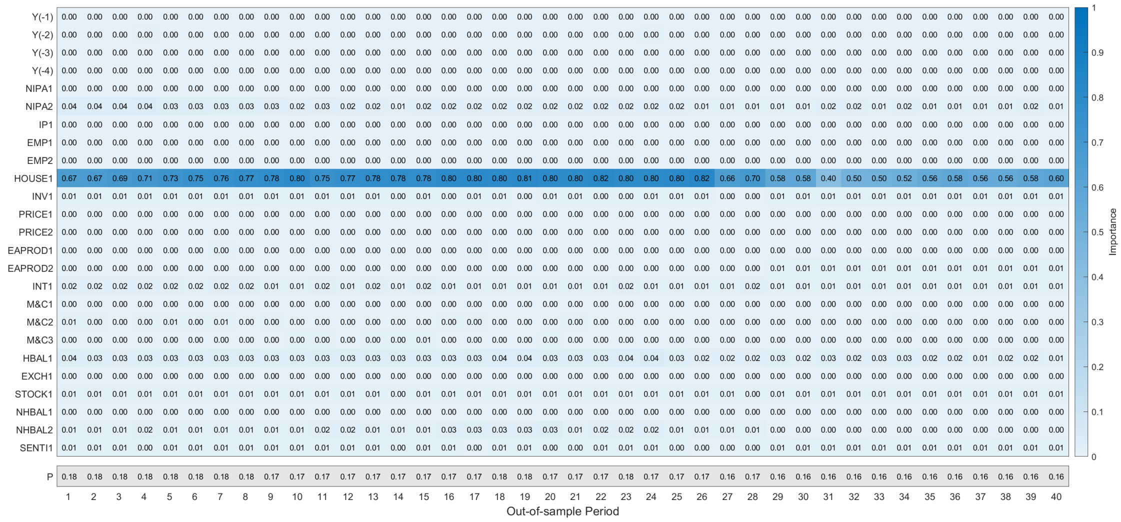
(a) **Dirac Model Case**



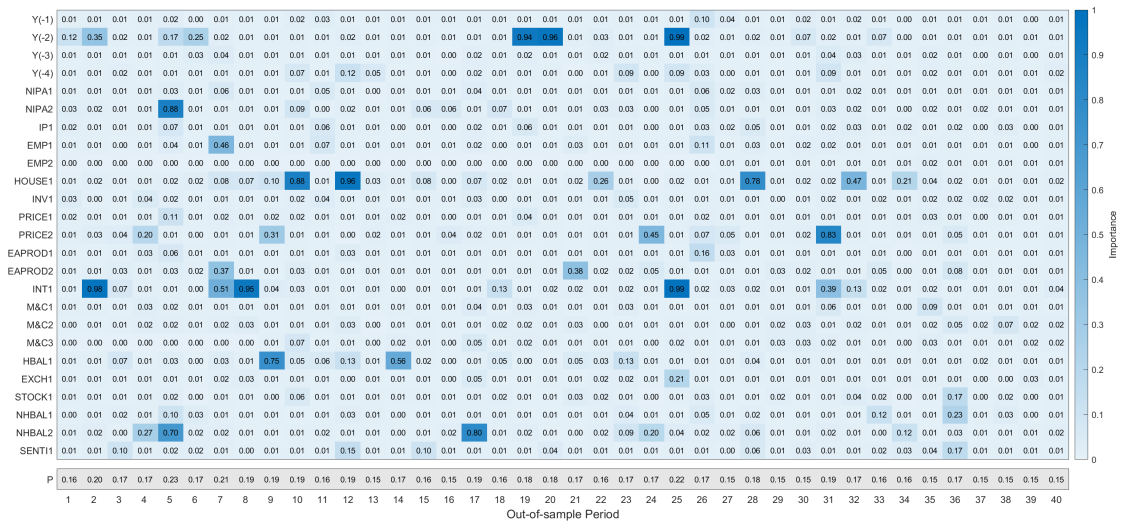
(b) **B-DART Model Case**

Figure P: Importance for Forecasting 6-quarter-ahead 10-year Maturity Interest Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 6-quarter-ahead 10-year Maturity Interest Rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



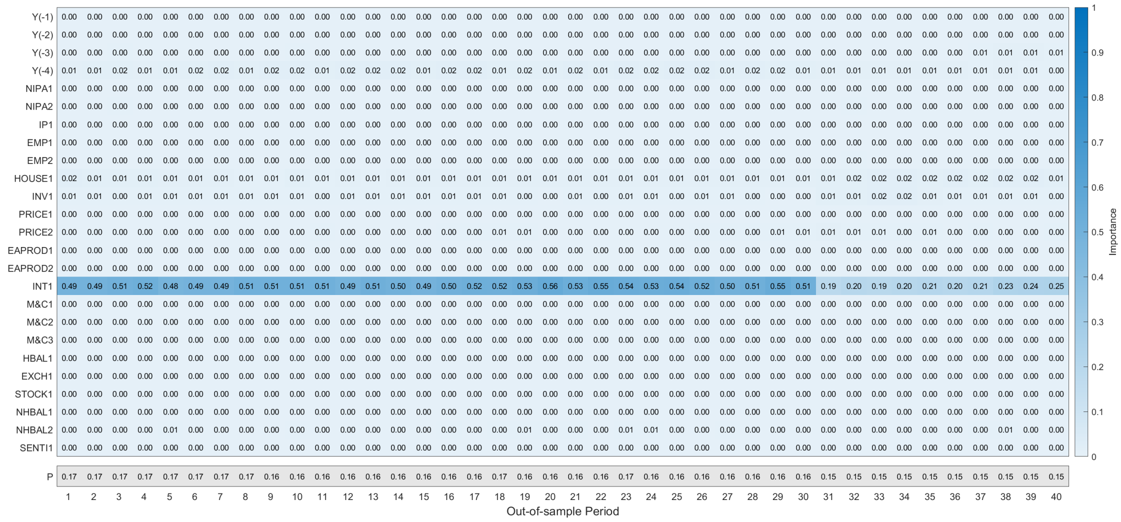
(a) **Dirac Model Case**



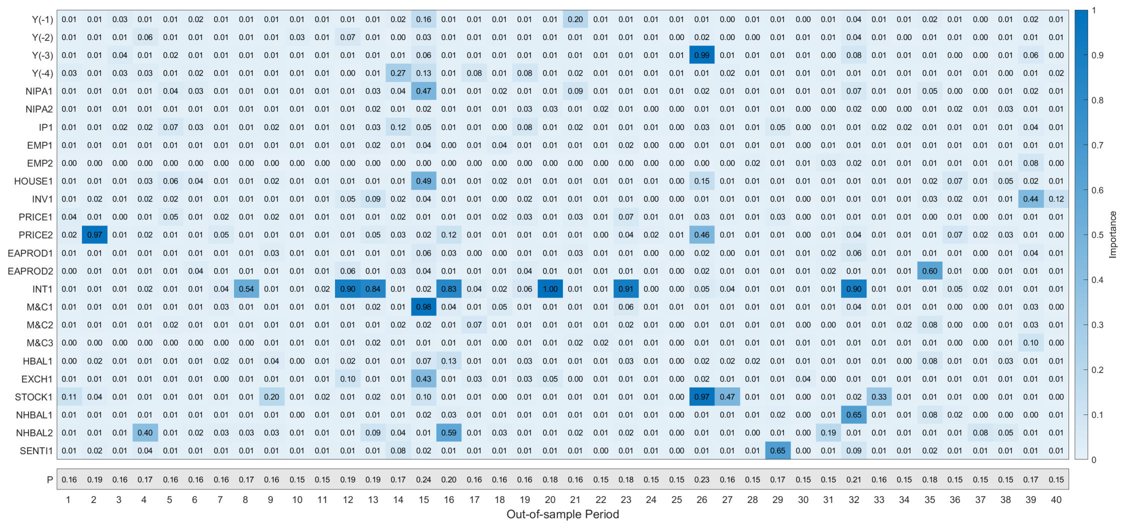
(b) **B-DART Model Case**

Figure Q: Importance for Forecasting 3-quarter-ahead Real GDP Growth Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 3-quarter-ahead Real GDP Growth Rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.



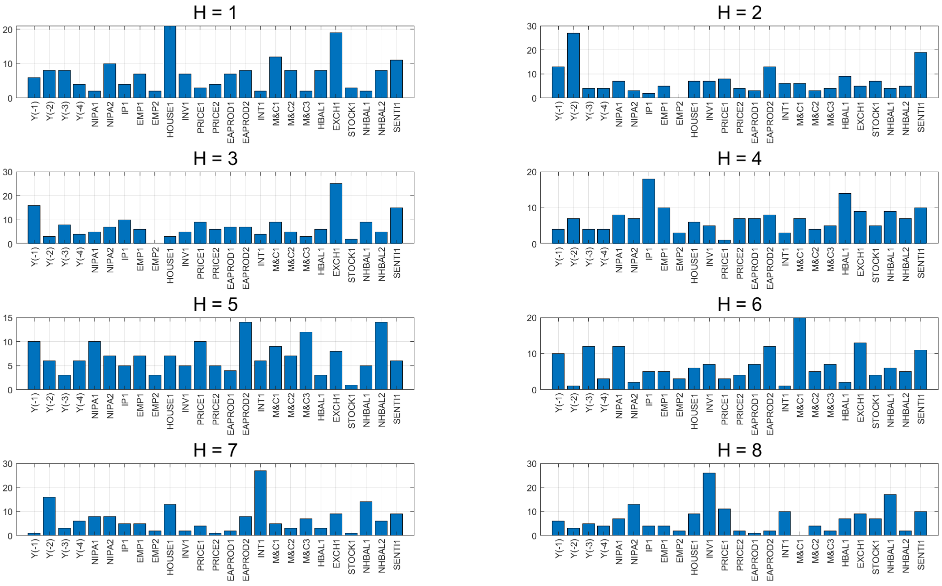
(a) **Dirac Model Case**



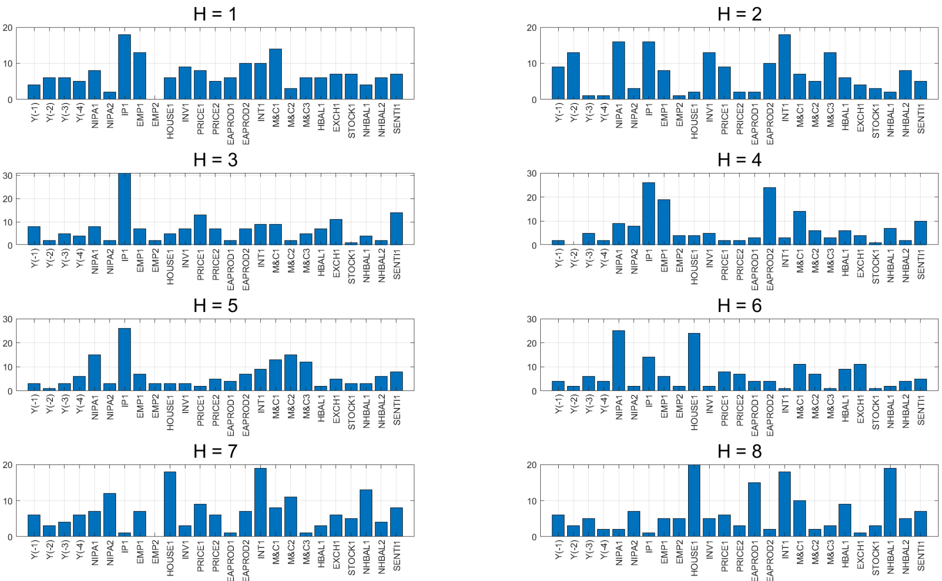
(b) **B-DART Model Case**

Figure R: Importance for Forecasting 5-quarter-ahead Real GDP Growth Rate

Note. The figure above depicts the importance of each predictor variable for 40 out-of-sample periods when forecasting the 5-quarter-ahead Real GDP Growth Rate. The heatmap becomes darker as the estimated importance of the predictor variable approaches 1. The last row is the estimated inclusion probability.

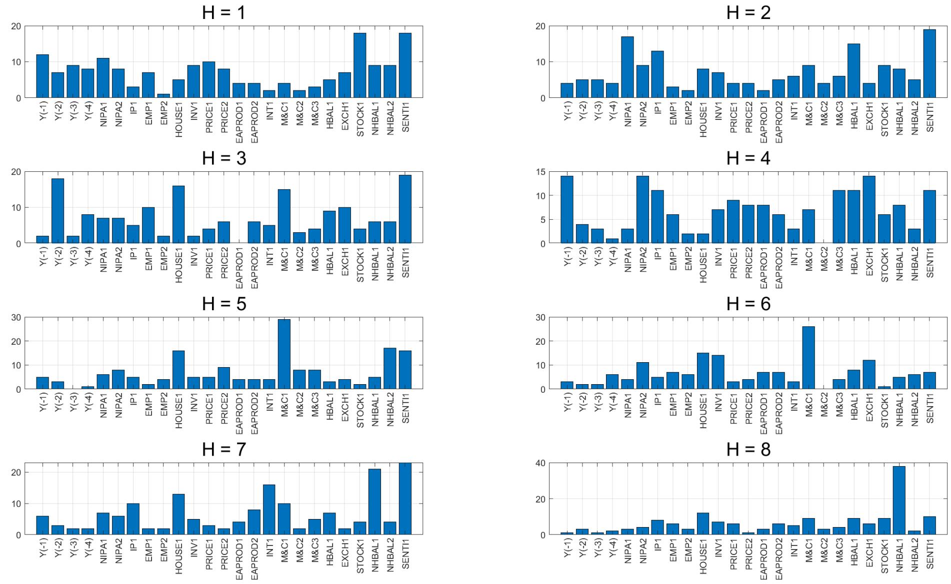


(a) **PCE Inflation Case**

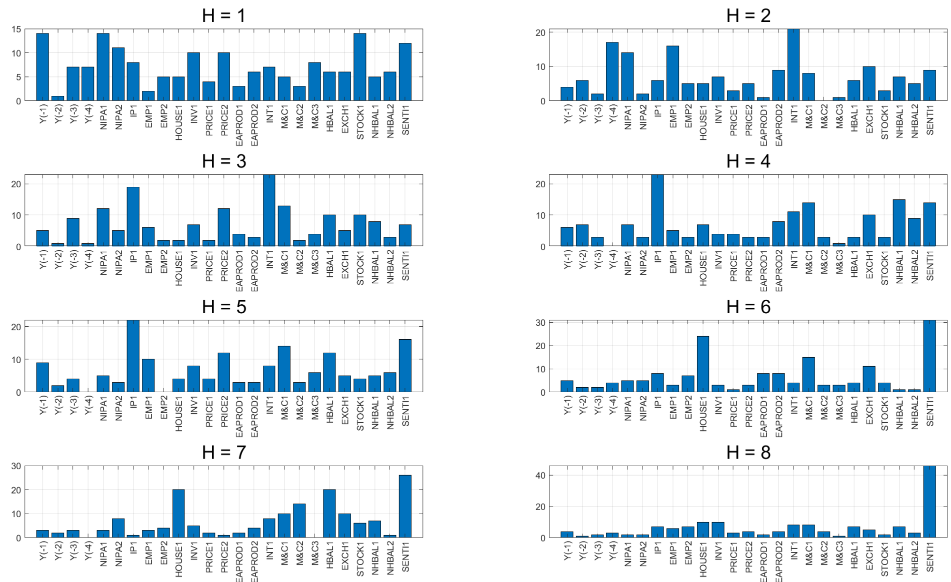


(b) **CPI Inflation Case**

Figure S: Frequency of Predictors used as the Splitting Criteria

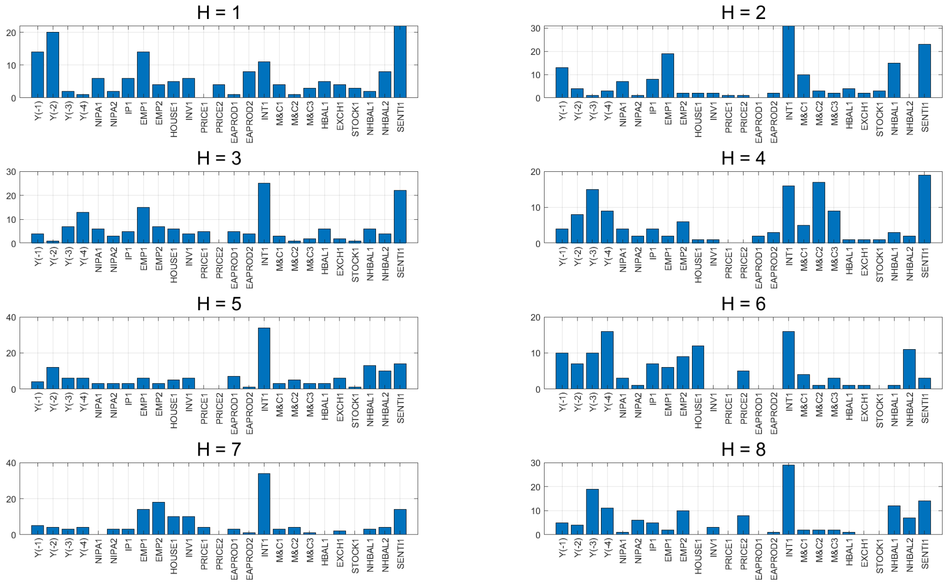


(a) PPI Inflation Case

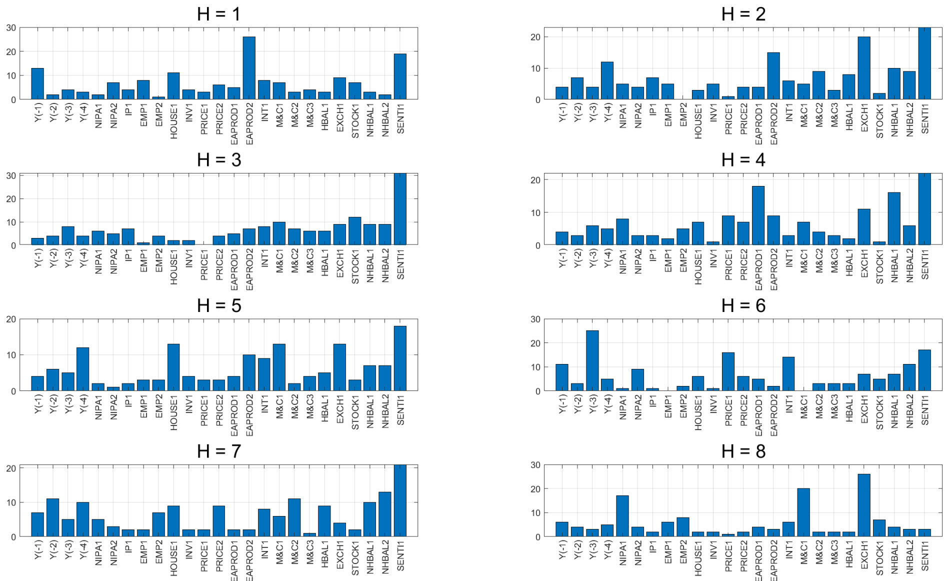


(b) Crude Oil Price Case

Figure T: Frequency of Predictors used as the Splitting Criteria

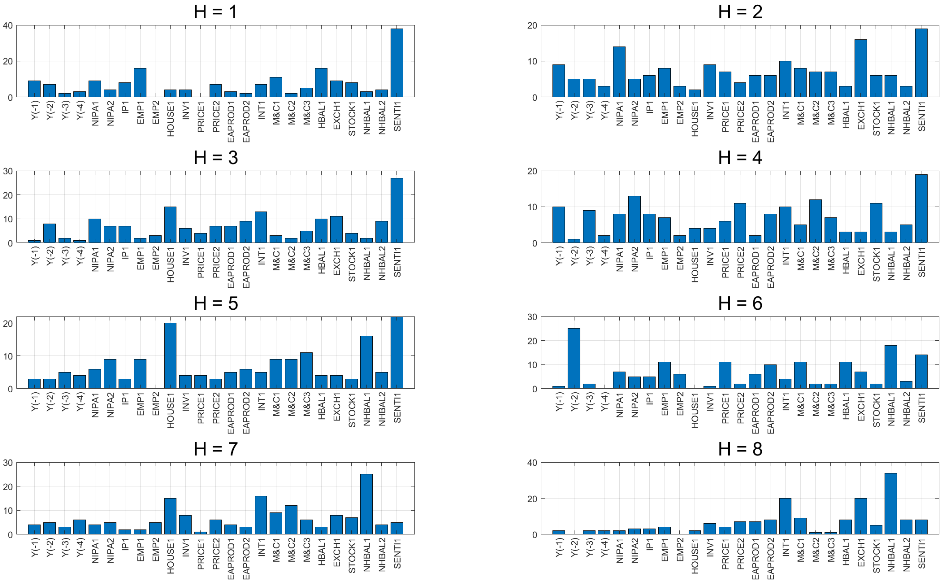


(a) Federal Funds Effective Rate Case



(b) 10-year Maturity Interest Rate Case

Figure U: Frequency of Predictors used as the Splitting Criteria



(a) Real GDP Growth Rate Case

Figure V: Frequency of Predictors used as the Splitting Criteria