HELLENIC REPUBLIC
UNIVERSITY OF MACEDONIA

ISSN 1791-3144

University of Macedonia
Department of Economics

Discussion Paper Series

A Mixed Frequency Approach for Stock Returns and Valuation Ratios

Theologos Dergiades, Costas Milas and Theodore Panagiotidis

Discussion Paper No. 8/2019

Department of Economics, University of Macedonia, 156 Egnatia str, 540 06 Thessaloniki, Greece, Fax: + 30 (0) 2310 891292
http://www.uom.gr/index.php?newlang=eng&tmima=3&categorymenu=2

A Mixed Frequency Approach for Stock Returns and Valuation Ratios

Theologos Dergiades
Department of International & European Studies
University of Macedonia, Greece
e-mail: dergiades@uom.edu.gr

Costas Milas*
Management School
University of Liverpool, United Kingdom
e-mail: costas.milas@liverpool.ac.uk

Theodore Panagiotidis
Department of Economics
University of Macedonia, Greece
e-mail: tpanag@uom.edu.gr

October 18, 2019

Abstract

We employ a Mixed-Frequency VAR to study the effect of four valuation ratios (the price-dividend ratio, the price-earnings ratio, the Cyclically Adjusted Price Earnings Ratio and the Total Return Cyclically Adjusted Price Earnings Ratio) on the US stock market. We quantify the interaction between high and low frequency data. We show that all valuation ratios (observed at a monthly frequency) significantly affect stock market returns (observed at a daily frequency) at both long and short horizons.

***Keywords*:** Stock Index Returns; Valuation Ratios; MF-VAR; Impulse Response Analysis.

JEL: G1, C12; C13.

* Corresponding author: Costas Milas, Management School, University of Liverpool, Chatham Street, L69 7ZH, Liverpool, UK. E-mail: costas.milas@liverpool.ac.uk.

1. Introduction

Campbell and Shiller (1988) pioneered work on the influence of cash flows measures on aggregate stock market portfolio returns according to which increases in the price-dividend ratio and the price-earnings ratio should decrease future returns. Their work was followed by a number of studies including Lettau and Ludvigson (2001, 2005), Rapach and Wohar (2006), Ang and Bekaert (2006), Bollerslev *et al.* (2015), Choi *et al.* (2017) and Jagannathan and Liu (2019). The main body of the literature suggests that stock market index returns are significantly affected by cash flow measures at long horizons. However, significant impact at shorter horizons (less than a year) also emerges once empirical specifications allow for either the presence of market risk (such as the panel defined tail risk measure of Kelly and Jiang, 2014) or investors' learning about existing market risks (Jagannathan and Liu, 2019).

Empirical disagreement on the timing of stock market returns predictability relates to the adopted methodological framework, the set of predictors or even the choice of data frequency. Most of the literature uses annual, quarterly or monthly data and the analysis is carried out at the lower frequency after aggregating higher frequency observations. As a result, valuable information is smoothed out therefore failing to take full advantage of the underlying data dynamics (Ghysels, 2016).

This paper focuses on the short-run and long-run impact of different cash flow measures for stock market index returns. We deviate from previous work by considering the mixed frequency nature of the involved variables. We focus on four US cash flow measures (the price-dividend ratio, the price-earnings ratio, the Cyclically Adjusted Price Earnings Ratio (CAPE) and the Total Return Cyclically Adjusted Price Earnings Ratio (TRCAPE); all observed at a monthly frequency) and assess their impact on the S&P 500 index returns (observed at a daily frequency). We follow the Ghysels (2016) Mixed Frequency Vector Autoregression (hereafter MF-VAR) impulse response analysis framework to evaluate the timing at which the above valuation ratios impact on aggregate market returns. We differentiate between the short and long-run dynamics of the responses among the trading days belonging to the low frequency time index.

The paper proceeds as follows. Section 2 outlines the methodology and Section 3 discusses the data. Section 4 reports the findings and Section 5 concludes.

2. Methodology

Ghysels (2016) and Ghysels *et al.* (2016) draw inference from a set of random processes with mixed observational frequency. Within this context, both high $\{r_{HF}(\tau_{LF}, j)_{j=1}^{m(\tau)}\}_{\tau_{LF}}$ and low $\{d_{LF}(\tau_{LF})\}_{\tau_{LF}}$ frequency (hereafter *HF* and *LF*) processes are concatenated into the $\underline{\mathrm{x}}(\tau_{LF})$ vector by stacking the *HF* process at the *LF* time index $\tau_{LF} \in \{0...T\}$. For the $\underline{\mathrm{x}}(\tau)$ vector, the K-dimensional VAR process of order p is written as:

$$\underline{\mathrm{x}}(\tau_{LF}) = \mathbf{A}_0 + \sum_{k=1}^{p} \mathbf{A}_k \underline{\mathrm{x}}(\tau_{LF} - k) + \underline{\varepsilon}(\tau_{LF}) \tag{1}$$

where, $K = K_{LF} + mK_{HF}$, K_{LF} and K_{HF} is the number of the *LF* and *HF* variables, respectively, and m is the number of observations for the *HF* variable that belong to the *LF* time index. $\mathbf{A}_k$ are the $(K \times K)$ coefficient matrices and $\mathbf{A}_0$ and $\underline{\varepsilon}(\tau_{LF})$ are the $(K \times 1)$ vectors of intercepts and errors, respectively.

By introducing the *LF* lag operator $\mathcal{L}_L$, that is $\mathcal{L}_L \underline{\mathrm{x}}(\tau_{LF}) = \underline{\mathrm{x}}(\tau_{LF} - 1)$, *Eq.* (1) is re-written more compactly as:

$$\mathrm{A}(\mathcal{L}_L)(\underline{\mathrm{x}}(\tau_{LF}) - \mu_{\underline{\mathrm{x}}}) = \underline{\varepsilon}(\tau_{LF}) \qquad \text{with} \qquad \begin{aligned} \mathrm{A}(\mathcal{L}_L) &= I - \sum_{k=1}^{p} \mathbf{A}_k \mathcal{L}_L^k \\ \mu_{\underline{\mathrm{x}}} &= \left(I - \sum_{k-1}^{p} \mathbf{A}_k \right)^{-1} \mathbf{A}_0 \end{aligned} \tag{2}$$

Eq. (2) is estimated by OLS at the *LF* time index and the respective impulse responses are calculated through the usual iterative approach based on the Cholesky identification scheme given by *Eq.* (3) below:

$$\begin{aligned} (\underline{\mathrm{x}}(\tau_{LF}) - \mu_{\underline{\mathrm{x}}}) &= \left(I - \sum_{k=1}^{p} \mathbf{A}_k^h \mathcal{L}_L^k \right)^{-1} \underline{\varepsilon}(\tau_{LF}) \\ &= \sum_{k=1}^{\infty} F_k \varepsilon(\tau_{LF} - k) \equiv F \mathcal{L}_L^k \underline{\varepsilon}(\tau_{LF}) \end{aligned} \qquad \text{with} \qquad I = (\mathrm{A}(\mathcal{L}_L)) F(\mathcal{L}_L) \tag{3}$$

Although the Cholesky identification scheme is frequently questionable within the standard VAR framework, this is not the case for MF-VAR models. Given that for every τ_{LF} the *HF* observations lead the corresponding *LF* values in terms of publication, the variables are ordered according to their release time. Therefore, the *HF* variable precedes the *LF* one. Hence, the identification of

the underlying shocks based on the Cholesky scheme is a justified choice (Ghysels, 2016). Finally, the MF-VAR specification describes better the system dynamics as it allows to observe a different response for every stacked *HF* observation that belongs to the *LF* time index.

3. Data

We use daily S&P 500 closing prices and monthly data for (a) U.S. dividends, (b) U.S. earnings, (c) the Cyclically Adjusted (over 10 years of earnings) Price Earnings Ratio, and (d) the Total Return Cyclically Adjusted Price Earnings Ratio over the 1950:1-2019:7 period.[1] Let $\mathbb{S}_{HF}(\tau, j)_{j=1}^{m(\tau)}$, $\mathbb{D}_{LF}(\tau)$ and $\mathbb{E}_{LF}(\tau)$ to denote the S&P 500, dividend and earnings series, respectively. We compute the annualized S&P 500 returns $r_{HF}(\tau, j)_{j=1}^{m(\tau)}$ as:

$$r_{HF}(\tau,j)_{j=1}^{m(\tau)} = ln\left(\frac{\mathbb{S}_{HF}(\tau,j)_{j=1}^{m(\tau)} + \sum_{i=1}^{12}\mathbb{D}_{LF}(\tau\text{-}i)}{\mathbb{S}_{HF}(\tau,\text{-}12\cdot\bar{m}(\tau))}\right) \tag{4}$$

where $\sum_{i=1}^{12}\mathbb{D}_{LF}(\tau\text{-}i)$ refers to the past twelve months sum of $\mathbb{D}_{LF}(\tau)$ and $\mathbb{S}_{HF}(\tau,\text{-}12\cdot\bar{m}(\tau))$ is the lag of the *HF* variable $12\cdot\bar{m}(\tau)$ trading days back, with $\bar{m}(\tau)$ being the average number of the *HF* observations for all sample months (21 in our case).

The price-dividend ratio, $d_{LF}(\tau)$, and the price-earnings ratio, $e_{LF}(\tau)$, are computed as:

$$d_{LF}(\tau) = ln\left(\frac{\sum_{j=1}^{m(\tau)}\mathbb{S}_{HF}(\tau,j)/m(\tau)}{\sum_{i=1}^{12}\mathbb{D}_{LF}(\tau-i)}\right) \quad \text{and} \quad e_{LF}(\tau) = ln\left(\frac{\sum_{j=1}^{m(\tau)}\mathbb{S}_{HF}(\tau,j)/m(\tau)}{\sum_{i=1}^{12}\mathbb{E}_{LF}(\tau-i)}\right) \tag{5}$$

where $\sum_{i=1}^{12}\mathbb{E}_{LF}(\tau-i)$ is the past twelve months sum of $\mathbb{E}_{LF}(\tau)$ and $\sum_{j=1}^{m(\tau)}\mathbb{S}_{HF}(\tau,j)/m(\tau)$ is the aggregation of the *HF* variable at the *LF* time index. Furthermore, as the number of business days $m(\tau)$ differs in each month of the sample τ, the *HF* observations are balanced according to the following modification:

[1] Changes in corporate payout policy (i.e. in the form of share repurchases rather than dividends; see e.g. Jivraj and Shiller, 2017) may affect the level of the CAPE ratio. The TRCAPE variable corrects for this bias through reinvesting dividends into the price index and scaling the earnings per share. All data are expressed in real terms and come from R.J. Shiller's website (available at: http://www.econ.yale.edu/~shiller/data.htm), except the daily S&P 500 closing prices which were retrieved from Yahoo Finance. The price-earnings variable ends in 2019:4 and the TRCAPE one in 2018:9 due to data availability.

$$r^*_{HF}(\tau, j) = \begin{cases} r_{HF}(\tau, j) & \text{for} \quad j = 1, \ldots, 16 \\ \dfrac{1}{m(\tau) - 16} \sum_{k=17}^{m(\tau)} r_{HF}\left(\tau_{LF}, k\right) & \text{for} \quad j = 17 \end{cases} \tag{6}$$

The annualized S&P 500 index returns (*Eq.* 4), the price-dividend ratio (*Eq.* 5), the price-earnings ratio (*Eq.* 5) and the natural logarithm of CAPE and TRCAPE ratios, all are presented jointly in the double-scaled **Figure 1**. The *HF* variable of our dataset (S&P 500 index returns), is presented as an embedded graph in **Figure 1**.

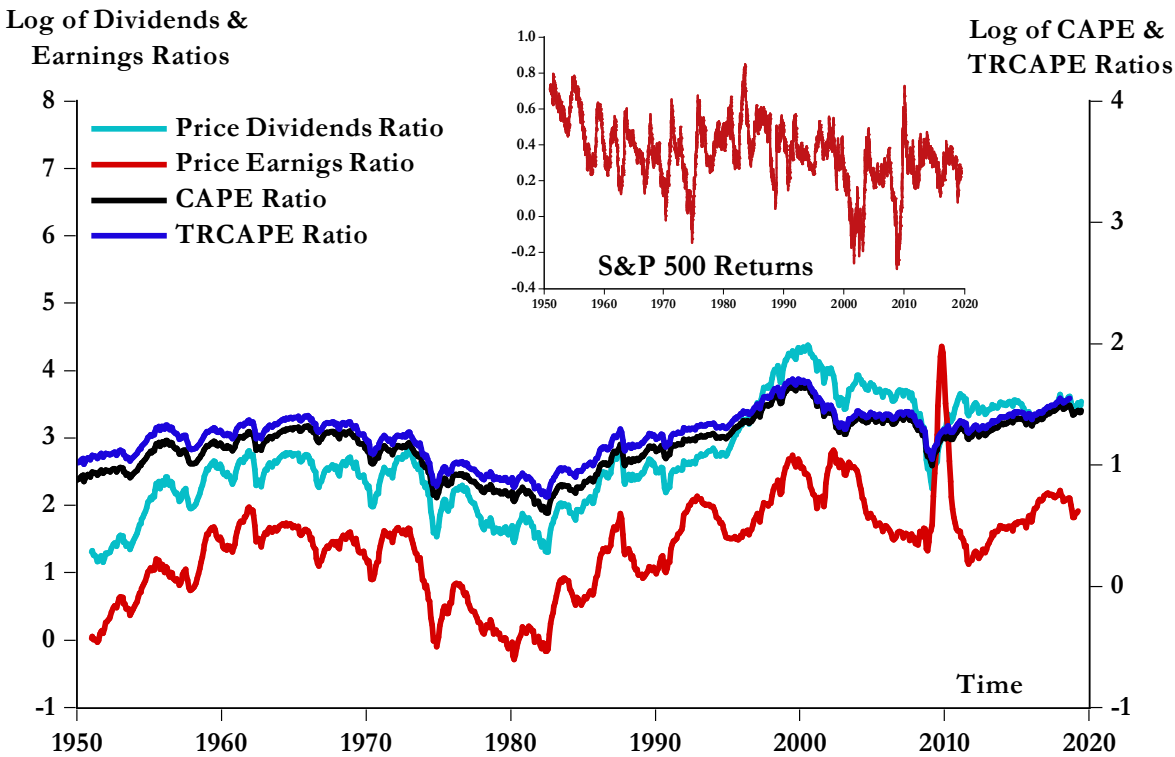


Figure 1. Valuation ratios and S&P 500 index returns

4. Impulse Response Analysis

We run bivariate time-stamped MF-VAR models and report the impact on the S&P 500 returns per trading (or business) day of a one standard deviation shock to each of the four monthly valuation ratios.[2] The results are reported in a three-dimensional surface plot. The middle surface is shaped by the impulse responses per business day and is colored according to the magnitude of the response,[3] whereas the upper and lower surfaces indicate the upper and lower bootstrapped

[2] Effective estimation is from 1951:2 onwards. For all models, an optimal lag length of two months is chosen by the Akaike Information Criterion. All variables are demeaned since the estimated specifications do not include the constant term. All estimations were done in Matlab.

[3] See the colorbar at the upper right-hand side of **Figure 2**.

95% confidence bands (based on 999 replications), respectively. The vertical left-axis depicts the magnitude of the response; the bottom central horizontal axis displays the impulse horizon in months, and the bottom horizontal right-axis portrays the business day that each derived impulse response refers to.

The responses of the stacked S&P 500 returns to a shock in the price-dividend ratio are reported in **Figure 2**. In total, the sample includes 822 months with 17 business days (see *Eq.* 6), implying 13,974 observations for $r^*_{HF}(\tau, j)$. For all business days within the low frequency time index, the paths of the responses are predominantly positive and statistically insignificant over the first five horizons (months). From the sixth horizon onwards, the responses of returns turn negative and persistent. The region of all likely trajectory paths, defined by the bootstrapped confidence surfaces, suggests that the responses are statistically different from the zero threshold from the tenth horizon onwards (indicated by the red impulse horizon of the central horizontal axis). In addition, while for various business days the short-run dynamics of the responses differentiate, at longer horizons the responses comove in a very similar manner.

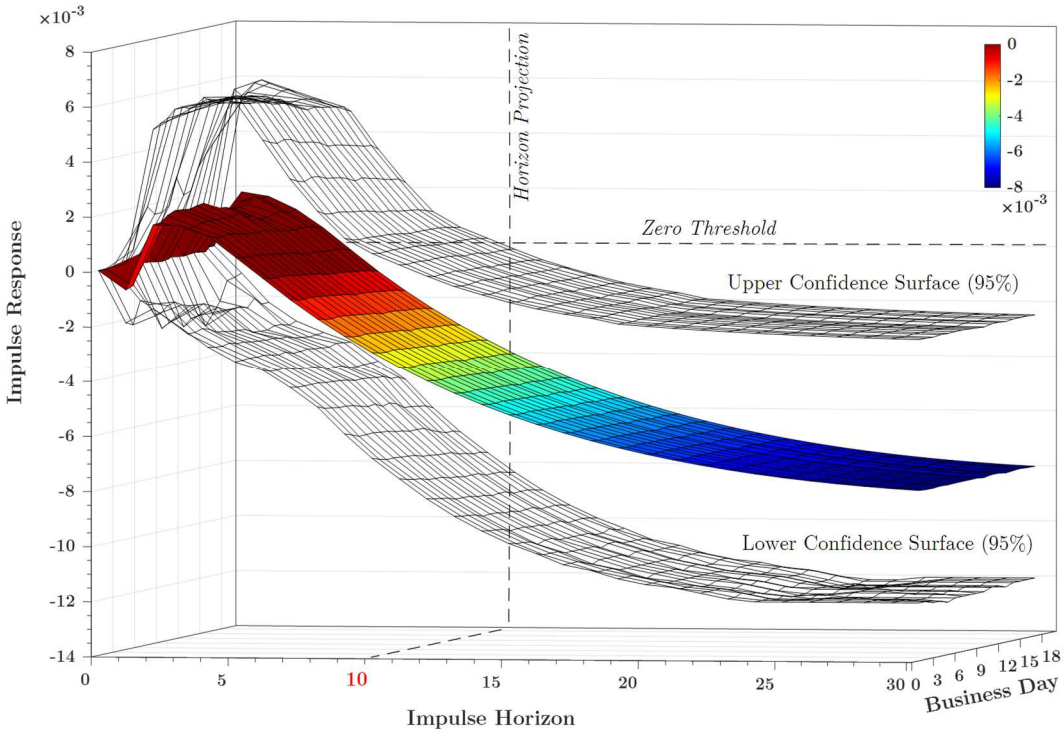


Figure 2. Responses of returns to a one standard deviation price-dividend shock (1951:2-2019:7)

The responses of the stacked S&P 500 returns to a shock in the price-earnings ratio are reported in **Figure 3**. The shape of the response surface for all business days, is qualitatively similar

to the price-dividend case. Indeed, the responses are mainly negative; the short-run dynamics imply weak and insignificant variation per business day, whereas this variation disappears at longer horizons. The main exception is the horizon at which the responses become statistically significant; these are different from zero at horizon 20 and beyond.

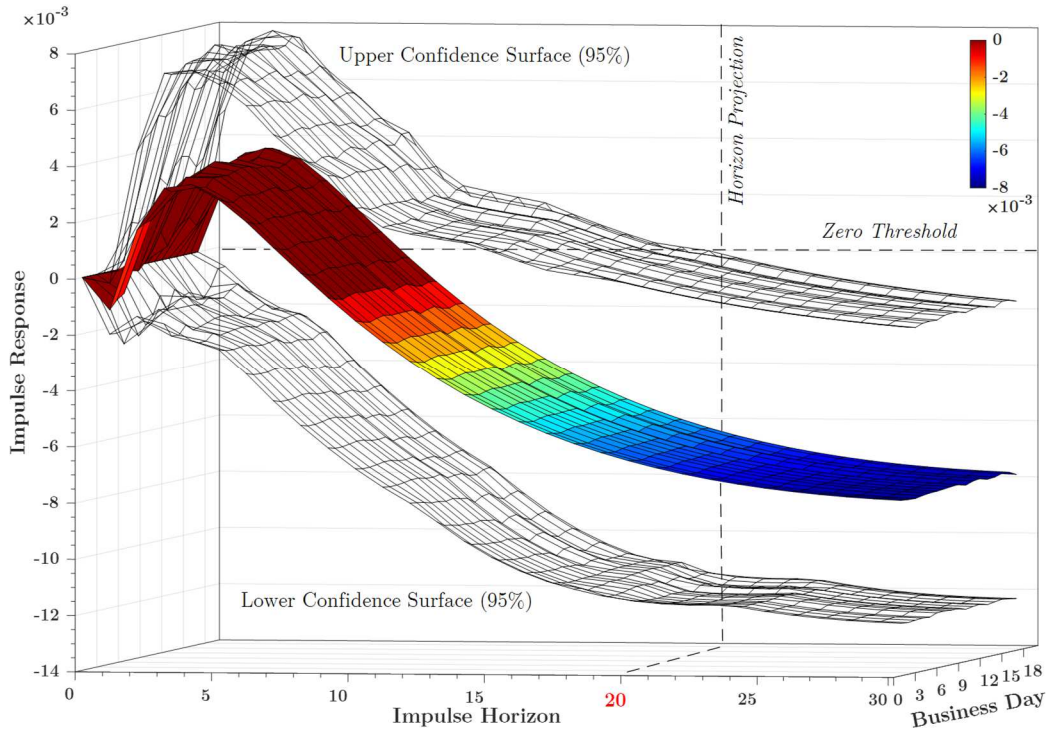


Figure 3. Responses of returns to a one standard deviation price-earnings shock (1951:2-2019:4)

Overall, the impulse response analysis shows that market returns are sensitive to the price-dividend and price-earnings ratios not only at long but also at short horizons. Although the magnitude of the response is similar, the price-dividend shock becomes significant 10 months earlier than the price-earnings shock which implies that investors pay more attention to dividends than earnings in valuing stocks. [4]

The responses of returns to one standard deviation CAPE (TRCAPE) shocks are statistically significant from month 9 (month 10) onwards; see **Figure 4** and **Figure 5**, respectively. Therefore, cyclically adjusted earnings measures produce shocks that match the timing of price-dividend shocks.[5]

[4] Within the dividend signaling theory, Aharony and Swary (1980) find that dividends convey information over and above earnings. Gopalan and Jayaraman (2012) find evidence for the superiority of dividends over earnings and flag the debate on earnings management being prone to information manipulation.

[5] For robustness reasons, we examine the impact of valuation ratio shocks on stock returns over time by running our models over the 1951:2-1986:1 period and then expanding our sample sequentially by 10 years. Results (available on request) were qualitatively similar to what we report here.

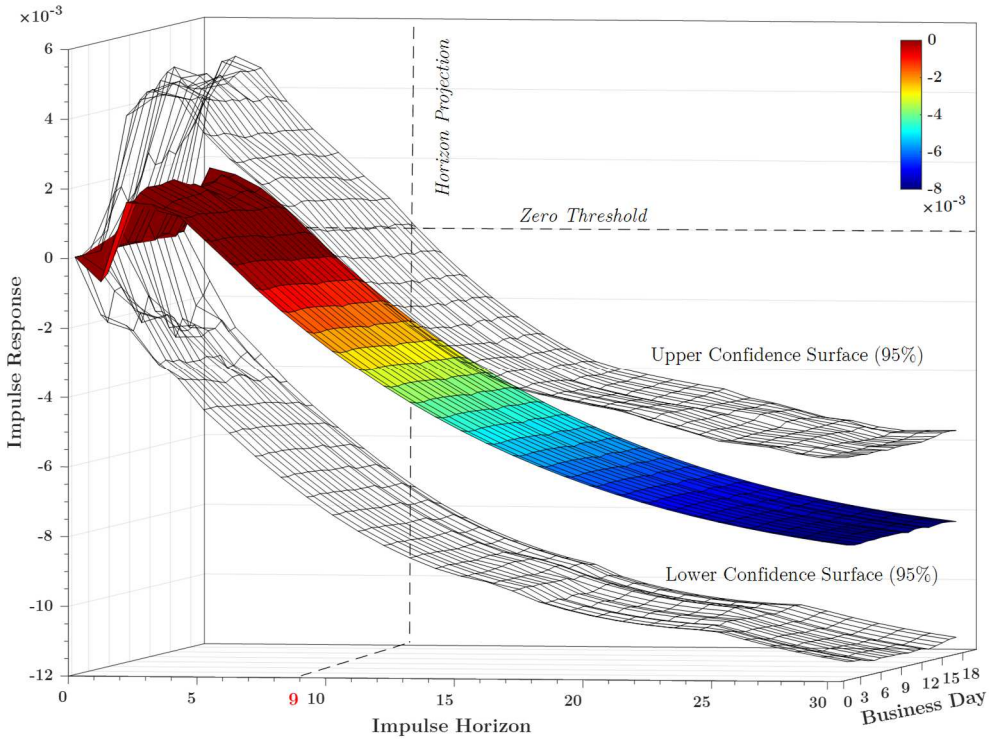


Figure 4. Responses of returns to a one standard deviation CAPE shock (1951:2-2019:7)

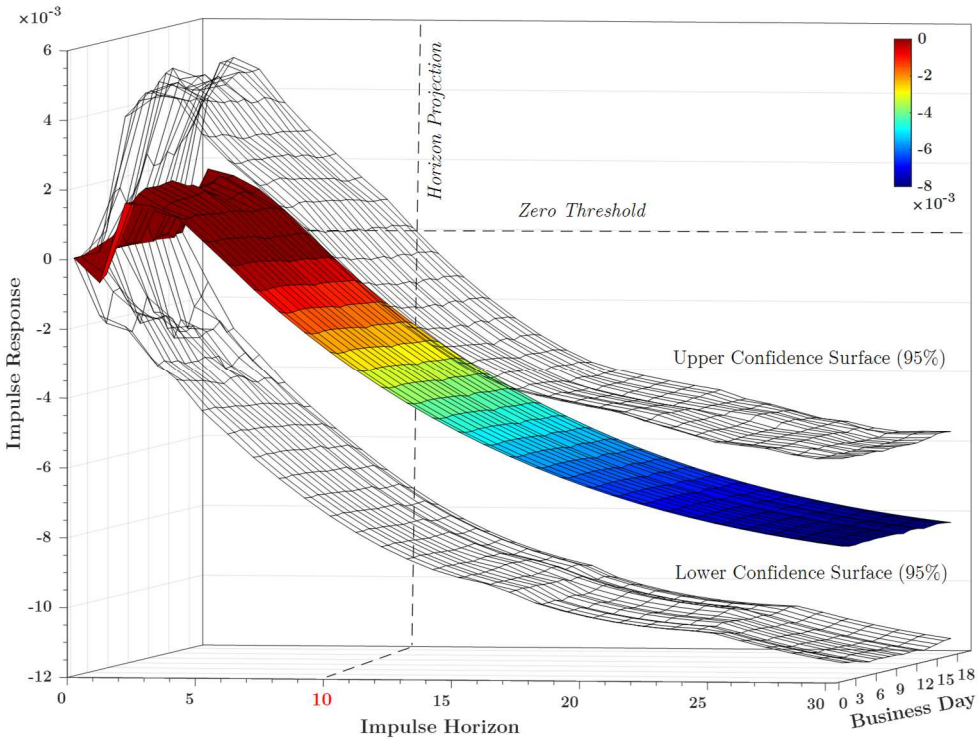


Figure 5. Responses of returns to a one standard deviation TRCAPE shock (1951:2-2018:9)

5. Conclusions

This paper focuses on the influence of cash flow measures on stock market index returns by considering the mixed frequency nature of the financial data to show that price-dividend and price-earnings ratios impact significantly on stock returns both at long and short horizons.

References

Aharony, J. and I. Swary (1980). Quarterly Dividend and Earnings Announcements and Stockholders' Returns: An Empirical Analysis. *Journal of Finance* **35**, 1-12.

Ang, A. and G. Bekaert (2006). Stock Return Predictability: Is it There? *The Review of Financial Studies* **20**, 651-707.

Bollerslev, T., L. Xu and H. Zhou (2015). Stock Return and Cash Flow Predictability: The Role of Volatility Risk. *Journal of Econometrics* **187**, 458-471.

Campbell, J.Y., and R.J. Shiller (1988). Stock Prices, Earnings, and Expected Dividends. *Journal of Finance* **43**, 661-676.

Choi K.H., C. Kim and C. Park (2017). Regime Shifts in Price-Dividend Ratios and Expected Stock Returns: A Present-Value Approach. *Journal of Money, Credit and Banking* **49**, 417-441.

Ghysels, E. (2016). Macroeconomics and the Reality of Mixed Frequency Data. *Journal of Econometrics* **193**, 294-314.

Ghysels, E., J.B. Hill and K. Motegi (2016). Testing for Granger Causality with Mixed Frequency Data. *Journal of Econometrics* **192**, 207-230.

Gopalan, R. and S. Jayaraman (2012). Private Control Benefits and Earnings Management: Evidence from Insider Controlled Firms. *Journal of Accounting Research* **50**, 117-157.

Jagannathan R. and B. Liu (2019). Dividend Dynamics, Learning, and Expected Stock Index Returns. *Journal of Finance* **74**, 401-448.

Jivraj, F. and R.J. Shiller (2017). The Many Colours of CAPE. Yale ICF Working Paper No. 2018-22.

Kelly, B. and H. Jiang (2014). Tail Risk and Asset Prices. *The Review of Financial Studies* **27**, 2841-2871.

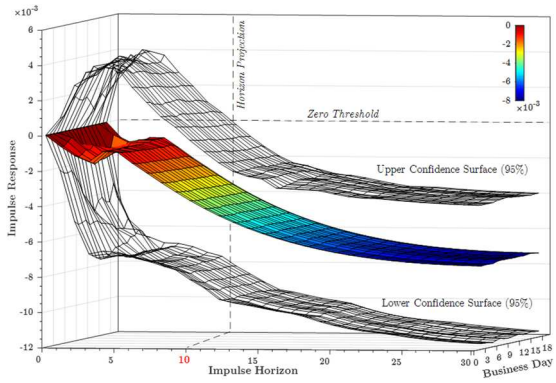
Lettau, M. and S. Ludvigson (2001). Consumption, Aggregate Wealth and Expected Stock Returns. *Journal of Finance* **56**, 815-849.

Lettau, M. and S. Ludvigson (2005). Expected Returns and Expected Dividend Growth. *Journal of Financial Economics* **76**, 583-626.

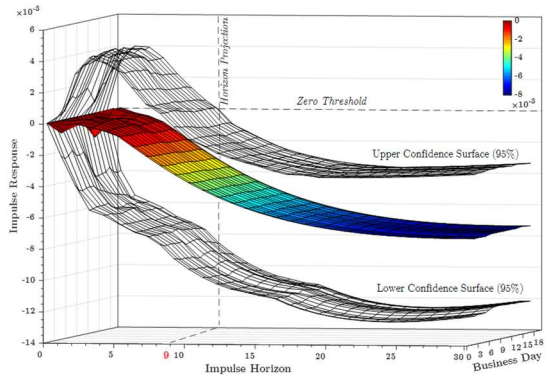
Rapach D.E. and M.E. Wohar (2006). In-sample vs. Out-of-sample Tests of Stock Return Predictability in the Context of Data Mining. *Journal of Empirical Finance* **13**, 231-247.

FOR THE ATTENTION OF REFEREES ONLY-NOT FOR PUBLICATION

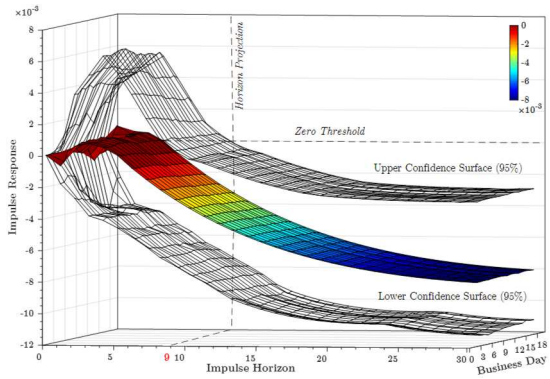
Appendix A: Sub-Sample Impulse Response Analysis



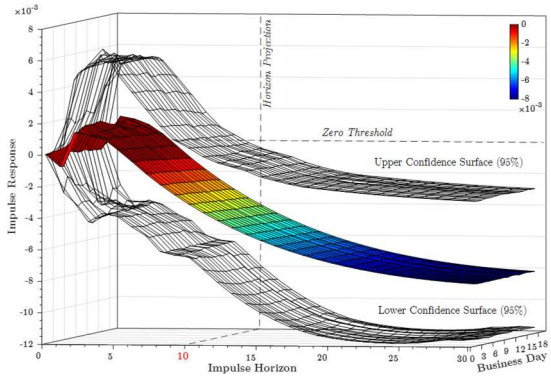
A1.a. Sample 1951:2-1986:1



A1.b. Sample 1951:2-1996:1

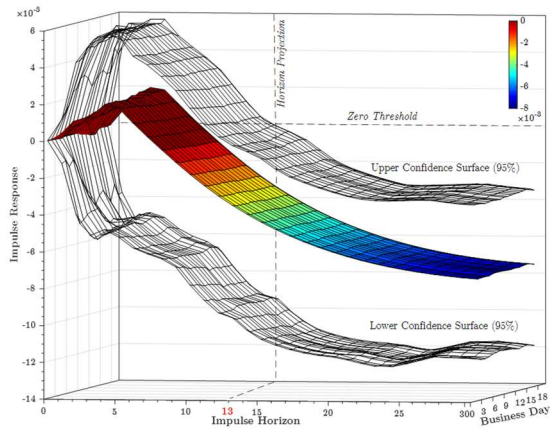


A1.c. Sample 1951:2-2006:1

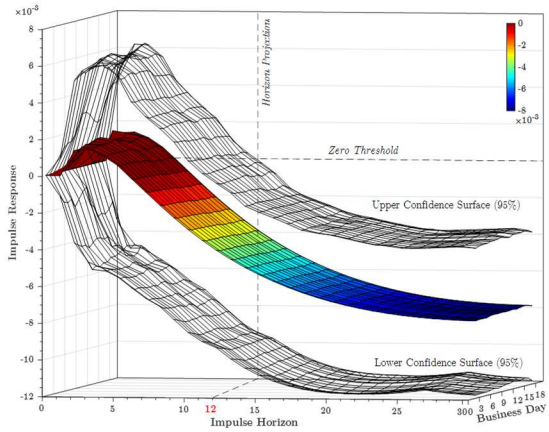


A1.d. Sample 1951:2-2016:1

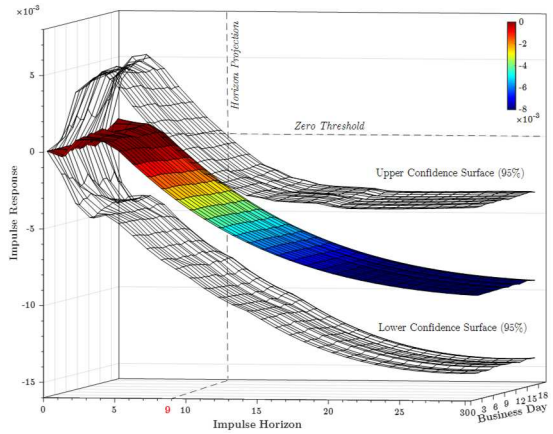
Figure A1. Sub-sample responses of returns to a one standard deviation price-dividend shock



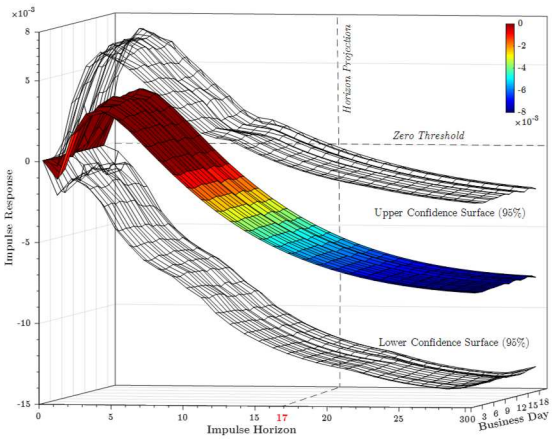
A2.a. Sample 1951:2-1986:1



A2.b. Sample 1951:2-1996:1

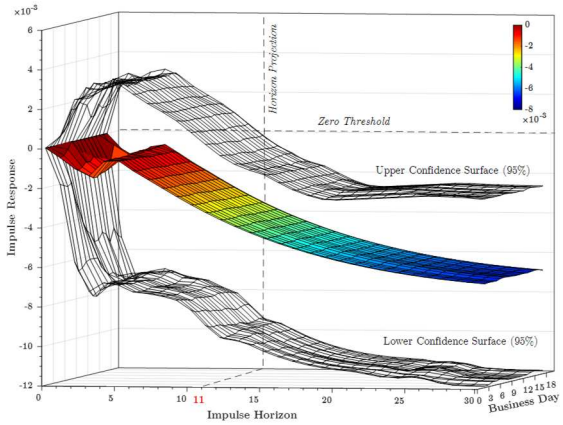


A2.c. Sample 1951:2-2006:1

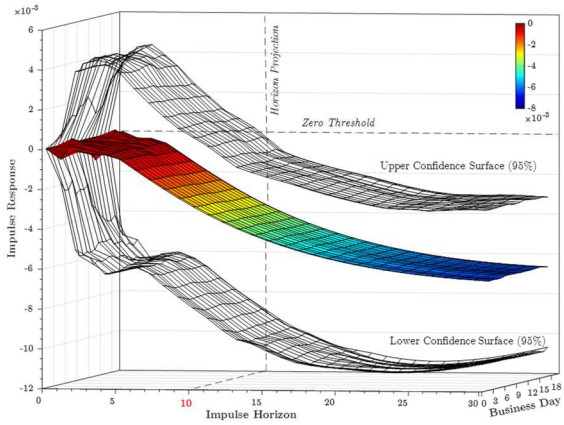


A2.d. Sample 1951:2-2016:1

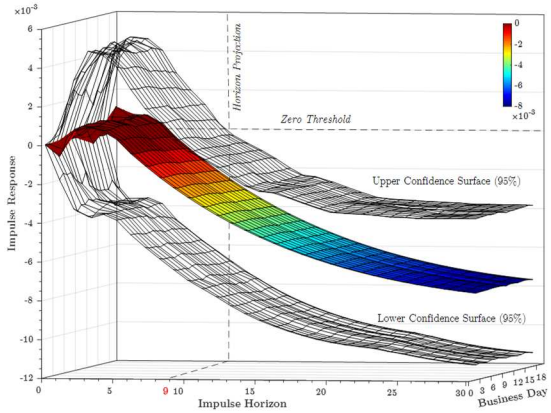
Figure A2. Sub-sample responses of returns to a one standard deviation price-earnings shock



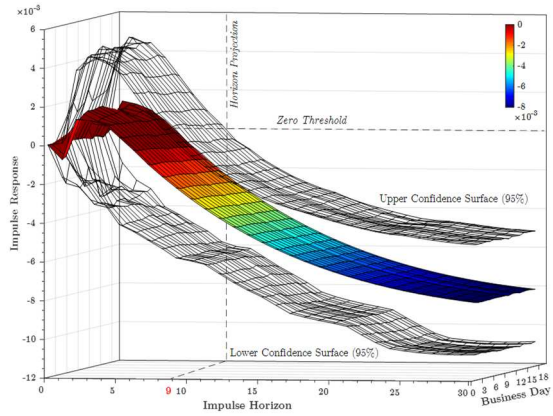
A3.a. Sample 1951:2-1986:1



A3.b. Sample 1951:2-1996:1

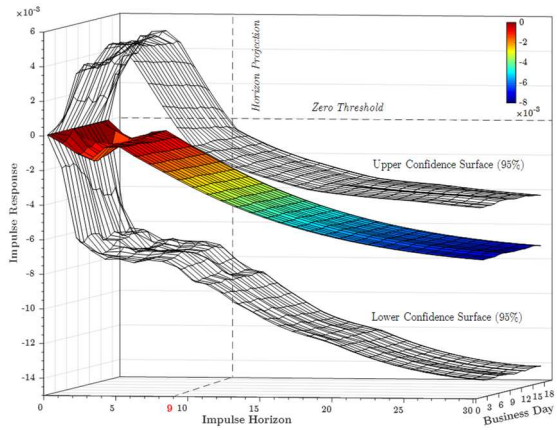


A3.c. Sample 1951:2-2006:1

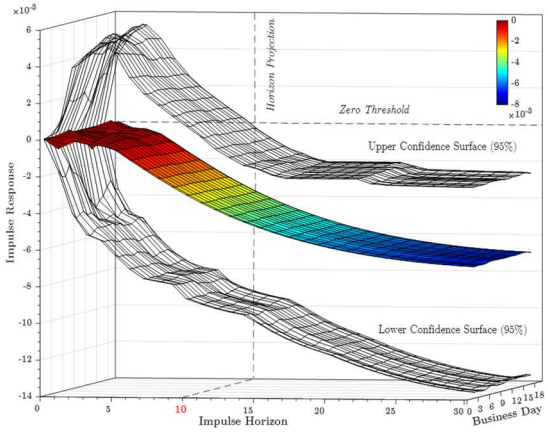


A3.d. Sample 1951:2-2016:1

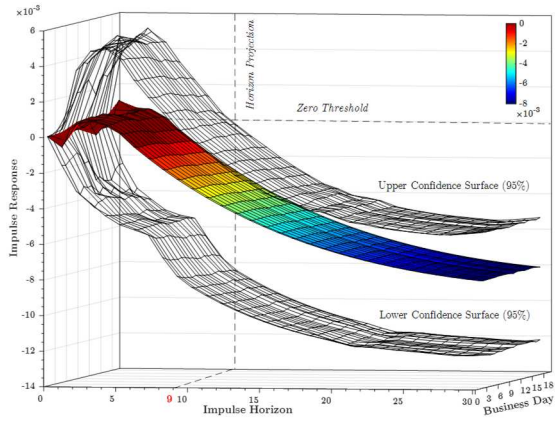
Figure A3. Sub-sample responses of returns to a one standard deviation CAPE shock



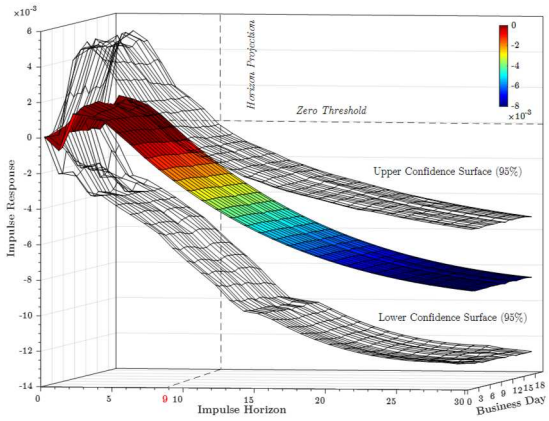
A4.a. Sample 1951:2-1986:1



A4.b. Sample 1951:2-1996:1



A4.c. Sample 1951:2-2006:1



A4.d. Sample 1951:2-2016:1

Figure A4. Sub-sample responses of returns to a one standard deviation TRCAPE shock