

Execution Report

Title: **Forking paths in empirical studies**

Author: **Guillaume Coqueret**

Full reference: Coqueret, Guillaume "Forking paths in empirical studies", Working Paper, February 3, 2022.

The structure and contents of this execution report provided by **cascad** for the certification are similar to those recommended by the [AEA Data Editor](#).

1. DATA DESCRIPTION

This paper proposes a theoretical framework that characterizes the range of outcomes in empirical studies, and applies it using the public dataset on equity premium prediction from [Goyal et al. \(2021\)](#).

2. CODE DESCRIPTION

For the purpose of this certification, we aimed to check the results displayed in Table 2 and in Figures 4-10. Note that the Table 1 and Figures 1-3 contain no computational results and hence are not checked.

The replication materials store one single R markdown file. It imports the data stored in a [Github repository](#), cleans them, generates the models and eventually creates the Tables and Figures.

3. REPLICATION STEPS

The code was downloaded from the cascade website, and run using R 4.1.2 on a computer with 256 GB RAM, Intel® Xeon™ Silver 4210R CPU @2.40-2.39GHz, and Windows 10 OS. We encountered no issues during the replication process.

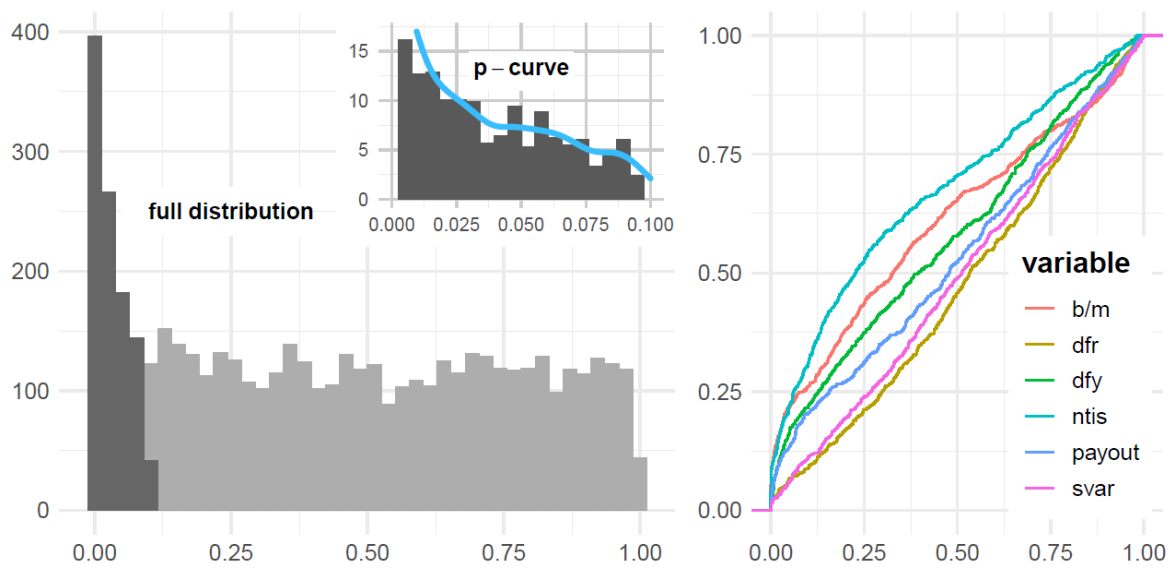
4. FINDINGS

We reproduced Table 2 and Figures 4-6, 8-10 with accuracy.

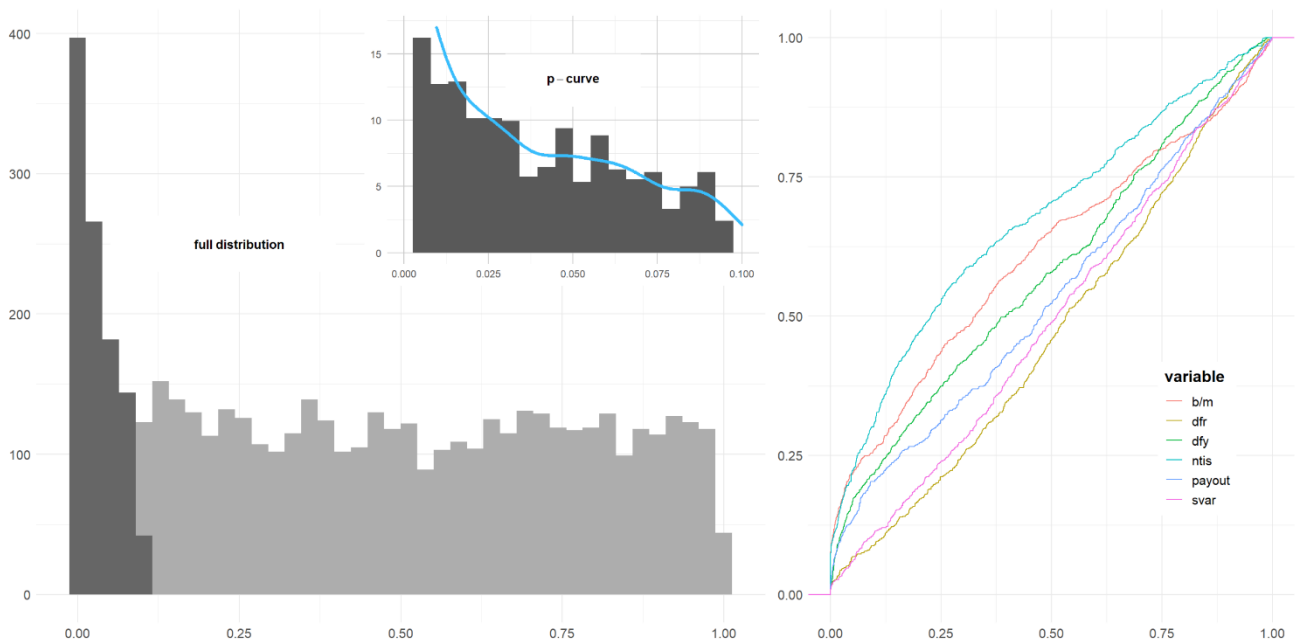
We reproduced Figure 7 with some discrepancies. After talking to the author and looking at the code, we realized that the values measured on the x-axis are random as the `geom_facet` layer in `ggplot2` generates some noise. While the two figures differ, the information they show, and the conclusions they carry, seem identical.

4.1. FIGURE 4: DISTRIBUTION OF P-VALUES.

Original:

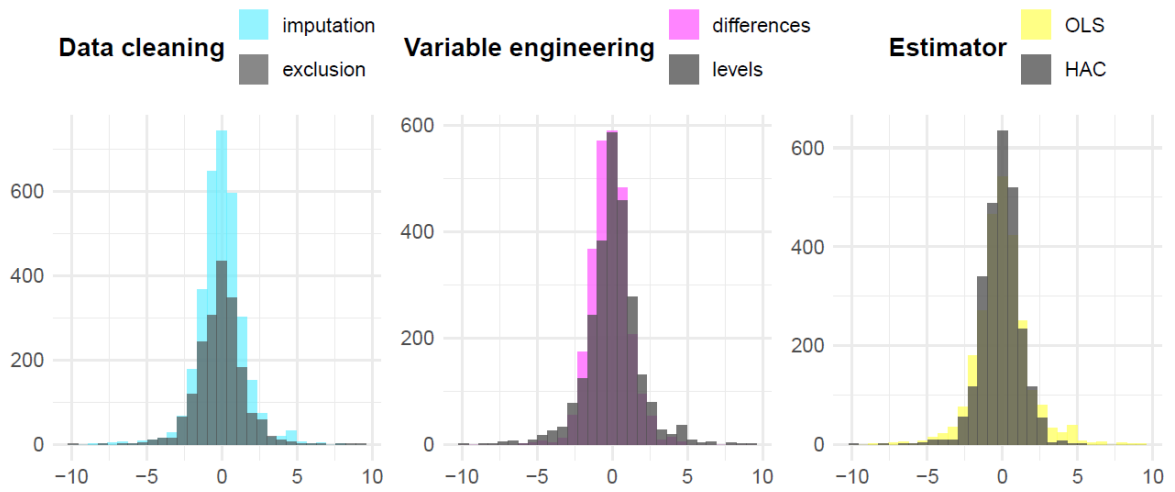


Reproduced:

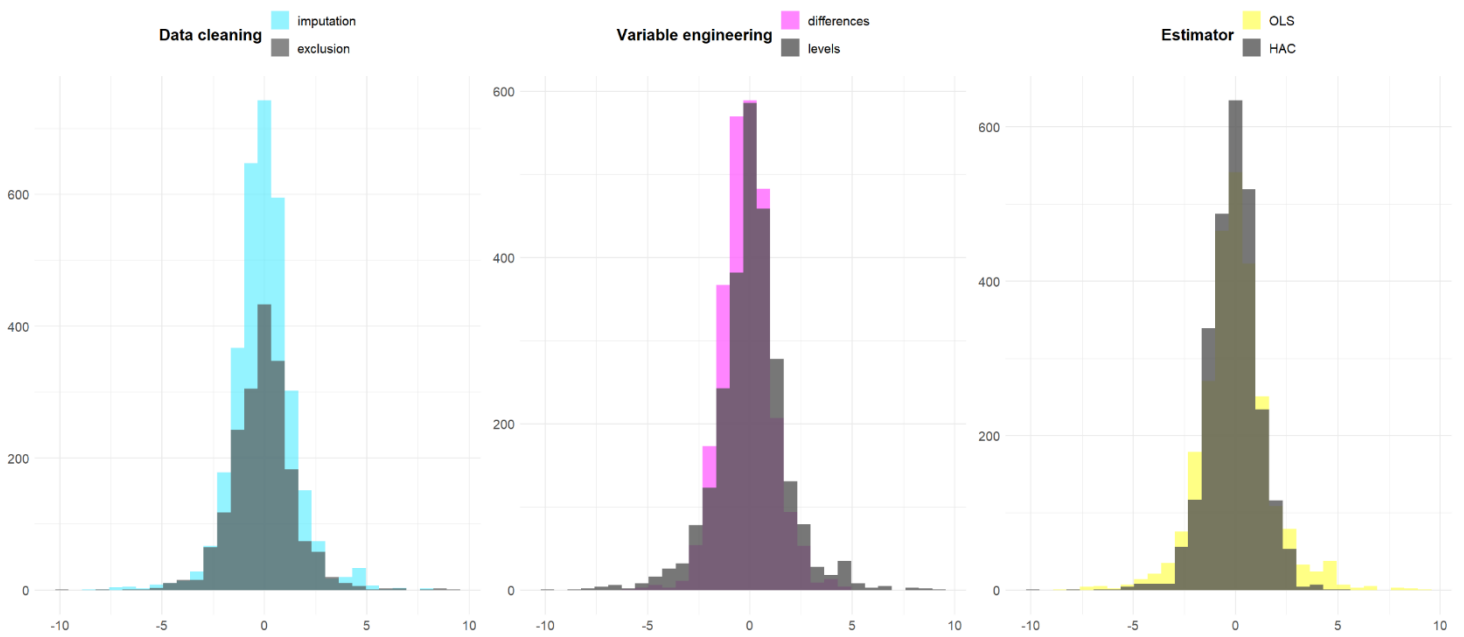


4.2. FIGURE 5: IMPACT OF MAPPINGS: ROBUSTNESS CHECKS.

Original:

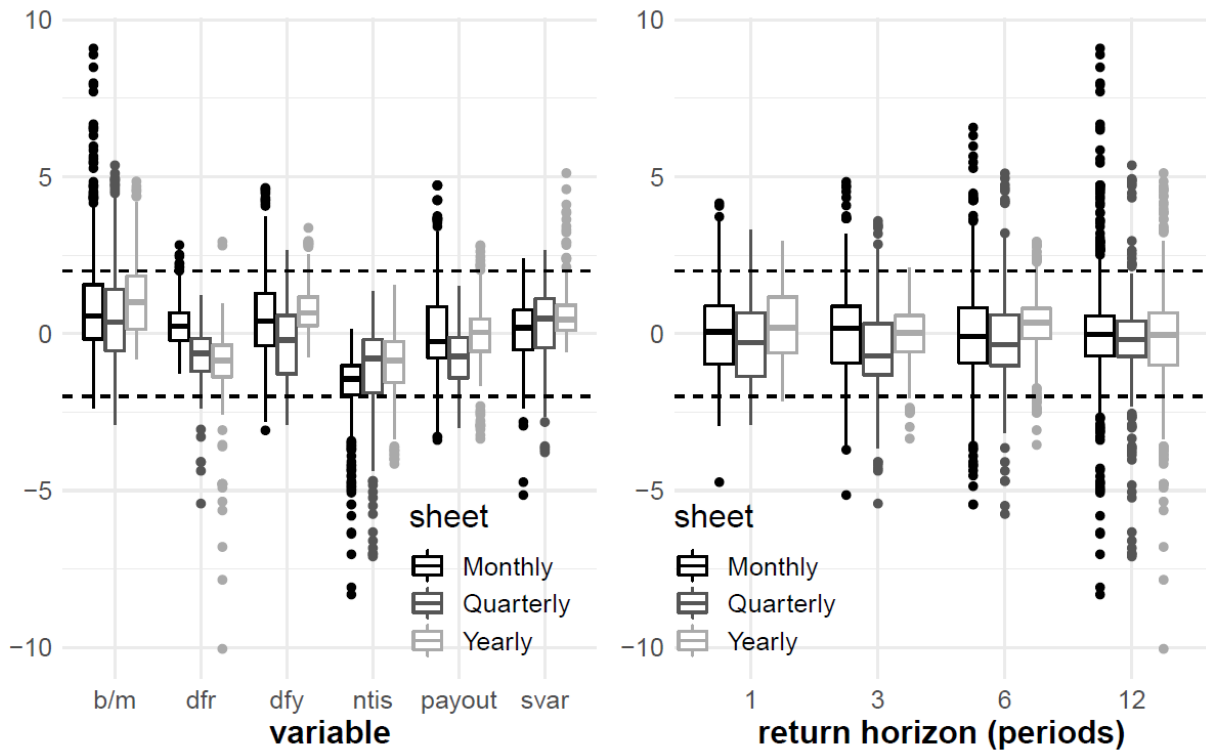


Reproduced:

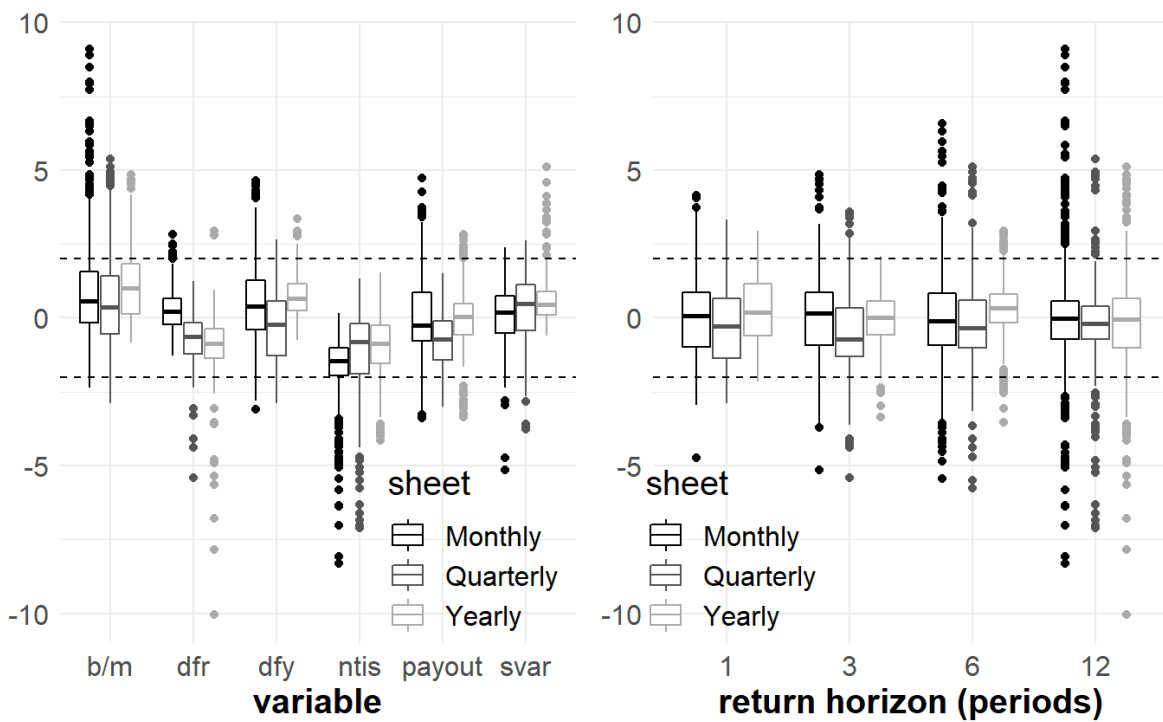


4.3. FIGURE 6: DRIVERS OF SCALAR OUTPUT: MODELLING ASSUMPTIONS.

Original:



Reproduced:



4.4. TABLE 2: TEST FOR THE VARIABLE CHOICE.

Original:

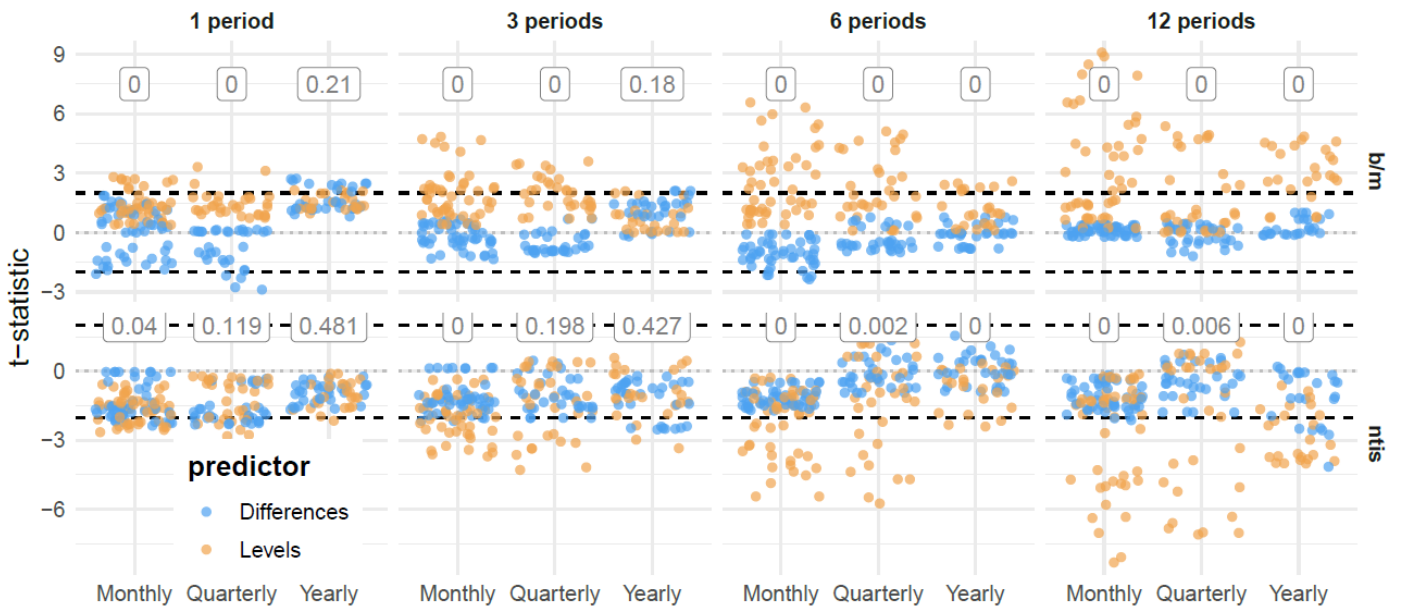
		pair-wise tests						simple test
		payout	b/m	svar	dfr	dfy	ntis	
pair-wise tests	payout		0.000	0.000	0.335	0.000	0.000	0.000
	b/m	0.000		0.000	0.000	0.000	0.000	0.000
	svar	0.003	0.000		0.000	0.281	0.000	0.000
	dfr	0.000	0.000	0.057		0.000	0.000	0.000
	dfy	0.008	0.054	0.000	0.000		0.000	0.000
	ntis	0.000	0.000	0.000	0.000	0.000		0.000
	simple test	0.011	0.000	0.000	0.000	0.043	0.000	

Reproduced:

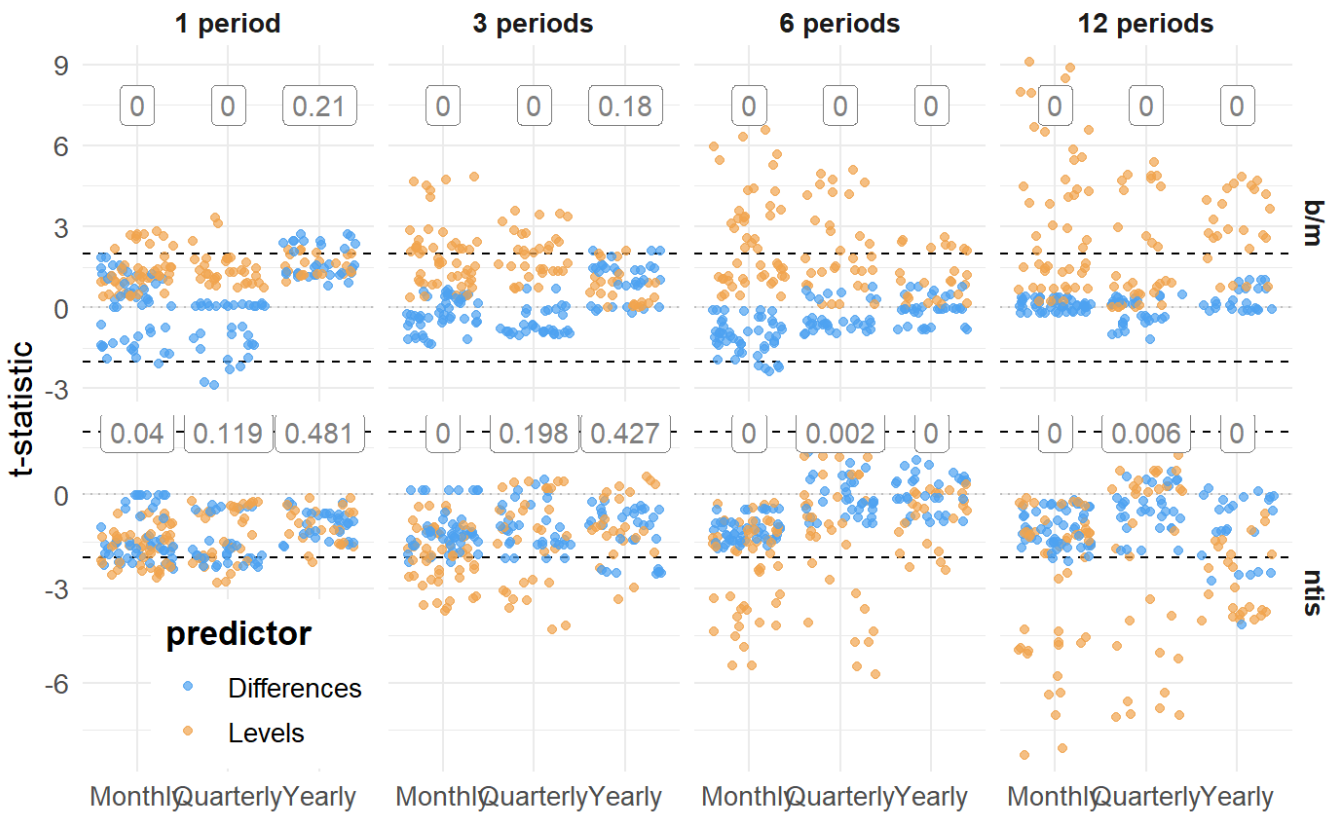
	payout	b/m	svar	dfr	dfy	ntis	t_stat_test
payout	NA	0.000	0.000	0.335	0.000	0	0
b/m	0.000	NA	0.000	0.000	0.000	0	0
svar	0.003	0.000	NA	0.000	0.281	0	0
dfr	0.000	0.000	0.057	NA	0.000	0	0
dfy	0.008	0.054	0.000	0.000	NA	0	0
ntis	0.000	0.000	0.000	0.000	0.000	NA	0
	0.011	0.000	0.000	0.000	0.043	0	NA

4.5. FIGURE 7: FOCUS ON B/M AND NTIS.

Original:

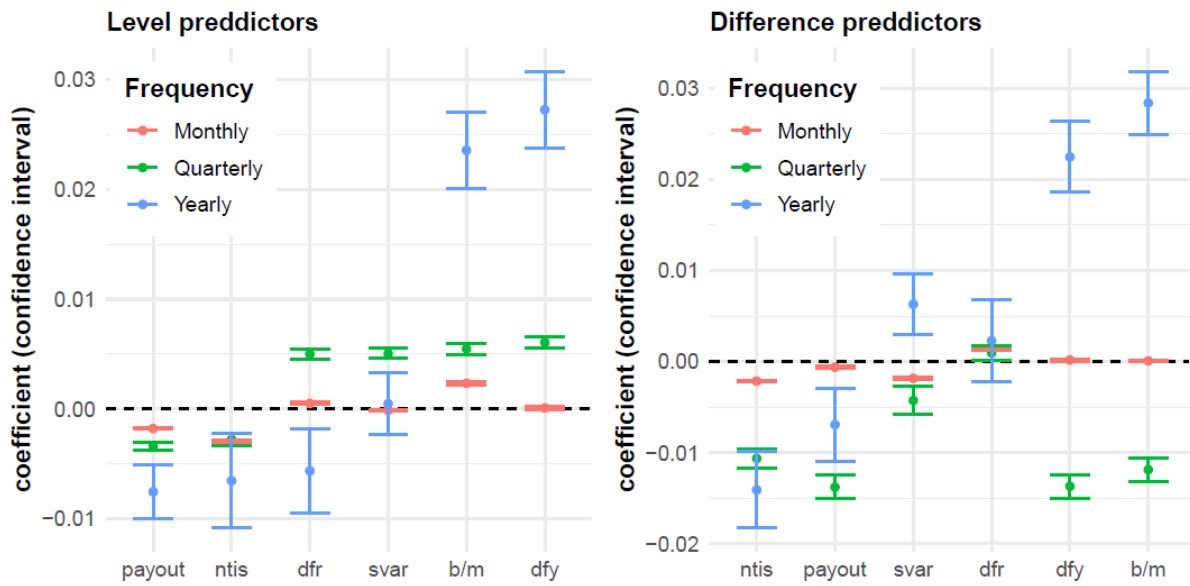


Reproduced:

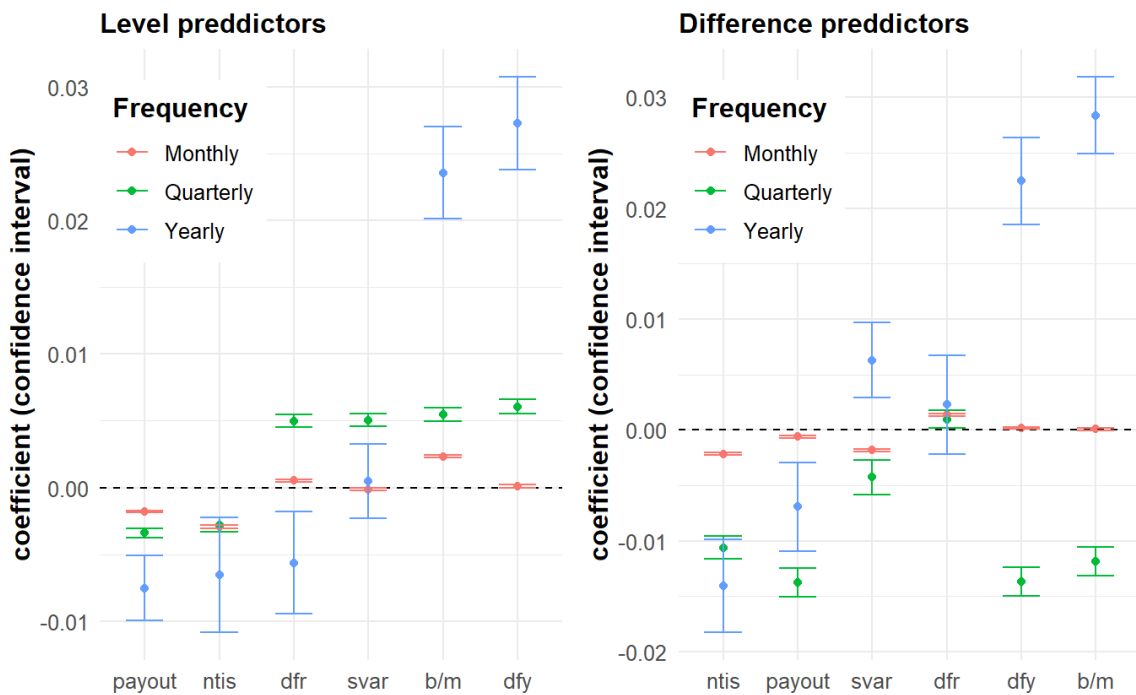


4.6. FIGURE 8: FREQUENTIST MODEL AVERAGING.

Original:

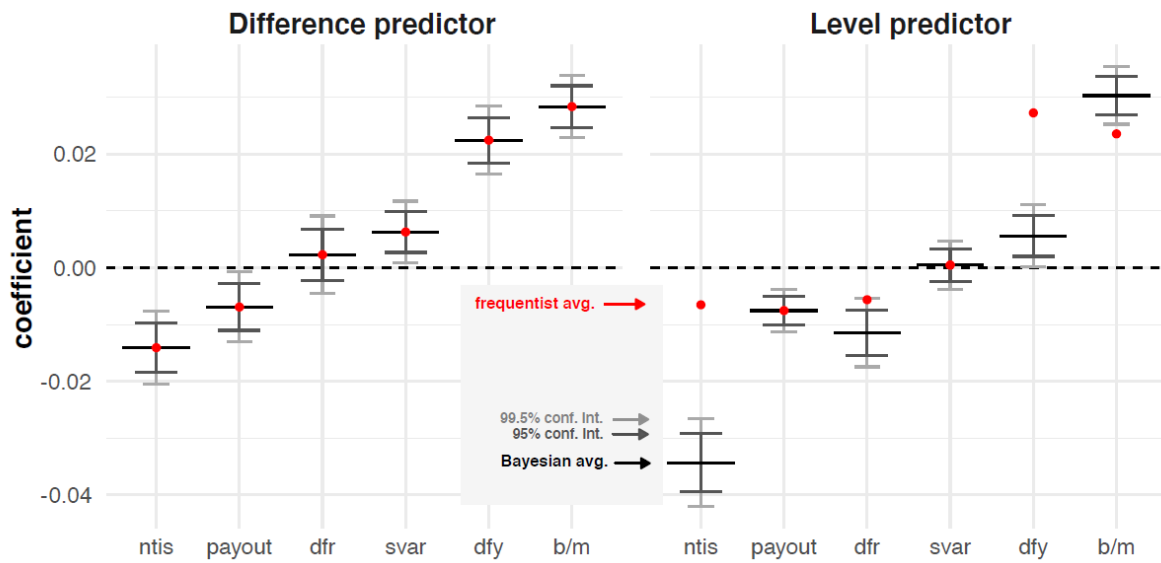


Reproduced:

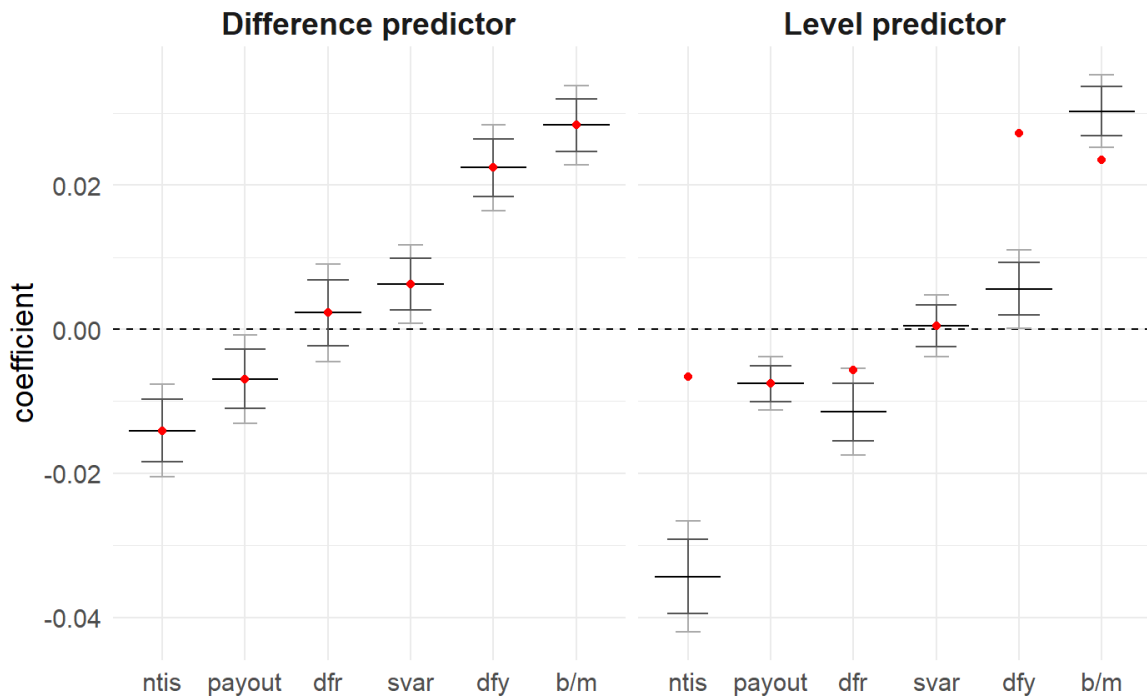


4.7. FIGURE 9: BAYESIAN MODEL AVERAGING.

Original:

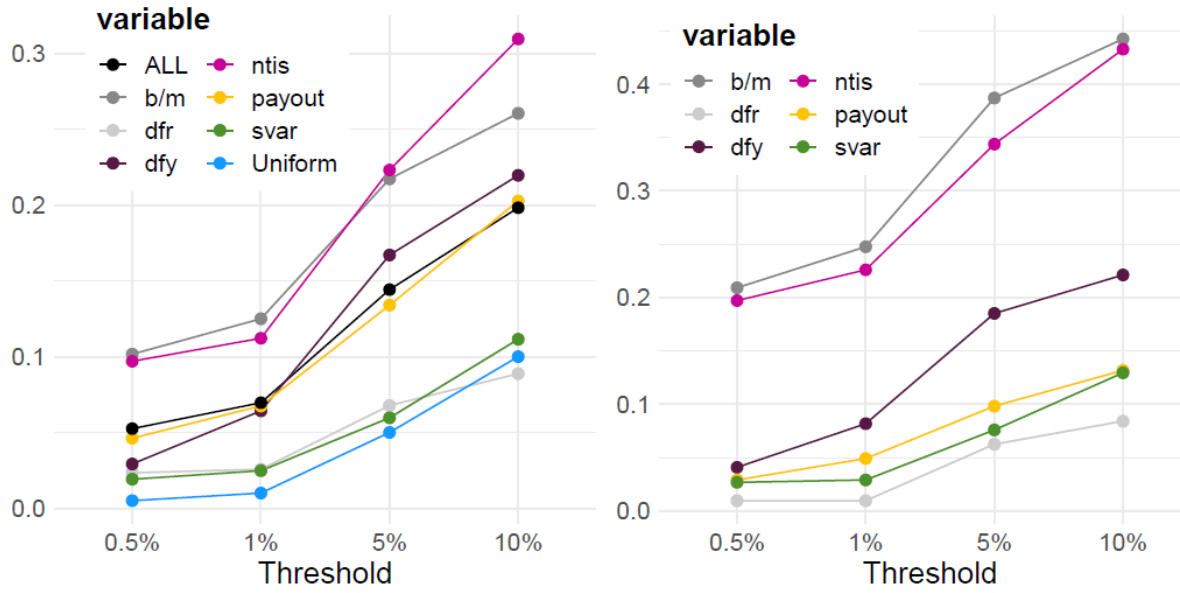


Reproduced:

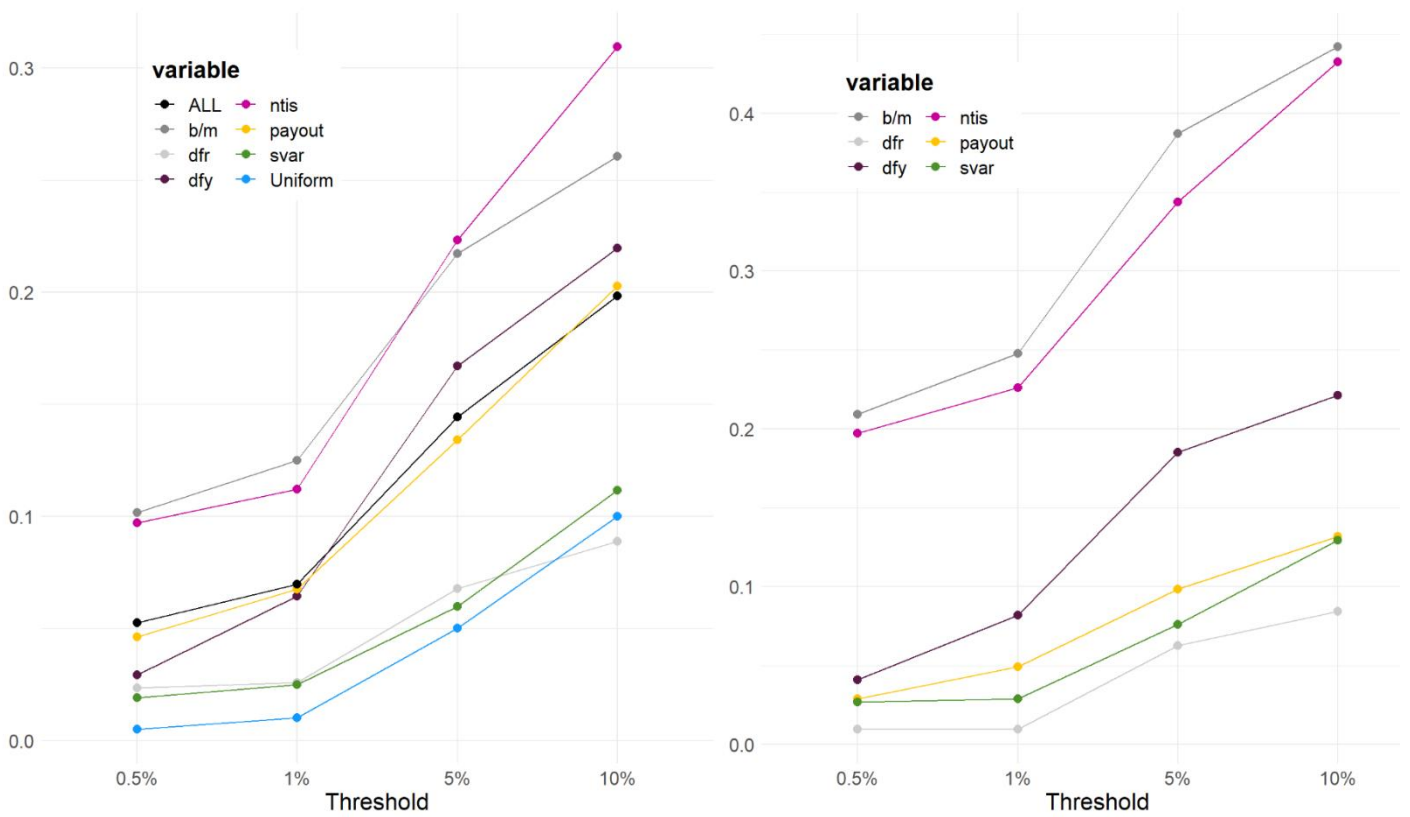


4.8. FIGURE 10: REJECTION RATES.

Original:



Reproduced:



5. REFERENCES

Goyal, A., I. Welch, and A. Zafirov (2021). A comprehensive look at the empirical performance of equity premium prediction II. *SSRN Working Paper 3929119*.