Sources of Return Predictability

BEATA GAFKA, PAVEL SAVOR, and MUNGO WILSON*

March 15, 2023

Abstract

A large literature establishes a set of predictors that robustly forecast future market returns, raising questions about these predictors' origins. We develop an approach to determine whether a particular predictor represents a proxy for fundamental risk, which is based on an intuitive assumption that risk-based predictors should be linked to new information about economic conditions. We show that an overwhelming majority of predictors forecast returns *either* on days with macroeconomic announcements *or* on the remaining days. This suggests that the sources of return predictability differ across predictors with some driven by fundamental risk and others having different origins. Shiller's excess volatility puzzle is confined to non-announcement days, indicating that the ability to forecast the noise component of stock market movements underlies much of the stock market return predictability documented in the literature.

Keywords: Announcements; Returns' predictability; Risk

JEL Classification: G12, G14

*Beata Gafka is at the Ivey Business School at University of Western Ontario, bgafka@ivey.ca. Pavel Savor is at the Kellstadt Graduate School of Business at DePaul University, psavor@depaul.edu. Mungo Wilson is at the Said Business School at Oxford University, Mungo.Wilson@sbs.ox.ac.uk.

1 Introduction

Stock returns are predictable, as shown by a large and still-growing literature.[1] While there exists little dispute about this basic result, the interpretation of many of the best-known predictors is another matter, with proposed explanations ranging from behavioral theories to various frictions to risk-based theories.

In this paper, we develop an approach for identifying predictors that represent proxies for fundamental risk, which is based on an intuitive assumption that such predictors should be linked to new information about economic fundamentals. We hypothesize that days when important macroeconomic news is scheduled to be announced (announcement days) are more likely to coincide with releases of such information than other days (non-announcement days).[2] Thus, predictors whose forecasting power is concentrated on announcement (non-announcement) days are more (less) likely to represent proxies for fundamental risk.

As in Savor and Wilson (2013) and subsequent papers, we define as announcement days those trading days when news about inflation, unemployment, or Federal Open Market Committee (FOMC) interest rate decisions is scheduled to be released (A-days) and all other trading days as non-announcement days (N-days). Using quarterly returns (aggregated separately for A- and N-days), we find that many widely used stock market predictors forecast returns only on non-announcement days. For example, log price/dividend ratio (pd_t, Campbell (1996), Litzenberger and Ramaswamy (1979)), long government yield (lty, Fama and French (1989)), treasury bill yield (tbl, Campbell (1987)), investment-to-capital ratio (i/k, Cochrane (1991)), production output gap ($ogap$, Cooper and Priestley (2009)), cyclical consumption (pce, Atanasov, Moller, and Priestley (2020), consumption fluctuations ($skew$, Colacito, Ghysels, Meng, and

[1] See Cochrane (2007), Goyal and Welch (2007), Campbell and Thompson (2007), and Harvey, Liu, and Zhu (2015), among others.

[2] Previous literature finds that announcement days are indeed special, both in terms of time-series (Savor and Wilson, 2013) and cross-sectional (Savor and Wilson, 2014) return patterns.

Siwasarit (2016)), and year-end economic growth characteristics (*gpce*, *gip*, Møller and Rangvid (2015)) forecast future returns accrued on N-days with a negative sign, while term spread (*tms*, Fama and French (1989); Campbell (1987)), and long government return spread (*ltr*, Fama and French (1989)) forecast future N-day returns with a positive sign. These predictors do not exhibit comparable ability to predict A-day returns, and most of the time the point estimates for these predictors have opposite signs (none are statistically significant). By contrast, default yield spread (dfy, Fama and French (1989)), stock return variance (sum of squared daily returns on the S&P 500, *svar*, Guo (2006)), and oil price changes (*wtexas*, Driesprong, Jacobsen, and Maat (2008)) forecast stock returns accrued on A-days with a positive sign but do not work on N-days. Similarly, although nearness to 52-week Dow high (*dtoy*, Li and Yu (2012)) predicts next quarter's stock returns accrued on A-days with a negative sign, the variable lacks predictive power for N-day returns. Only nearness to all-time Dow high (*dtoat*, Li and Yu (2012)) and average correlation of stock returns (*avgcor*, Pollet and Wilson (2010)) predict returns on both types of days. Even then, the magnitude of this relationship and its statistical significance are much higher on N-days than on A-days.

The above evidence shows a clear dichotomy exists between A-days and N-days with respect to their return predictors. Strikingly, the popular in the literature predictors overwhelmingly forecast returns only on non-announcement days, which supports the hypothesis underlying our approach that the two sets of days are different. Furthermore, it means that we can group predictors into those that are linked to economic news and those that are not. The predictors that are based on direct measures of the amount of risk in the economy (like *svar*), which according to asset pricing theory should forecast returns (but fail to do so in reality), forecast quarterly returns accrued on A-days (but not on N-days). For the predictors historically documented to forecast future stock returns, we show that while they forecast N-day returns, they do not exhibit explanatory power for A-day returns. Overall, these findings are consistent with the hypothesis that sources of return predictability differ across predictors, with direct risk-based measures driven by economic fundamentals and most of the others having different origins.

If not fundamental risk, what is the source of predictability for the wide range of variables forecasting only N-day returns? In order to answer this question, we revisit the 1980s excess volatility puzzle which claims that the observed price movements cannot be justified by subsequent events (Shiller, 1981). In particular, they cannot be explained by the stream of subsequent dividends since the realized prices move too much compared to the time-series of ex-post rational price (fundamental value) realization. Our tests show that the excess volatility puzzle defined in this way is very strong on N-days and limited on A-days. We use multivariate regressions ($Y_{t+1} = \beta_0 + \beta_1 r_t^A + \beta_2 r_t^N$) of future log changes in ex-post rational price (Y_{t+1}) on past returns accrued on both A-days (r_t^A) and N-days (r_t^N). The results highlight, once again, a clear difference between these two types of days. Quarterly and annual returns accrued on A-days are positively related to future changes in fundamental value with $\beta_1 = 0.01$ ($t(\beta_1) = 1.81$) and $\beta_1 = 0.046$ ($t(\beta_1) = 1.7$) in regressions using quarterly and annual frequency data, respectively. This relation does not hold for their N-day counterparts with β_2 not significantly different from zero at any frequency. Hence, although it may be true that the price movements on N-days are too big to be justified by subsequent dividends, this is not the case for price movements experienced on A-days. Using the Campbell and Shiller (1988) decomposition, we further show that *both* the excess volatility (with respect to the dividend discount model) puzzle *and* the residual volatility (with respect to the conditional CAPM) puzzles are confined to non-announcement days and absent from announcement days. This further confirms that variables predicting A-day returns and those predicting N-day returns are crucially different with respect to the source of their ability to predict stock market returns. While A-day predictors are driven by future fundamentals, N-day predictors seem to be predicting the "noise" component of stock market movements.

The combined evidence presented confirmes that A-day returns are in fact crucially different from their N-day counterpart. The returns accrued on A-days are more fundamentals-driven, explained by the amount of physical risk in the economy, and justified by subsequent changes in fundamental value. The same is not true for returns accrued on N-days, which are in turn where excess and residual volatility of the stock market are confined to. Consequently, predictors which forecast stock market returns accrued on A-days are crucially different to those which forecast stock market returns accrued on N-days when it comes to the source of the return predictability. Our methodology can be applied to other predictors discovered in the future.

Related literature. This paper relates to a strain of literature documenting the differential behavior of asset returns on A-days and N-days and the potential explanations for these findings. Savor and Wilson (2014) show that stock returns on said A-days are significantly higher and their patterns easier to reconcile with known asset pricing theories than their N-day counterparts. In particular, they show that while the CAPM holds on A-days, it fails to hold on N-days. Similarly Brooks, Katz, and Lustig (2018) show that while the expectations hypothesis holds on A-days, it fails thereafter and this failure increases in the length of the window considered. Attempts have been made to reconcile these differences within standard asset pricing models (Savor and Wilson, 2014) and to find the drivers behind the phenomenon pointing to increases in the price of risk as opposed to the quantity of risk on A-days (Savor, Wilson, and Puhl, 2015). Meanwhile, other authors argued this differential behavior of A-day returns is only a by-product of high ex-post returns on those days rather than an evidence of them being in any way special (Ernst, Gilbert, and Hrdlicka, 2019). The results in our paper show the dichotomy between announcement and non-announcement days is multifaceted with excess volatility and residual volatility puzzles confined to non-announcement days and absent from announcement days. We show that this striking dichotomy can be used to shed light on other asset pricing phenomena like predictability. Consequently, the paper also relates to the wider literature on returns predictability.

2 Data and summary statistics

The macroeconomic announcements considered are in line with those used by Savor and Wilson (2013) and Savor and Wilson (2014). As in these papers, inflation and unemployment announcement dates come form the US Bureau of Labor Statistics (https://www.bls.gov/) with the available time series staring in 1958. We follow the authors in using consumer price index (CPI) announcements up to and including February 1972. Producer price index (PPI) announcements are used between March 1972 and January 2018 (inclusive). This is because in that time period PPI numbers are reported a few days prior to the CPI ones thus diminishing the informational content of the CPI numbers. Between February 2018 and December 2019, for some months CPI is again released before PPI. In our analysis we use the date of the earlier of these two announcements. FOMC interest rate announcement days come from the Federal Reserve website and are available from 1978 onwards. Unscheduled FOMC meetings are excluded from the sample.

Data on stock market returns comes from Center for Research in Security Prices (CRSP). Data on risk-free interest rate comes from Professor Kenneth French's website. Our main stock market proxy is the CRSP NYSE, Amex, and Nasdaq value-weighted index of all listed shares. We collect daily values of this index between January 1953 and December 2022. We use those valuations and the daily risk free rate to construct log daily excess returns over this time period. These are then aggregated on a quarterly basis for all trading days in the given quarter ($r_t^{A\&N}$), all A-days in a given quarter (r_t^A), and all N-days in a given quarter (r_t^N). Panel A of Table 2 presents the summary statistics for these returns. Since our predictor variables run between 1953Q1 and 2021Q4, we focus on that time period here. We see that the average quarterly return on N-days over this time period (1.2%) is almost twice as large as the average quarterly return on A-days (0.6%). Compared to N-day quarterly returns, A-day returns are less volatile (0.03 vs. 0.08), exhibit lower autocorrelation (-0.02 vs. 0.1), and are less negatively skewed (-0.64 vs. -0.75).

We use a wide range of variables historically documented to be predictors of stock returns. Table 1 provides a summary of those, their abbreviations used throughout the paper, references to papers that introduced them to the literature, and the frequency at which the variables are computed. With the exception of the price dividend ratio, the variables' time series are courtesy of Amit Goyal, Ivo Welch, and Athanasse Zafirov, who have kindly shared their data with us. Panel B and Panel C of Table 2 provide summary statistics of the predictors considered.

3 Evidence A: Univariate predictive regressions

In this section we analyse the ability of the various predictors considered to forecast future aggregate quarterly stock market returns, their part accrued on A-days, and their part accrued on N-days.

Among the predictive variables summarized in Table 1, only *svar* constitutes a measure of the amount of physical risk on the market. As a result, it should be the only variable able to explain future stock returns since according to fundamental asset pricing theories the amount of physical risk should be the only driver behind stock market returns. Although the remaining variables are not such proxies for the physical amount of risk on the market, they have been historically found to predict returns.

In our analysis we use the univariate linear regression framework. We regress the relevant quarterly returns on the various predictors considered lagged by one quarter. The regression can be summarized as follows:

$$r_{t+1}^{i} = \alpha + \beta^{i} x_t, \quad i = (A\&N, A, N),$$

where $r_{t+1}^{A\&N}$, r_{t+1}^{A}, and r_{t+1}^{N} are the aggregate quarterly return, quarterly return accrued on A-days in a given quarter, and quarterly returns accrued on N-days in a given quarter, respectively. x_t is one of the predictors outlined in Table 1.

Panel A of Table 3 summarizes the regression results for variables which were found to be A-day but not N-day return predictors at quarterly frequency. We see that default yield spread (dfy) and oil price changes ($wtexas$) forecast returns accrued on A-days but lack predictive power for returns accrued over the whole quarter. A one percentage point increase in dfy ($wtextas$) leads to 79 (4) basis points *increase* in next quarter's A-day return. On the other hand, stock return variance ($svar$) nearness to 52-week Dow high (*dtoy*) forecast returns accrued on A-days and over the entire quarter as a whole despite lacking the predictive power for returns accrued on N-days (which constitute the largest share of days in any given quarter). In fact a one percentage point *increase* in $svar$ (*dtoy*) leads to 58 (6) basis points *increase* (*decrease*) in next quarter's A-day return and 90 (14) basis points *increase* (*decrease*) in next quarter's return accrued on both types of days.

Panel B of Table 3 shows that in-sample return predictability (if present) is overwhelmingly an N-day phenomenon. Term spread (*tms*) and long government return spread (ltr) are positively correlated with future quarterly N-day returns. An *increase* of one percentage point in *tms* (ltr) leads to 52 (21) basis points *increase* in next quarter's N-day returns. The relationship between the remaining predictors and future quarterly N-day returns is negative. In particular, some well-established return predictors such as log price/dividend ratio ($logPD$), long government yield (*lty*), treasury bill rate (*tbl*), and investment to capital ratio (*ik*) are negatively correlated with future stock market returns. A percentage point *increase* in *lty* (*tbl*) leads to a 31 (36) basis points *decrease* in next quarter's N-day return. Similarly, a percentage point *increase* in *ik* and *logPD* lead to 548 and 2 bps *decrease* in next quarter's N-day return, respectively. More recently discovered return predictors such as production output gap (*ogap*), cyclical consumption (*pce*), consumption fluctuations (*skew*), and year-end economic growth characteristics (*gpce*, *gip*) also forecast future N-day returns with negative sign. Overwhelming majority of N-day return predictors also forecast quarterly returns accrued on both types of days. This is unsurprising since N-days constitute the vast majority of days in a any given quarter.

Panel C shows that two of the predictors considered were statistically significant in predicting the quarterly returns accrued on *both* A-days and N-days between 1953 and 2021. These were nearness to all-time Dow high (*dtoat*) and average correlation of stock returns (*avgcor*). The magnitude of both relationships is higher for N-day returns than for A-day returns with $\beta^A = -0.03$, $\beta^N = -0.11$ for *dtoat* and $\beta^A = 0.05$ and $\beta^N = 0.13$ for *avcor*. Finally, *Panel D* shows that another fifteen variables historically found to predict stock market returns failed to do so in sample between 1953 and 2021.

The results outlined above highlight a startling dichotomy between predictors of returns accrued on A-days and those accrued on N-days. We observe that the vast majority of variables historically documented to forecast stock market returns, if at all statistically significant in univariate regressions between 1953 and 2021, predict the part of quarterly returns accrued on N-days but lack predictive power for their part accrued on A-days. On the contrary, variables which have roots in fundamental asset pricin theories and are proxies for physical risk on the market, such as stock variance (*svar*) are both economically and statistically significant predictors of future returns accrued on A-days but not on N-days.

These results are highly suggestive and intriguing. They seem to indicate that great many of well-known stock returns predictors suggested in the literature are not related to new information about fundamentals being revealed to the market (A-days) but seem to be predicting something different instead. In order to answer the question of what the source of this N-day predictability is, we proceed to concentrate on analysing the relationship between the changes in prices on those two types of days and the changes in ex-post rational prices (fundamental value) as calculated by Shiller (1981).

4 Evidence B: Excess and residual volatility – an N-day puzzle

Shiller (1981) shows that realized stock prices move too much to be justified by the subsequent changes in dividends. In this section we revisit the relationship between these price changes and the subsequent dividends independently for A-days and N-days. We show that quarterly (annual) returns accrued on A-days forecast changes in next quarter (year) ex-post rational price. The same is not true for quarterly (annual) returns accrued on N-days. In regressions of fundamental value (understood as sum of discounted ex-post realized dividends) changes on lagged returns accrued on these two types of days, the coefficient for returns accrued on A-days is positive and significantly higher than that for returns accrued on N-days.

In what follows, We use the Campbell and Shiller (1988) decomposition to show that *both* excess volatility puzzle (with respect to the dividend discount model) *and* the residual volatility (with respect to the conditional CAPM) are strictly N-day puzzles. This shows that predictability of A-day returns must have different source to the predictability of N-day returns.

4.1 Shiller's excess volatility puzzle

We follow Shiller (1981) methodology very closely to calculate the real values of prices and dividends, and their de-trended counterparts. The only difference compared to the original paper is that we sample prices at the end of each period while the original paper records them for its beginning. Since this slightly alters the formulas and as we acknowledge the time that has passed since the original work, in what follows we briefly summarize the main idea behind the original excess volatility puzzle and the steps taken to arrive at it.

Revisiting Shiller (1981) approach. For the simple efficient markets model to be correct the real price P_t at the end of time period t should be equal to:

$$P_t = \sum_{k=1}^{\infty} \gamma^k E_t D_{t+k},^{3} \tag{1}$$

where the D_t is the real dividend paid at time t and γ is a constant real discount factor. As in the original work, we assume all dividends D_t occur at the end of the relevant time period t. The constant real interest rate r is defined such that $\gamma = 1/(1+r)$ and has the property that $r = E_t(H_t)$, where H_t is the holding period return $H_t \equiv (\Delta P_{t+1} + D_{t+1})/P_t$.[4]

It is possible to restate the relationship in Equation (1) using detrended prices and dividends. Such detrending is done by restating those time series as a proportion of the long-run growth factor: $p_t = P_t/\lambda^{t-T}$, $d_t = D_t/\lambda^{t-T}$,[5] where T is the last period for which we have observations (the base period) and λ^{t-T} is the growth factor. The growth factor is calibrated by estimating a long-run exponential growth path for the time series of real prices. To this end, we regress $ln\left(P_t\right)$ on a constant and time and set $\lambda = e^b$, where b is the coefficient on time in $ln\left(P_t\right) = a + b * t$. It can be shown that the following holds for such detrended time series:

$$p_t = \sum_{k=1}^{\infty} \overline{\gamma}^k E_t d_{t+k},^{6} \tag{2}$$

where $\overline{\gamma} \equiv \lambda\gamma$ is the constant discount factor appropriate for the detrended time series of p_t and d_t. The corresponding discount rate $\overline{r}$, $\overline{\gamma} = 1/(1+\overline{r})$, can be shown to be equal to the mean detrended dividend divided by the mean detrended price:$\overline{r} = E\left(d\right)/E\left(p\right)$.[7]

The above relationship (2) can be re-written in terms of *ex-post* rational price series

[3] $P_t = \sum_{k=0}^{\infty} \gamma^{k+1} E_t D_{t+k}$ in Shiller (1981).

[4] $H_t \equiv (\Delta P_{t+1} + D_t)/P_t$ in Shiller (1981).

[5] $d_t = D_t/\lambda^{t+1-T}$ in Shiller (1981).

[6] $p_t = \sum_{k=0}^{\infty} \overline{\gamma}^{k+1} E_t d_{t+k}$ in Shiller (1981).

[7] This follows from taking unconditional expectation of both sides of equation (2) and solving for $\overline{r}$. Compare: footnote 7, page 424 in (Shiller, 1981).

p_t^*. Such *ex-post* rational price is the present value of actual subsequent dividends:

$$p_t^* = \sum_{k=1}^{\infty} \overline{\gamma}^k d_{t+k}.^{8} \tag{3}$$

As pointed out in the original paper, although the summation extends to infinity, with long enough time series we can observe a reasonably accurate approximation of p_t^*. Subject to the choice of terminal (base year) value of the *ex-post* rational price, p_T^*, the entire time series can be determined recursively by

$$p_t^* = \overline{\gamma}\left(p_{t+1}^* + d_{t+1}^*\right)^{9}$$

working backwards from the base year.

Calibration. We follow the above process for the CRSP NYSE, Amex, Nasdaq value-weighgted index of all common shares. Prices (p_t) are assumed to be the time series of CRSP index level excluding dividends. We calculate dividends (d_t) at monthly intervals between January 1950 and December 2019. Both prices and dividends are deflated using the CPI values provided by Professor Shiller on his website. We estimate the long-run exponential growth path using daily frequency data in line with $ln\left(P_t\right) = a + bt$ and set λ as e^b. $\overline{r}$ is estimated using monthly frequency data as the mean of the detrended dividend divided by the mean of the detrended price. For the purpose of calculating p_t^*, the terminal value p_T^* is set as the average of the detrended real price over the sample. For the purpose of calculating P_t^*, the terminal value P_T^* is set to the terminal value of the real price process: P_T.

Excess volatility. Figure (1) shows the behaviour of detrended real prices (p_t) and detrended ex-post rational prices (p_t^*) of the CRSP NYSE, Amex, Nasdaq value-weighgted index between 1950 and 2022. This figure corresponds to Figures (1) and (2) in Shiller (1981). As in the original work, also here we observe that the realized prices seem too volatile for their movements to be driven by new information about the stream of

[8] $p_t^* = \sum_{k=0}^{\infty} \overline{\gamma}^{k+1} d_{t+k}$ in Shiller (1981).

[9] $p_t^* = \overline{\gamma}\left(p_{t+1}^* + d_t^*\right)$ in Shiller (1981).

subsequent dividends (i.e. for the efficient markets model as proposed by Equation (1) to accurately describe the price process).

4.2 Which returns forecast future fundamental value changes?

Having computed the fundamental value (*ex-post* rational price) time series we are now in a position to test whether there is a difference in the informational content of changes in prices on A-days and N-days. In particular, we can now test which of them is better able to forecast future changes in fundamental value.

Let $\widetilde{P}^*_{t+1} = log(P^*_{t+1}) - log(P^*_t)$ and $\widetilde{p}^*_{t+1} = log(p^*_{t+1}) - log(p^*_t)$. These are the log change in the real ex-post rational price and the log change in the detrended real ex-post rational price, respectively. Setting Y_{t+1} to either $\widetilde{P}^*_{t+1}$ or $\widetilde{p}^*_{t+1}$, we can run the following regression to predict these changes at monthly, quarterly, and annual frequencies:

$$Y_{t+1} = \beta_0 + \beta_1 r^A_t + \beta_2 r^N_t, \tag{4}$$

where r^A_t and r^N_t are the part of the lagged return of the relevant frequency accrued on A-days and N-days, respectively. Since we are not only interested in the economic and statistical significance of β_1 and β_2, but also in formally testing whether $\beta_1 > \beta_2$, we turn to the following regression:

$$Y_{t+1} = \delta_0 + \delta_1(r^A_t + r^N_t) + \delta_2(r^A_t - r^N_t). \tag{5}$$

In the above, we observe that: $\beta_1 = \delta_1 + \delta_2$ and $\beta_2 = \delta_1 - \delta_2$. As a result, the following is true:

1. if $\beta_1 > \beta_2$, then $\delta_1 + \delta_2 > \delta_1 - \delta_2$, and so $\delta_2 > 0$
2. since $\delta_1 = 0.5(\beta_1 + \beta_2)$ and $\delta_2 = 0.5(\beta_1 + \beta_2)$, $\delta_2 > 0$ implies $\beta_1 > \beta_2$.

Hence, $\delta_2 > 0$ *if and only if* $\beta_1 > \beta_2$.

We estimate the relationships in Equation 4 and 5 at monthly, quarterly, and annual frequencies for two time periods: 1953 – 2022 and 1953 – 2010. This is to account for the fact that starting from January 2011 the terminal values of p_T^* and P_T^* become non-negligible share of p_t^* and P_t^*. Table 4 summarizes the results for $Y_{t+1} = \widetilde{P}_{t+1}^*$ i.e. the real ex-post rational price. Table 5 summarizes the results for $Y_{t+1} = \widetilde{p}_{t+1}^*$ i.e the *detrended* real ex-post rational price. In the below, unless separately specified, we summarize the regression estimates from the 1953 – 2010 time period.

At *monthly frequency*, we observe that although the relationship between the returns and the future fundamental value changes is positive for the share of monthly returns accrued on A-days ($\beta_1 = 0.003$ when $Y_{t+1} = p_t^*$; $\beta_1 = 0.004$ when $Y_{t+1} = P_t^*$) and negative for the share of monthly returns accrued on N-days ($\beta_2 = -0.002$ when $Y_{t+1} = p_t^*$; $\beta_2 = -0.003$ when $Y_{t+1} = P_t^*$), neither relationship is statistically significant. Similarly, although positive (0.003 across both specifications), δ_2 is not statistically significant ($t(\delta_2)$ equal to 1.3 and 1.28 for $Y_{t+1} = P_t^*$ and $Y_{t+1} = p_t^*$, respectively). There is no evidence that cumulative monthly A-day returns forecast future one-month-ahead changes of fundamental value better than their N-day counterparts. The same is, however, no longer true at quarterly and annual frequencies.

Quarterly frequency data shows that cumulative A-day returns forecast next quarter's log change of fundamental value with a positive sign. In the shorter time period (1953 – 2012) β_1 is equal to 0.010 and 0.014 when $Y_{t+1} = p_t^*$ and $Y_{t+1} = P_t^*$, respectively. This means that a one percentage point increase in the cumulative A-day return in a given quarter leads to a 1.4 basis points (1 basis point) increase in the next quarter's log change in (de-trended) real ex-post rational price P_t^* (p_t^*). This relationship is statistically significant at 10% confidence level ($t(\beta_1) = 1.8$ when $Y_{t+1} = p_t^*$ and $t(\beta_1) = 1.81$ when $Y_{t+1} = P_t^*$). The relationship between N-day returns and future fundamental value changes is negative but not statistically significantly so ($t(\beta_2) = -0.001$ for both $Y_{t+1} = p_t^*$ and $Y_{t+1} = P_t^*$). Estimates of equation (5) indicate that $\beta_1 > \beta_2$ at quarterly frequency: $\delta_2 = 0.006$ for $Y_{t+1} = p_t^*$ and $\delta_2 = 0.008$ for $Y_{t+1} = P_t^*$ and statistically significant at 10% level ($t(\delta_2) = 1.78$ using p_t^*; $t(\delta_2) = 1.82$ using P_t^*).

We find similar results using *annual frequency* data. Cumulative A-day returns at this frequency forecast next year's changes in fundamental value with a positive sign: $\beta_1 = 0.035$ for $Y_{t+1} = p_t^*$ and $\beta_1 = 0.046$ for $Y_{t+1} = P_t^*$. Both estimates are significant at 10% confidence level. As such, a one percentage point increase in the cumulative A-day return in a given year leads to a 4.6 (3.5) basis points increase in next year's log change in (detrended) real ex-post rational price P_t^* (p_t^*). As in the quarterly returns case, the relationship between cumulative annual N-day returns and next year's fundamental value changes is negative but not statistically significantly so. Finally, estimates of equation (5) indicate that $\beta_1 > \beta_2$ at annual frequency: $\delta_2 = 0.022$ for $Y_{t+1} = p_t^*$ and $\delta_2 = 0.029$ for $Y_{t+1} = P_t^*$ and statistically significant at 5% level ($t(\delta_2) = 1.97$ using p_t^*; $t(\delta_2) = 1.99$ using P_t^*).

These results are of crucial importance for the excess volatility puzzle (Shiller, 1981). Although it is true that price movements in general are too big to be attributed to new information about actual subsequent events, a clear dichotomy exists in this respect for aggregate movements on A-days and N-days. We show that, although the aggregate returns accrued on N-days can not be justified by subsequent changes in fundamental value, the same is not true for aggregate returns accrued on A-days. Quarterly returns accrued on A-days forecast future changes in ex-post rational price (detrended or not) with a positive sign. Furthermore, at both quarterly and annual frequency the ability to forecast such future changes in fundamental value is superior for A-day compared to N-day returns. This suggests that the movements in prices on A-days are not "too big relative to actual subsequent dividends" (Shiller, 1981).

4.3 Excess and residual volatility: an A/N-day decomposition

One well-known issue with Shiller (1981)'s fundamental value calculations is that they do not allow for time-varying discount rates. This may make the reader suspicious of the results derived above. Therefore, in what follows we use the Campbell and Shiller (1988) decomposition to further showcase that the excess volatility (i.e. excess volatility relative to the Dividend Discount Model) and residual volatility (i.e. excess volatility relative to the Conditional CAPM) phenomena are confined to N-days and (almost) absent from A-days.

Derivations Let us define the following variables:

$$pd_t = \ln\left(\frac{P_t}{D_t}\right)$$
$$r_{t+1} = \ln(1 + R_{t+1})$$
$$\Delta d_{t+1} = \ln\left(\frac{D_{t+1}}{D_t}\right)$$

$$\begin{aligned}
\sigma_R^2 &= Var[r_{t+1}] = Var[r_{t+1}^A + r_{t+1}^N] \\
\sigma_A^2 &= Var[r_{t+1}^A] \\
\sigma_N^2 &= Var[r_{t+1}^N]
\end{aligned}$$

$$\begin{aligned}
\overline{pd} &= E[pd_t] \\
\rho &= (1 + \exp(-\overline{pd}))^{-1} \\
k &= -\ln\rho - (1-\rho)\ln(1/\rho - 1)
\end{aligned}$$

Then Campbell-Shiller derive:

$$r_{t+1} \approx k + \Delta d_{t+1} + \rho pd_{t+1} - pd_t. \tag{6}$$

Rearranging (6) gives us

$$pd_t \approx k + \Delta d_{t+1} - r_{t+1} + \rho pd_{t+1},$$

which can be iterated forward to derive

$$pd_{t+1} \approx k + \Delta d_{t+2} - r_{t+2} + \rho pd_{t+2}.$$

Substituting this last expression into (6) gives

$$\begin{aligned}
r_{t+1} &\approx k + \Delta d_{t+1} + \rho(k + \Delta d_{t+2} - r_{t+2} + \rho pd_{t+2}) - pd_t \\
&= (1+\rho)k + \Delta d_{t+1} + \rho\Delta d_{t+2} - \rho r_{t+2} + \rho^2 pd_{t+2} - pd_t \\
&= \frac{1-\rho^{1+1}}{1-\rho}k + \Sigma_{j=0}^{1}\rho^j \Delta d_{t+1+j} - \Sigma_{j=1}^{1}\rho^j r_{t+1+j} + \rho^{1+1} pd_{t+1+1} - pd_t
\end{aligned}$$

Repeating these iterations T times gives:

$$r_{t+1} \approx \frac{1-\rho^{T+1}}{1-\rho}k + \Sigma_{j=0}^{T}\rho^{j}\Delta d_{t+1+j} - \Sigma_{j=1}^{T}\rho^{j} r_{t+1+j} + (\rho^{T+1} pd_{t+T+1} - pd_t). \quad (7)$$

Since, by construction:

$$r_{t+1} = r_{t+1}^{A} + r_{t+1}^{N},$$

the following holds:

$$r_{t+1} = r_{t+1}^{A} + r_{t+1}^{N} \approx \frac{1-\rho^{T+1}}{1-\rho}k + \Sigma_{j=0}^{T}\rho^{j}\Delta d_{t+1+j} - \Sigma_{j=1}^{T}\rho^{j} r_{t+1+j} + (\rho^{T+1} pd_{t+T+1} - pd_t).$$

Since the variance of the LHS of (7) equals the covariance of the LHS with the RHS, and since dividing both sides by $Var[r_{t+1}]$ yields both sides to equal (approximately) one, it can be shown that the following holds true:

$$\begin{aligned} & \frac{Cov\left[r_{t+1}^{A}, \Sigma_{j=0}^{T}\rho^{j}\Delta d_{t+1+j}\right]}{Var[r_{t+1}^{A}]}\frac{Var[r_{t+1}^{A}]}{Var[r_{t+1}]} + \frac{Cov\left[r_{t+1}^{N}, \Sigma_{j=0}^{T}\rho^{j}\Delta d_{t+1+j}\right]}{Var[r_{t+1}^{N}]}\frac{Var[r_{t+1}^{N}]}{Var[r_{t+1}]} \\ & -\frac{Cov\left[r_{t+1}^{A}, \Sigma_{j=1}^{T}\rho^{j} r_{t+1+j}\right]}{Var[r_{t+1}^{A}]}\frac{Var[r_{t+1}^{A}]}{Var[r_{t+1}]} - \frac{Cov\left[r_{t+1}^{N}, \Sigma_{j=1}^{T}\rho^{j} r_{t+1+j}\right]}{Var[r_{t+1}^{N}]}\frac{Var[r_{t+1}^{N}]}{Var[r_{t+1}]} \\ & +\frac{Cov\left[r_{t+1}^{A}, \rho^{T+1} pd_{t+T+1} - pd_t\right]}{Var[r_{t+1}^{A}]}\frac{Var[r_{t+1}^{A}]}{Var[r_{t+1}]} + \frac{Cov\left[r_{t+1}^{N}, \rho^{T+1} pd_{t+T+1} - pd_t\right]}{Var[r_{t+1}^{N}]}\frac{Var[r_{t+1}^{N}]}{Var[r_{t+1}]} \\ \approx\ & 1. \end{aligned} \quad (8)$$

The expressions of the form

$$\frac{Cov[r_{t+1}^{k}, y_{t+1}]}{Var[r_{t+1}^{k}]},\ k \in [A, N]$$

are just the betas from a univariate regression of the form

$$y_{t+1} = \alpha_k + \beta_k r_{t+1}^{k} + \varepsilon_{t+1}^{k}$$

or if $Cov[r_{t+1}^{A}, r_{t+1}^{N}] = 0$, they are also the betas from a bivariate regression of the form

$$y_{t+1} = \alpha + \beta_A r_{t+1}^{A} + \beta_N r_{t+1}^{N} + \varepsilon_{t+1}.$$

This decomposition outlined by (8) allows us to measure the shares of A-day and N-day returns in total stock market volatility. It also allows us to measure how much each type of market return contributes to excess volatility. At a high value of T, the share of excess volatility of returns under each regime are given by

$$S_{EV}^{A} = \frac{Cov\left[r_{t+1}^{A}, \rho^{T+1} pd_{t+T+1} - pd_{t}\right]}{Var[r_{t+1}]} - \frac{Cov\left[r_{t+1}^{A}, \Sigma_{j=1}^{T} \rho^{j} r_{t+1+j}\right]}{Var[r_{t+1}]} \tag{9}$$

and

$$S_{EV}^{N} = \frac{Cov\left[r_{t+1}^{N}, \rho^{T+1} pd_{t+T+1} - pd_{t}\right]}{Var[r_{t+1}]} - \frac{Cov\left[r_{t+1}^{N}, \Sigma_{j=1}^{T} \rho^{j} r_{t+1+j}\right]}{Var[r_{t+1}]}. \tag{10}$$

That is not the end of the story. A-day returns will account for a positive share of excess volatility under this measure. However, some of that share is likely due to changing measures of fundamental risk such as stock market return variance, which is somewhat persistent, but nowhere near as persistent as dividend-price ratios. If, following the Conditional CAPM, we believe that changing risk-free rates and changing measures of stock market variance are not really *irrational* drivers of changes in returns, then we can further decompose these measures into shares due to risk-free rates, market return variances, and a residual. It's only the residual that requires explanation beyond the CCAPM.

We show that the above residual is indeed almost entirely due to N-day returns. This is done by taking the CCAPM as a benchmark, and then allowing it to vary across regimes:

$$\begin{aligned} r_{t+1}^{A} &= r_{f,t+1}^{A} + \gamma^{A} Var_{t}[r_{t+1}] + \upsilon_{t+1}^{A} \\ r_{t+1}^{N} &= r_{f,t+1}^{N} + \gamma^{N} Var_{t}[r_{t+1}] + \upsilon_{t+1}^{N} \end{aligned} \tag{11}$$

where $Var_t[r_{t+1}]$ is the conditional expectation of the physical variance of market returns and υ_{t+1} is the residual.

As shown by Savor and Wilson (2013), γ^A is positive and significant while γ^N is not. Moreover, $r_{f,t+1}$ is slightly lower on A-days. Imposing (11) on (8) gives a rather lengthy expression which can be used to back out a residual volatility, as opposed to excess volatility, for returns under each regime.

From the above,

$$
\begin{aligned}
r_{t+1} &= r^A_{t+1} + r^N_{t+1} \\
&= r^A_{f,t+1} + \gamma^A Var_t[r_{t+1}] + \upsilon^A_{t+1} + r^N_{f,t+1} + \gamma^N Var_t[r_{t+1}] + \upsilon^N_{t+1} \\
&= r^A_{f,t+1} + r^N_{f,t+1} + \gamma^A Var_t[r_{t+1}] + \gamma^N Var_t[r_{t+1}] + \upsilon^A_{t+1} + \upsilon^N_{t+1}.
\end{aligned}
$$

Then the middle two terms in (8) can each be replaced:

$$
\begin{aligned}
&\frac{Cov\left[r^{k=A,N}_{t+1}, \Sigma^T_{j=1}\rho^j r_{t+1+j}\right]}{Var[r^{k=A,N}_{t+1}]} \\
= \; &\frac{Cov\left[r^{k=A,N}_{t+1}, \Sigma^T_{j=1}\rho^j r^A_{f,t+1+j}\right]}{Var[r^{k=A,N}_{t+1}]} + \frac{Cov\left[r^{k=A,N}_{t+1}, \Sigma^T_{j=1}\rho^j r^N_{f,t+1+j}\right]}{Var[r^{k=A,N}_{t+1}]} \\
&+ \frac{\gamma^A Cov\left[r^{k=A,N}_{t+1}, \Sigma^T_{j=1}\rho^j Var_{t+j}[r_{t+1+j}]\right]}{Var[r^{k=A,N}_{t+1}]} + \frac{\gamma^N Cov\left[r^{k=A,N}_{t+1}, \Sigma^T_{j=1}\rho^j Var_{t+j}[r_{t+1+j}]\right]}{Var[r^{k=A,N}_{t+1}]} \\
&+ \frac{Cov\left[r^{k=A,N}_{t+1}, \Sigma^T_{j=1}\rho^j \upsilon^A_{t+1+j}\right]}{Var[r^{k=A,N}_{t+1}]} + \frac{Cov\left[r^{k=A,N}_{t+1}, \Sigma^T_{j=1}\rho^j \upsilon^N_{t+1+j}\right]}{Var[r^{k=A,N}_{t+1}]}
\end{aligned}
$$

The resulting variance decomposition contains sixteen terms, but many of them will be zero. For example we know from previous results that the covariance of r_A with future returns at any horizon is zero once variance-driven effects have been removed. According to this line of reasoning, residual volatility (i.e. excess volatility relative to the Conditional CAPM) as opposed to excess volatility (i.e. excess volatility relative to the Dividend Discount Model) shares under each type of regime are then:

$$
S^A_{RV} = \frac{Cov\left[r^A_{t+1}, \rho^{T+1} pd_{t+T+1} - pd_t\right]}{Var[r_{t+1}]} - \frac{Cov\left[r^A_{t+1}, \Sigma^T_{j=1}\rho^j \upsilon^A_{t+1+j}\right]}{Var[r_{t+1}]} \tag{12}
$$

and

$$
S^N_{RV} = \frac{Cov\left[r^N_{t+1}, \rho^{T+1} pd_{t+T+1} - pd_t\right]}{Var[r_{t+1}]} - \frac{Cov\left[r^N_{t+1}, \Sigma^T_{j=1}\rho^j \upsilon^N_{t+1+j}\right]}{Var[r_{t+1}]}. \tag{13}
$$

In line with previous results, we would expect S_{RV}^{A} to be (almost) zero and S_{RV}^{N} to be (almost) one, for all T.

Evidence Equipped with the above derivations and hypothesis we proceed to analyse the A-day and N-day components of excess (with respect to the Dividend Discount Model) and residual (with respct to conditional CAPM) volatility using the wider CRSP universe index (spanning NYSE, Amex, and Nasdaq stocks) as a proxy.

The $Var_t[r_{t+1}]$ variable (*expected variance*) – the estimate of aggregate risk according to conditional CAPM – is calculated in line with Savor and Wilson (2014) as the conditional expectation of one-quarter-ahead variance of daily market returns on the corresponding index. This conditional forecast is computed as a function of contemporaneous (quarter t) excess returns accrued on a-days, excess returns accrued on n-days, realized variance (annualized average squared daily excess market return), and a constant. The corresponding coefficients are calibrated using constrained least squares (where the RV forecast is constrained to be non-negative) predictive regression of realized variance on the lagged variables mentioned before using quarterly data between 1964Q1 and 2022Q4. [10] Figure 2 compares the realized and expected variance between 1964 and 2022 computed for the wider CRSP universe index.

We then proceed to calculate the excess and residual volatility components given by equations 9, 10, 12, and 13 using quarterly data and wider CRSP universe index. Figure 3 shows the A-day and N-day components to excess volatility with respect to the dividend discount model. Figure 4 shows the A-day and N-day components to residual volatility with respect to conditional CAPM. The figures show that the contribution of A-day returns to *both* excess volatility with respect to dividend discount model and residual volatility with respect to the conditional CAPM is almost non-existent for high values of T (i.e. the number of quarters). In fact both the excess and residual volatility puzzles are almost entirely N-day phenomena.

[10]Savor and Wilson (2014) explain the model selection in greater detail.

All in all, A-day and N-day returns are visibly different on many levels as evidenced here and in previous work (Savor and Wilson (2013), Savor et al. (2015)). Hence, the sources of predictability of A-day and N-day returns must be different. The results in this paper suggest that while A-day returns are driven by economic fundamentals, their N-day counterparts are not. We argue that the excess and residual volatility results documented above suggest that it is the ability to forecast the "noise" component of stock market movements (as opposed to fundamental risk), that is the source of a lot of the stock market return predictability documented in the literature.

5 Conclusion

In this paper, we develop an approach to determine whether a particular predictor represents a proxy for fundamental risk or not. Our methodology is based on the intuitive assumption that risk-based predictors should be linked to new information about economic conditions. As our measure of such new information we use days when important macroeconomic announcements are released. In support of this hypothesis, we show that A-day returns are positively related to future changes in fundamental value (understood as a discounted sum of ex-post realized dividends) while N-days returns are not. Furthermore, we show that both the excess volatility (with respect to dividend discount model) and the residual volatility (with respect to the conditional CAPM) phenomena are limited to N-days and absent from A-days. We use this multifacet dichotomy to infer about the sources of return predictability.

We study a wide range of well-known predictors and find that (with very few exceptions) they forecast returns accrued either on days with macroeconomic announcements (A-days) or on days when no such announcements are made (N-days). In the limited cases when the predictor forecasts returns on both types of days, both the magnitude and the statistical significance of this relationship are overwhelmingly concentrated on N-days. These results allow us to group predictors into those that are linked to economic fundamentals and those that are not. More specifically, predictors based on direct measures of the amount of risk in the economy, which according to asset pricing theory should forecast returns, forecast the share of quarterly returns accrued on A-days but not their share accrued on N-days. The opposite holds for predictors historically documented to forecast future stock market returns – they forecast only the part of returns accrued on N-days but lack predictive power for their share accrued on A-days.

Together, these results suggest that the sources of return predictability differ across predictors. While direct risk-based measures are backed by future economic fundamentals, the remaining ones have different origins. We argue our excess and residual volatility results suggest that the N-day returns predictors possess superior ability to explain the "noise" component of stock market returns. The methodology and results presented can be further used to evaluate other predictors of asset returns.

References

Victoria Atanasov, Stig V. Moller, and Richard Priestley. Consumption fluctuations and expected returns. *The Journal of Finance*, 75(3):1677–1713, 2020. doi: https://doi.org/10.1111/jofi.12870. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12870.

Jordan Brooks, Michael Katz, and Hanno Lustig. Post-FOMC announcement drift in U.S. bond markets. *NBER Working Papers*, 25127, 2018.

John Y. Campbell. Stock returns and the term structure. *Journal of Financial Economics*, 18(2):373–399, 1987.

John Y. Campbell. Understanding risk and return. *Journal of Political Economy*, 104 (2):298–345, 1996.

John Y. Campbell and Robert J. Shiller. Stock prices, earnings, and expected dividends. *The Journal of Finance*, 43(3):661–676, 1988. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/2328190.

John Y. Campbell and Samuel B. Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531, 11 2007.

John Cochrane. Production-based asset pricing and the link between stock returns and economic fluctuations. *The Journal of Finance*, 46(1):209–237, 1991.

John H. Cochrane. The dog that did not bark: A defense of return predictability. *The Review of Financial Studies*, 21(4):1533–1575, 09 2007.

Riccardo Colacito, Eric Ghysels, Jinghan Meng, and Wasin Siwasarit. Skewness in expected macro fundamentals and the predictability of equity returns: Evidence and theory. *The Review of Financial Studies*, 29(8):2069–2109, 2016. ISSN 08939454, 14657368. URL http://www.jstor.org/stable/43866076.

Ilan Cooper and Richard Priestley. Time-varying risk premiums and the output gap. *The Review of Financial Studies*, 22(7):2801–2833, 2009. ISSN 08939454, 14657368. URL http://www.jstor.org/stable/40247688.

Gerben Driesprong, Ben Jacobsen, and Benjamin Maat. Striking oil: Another puzzle? *Journal of Financial Economics*, 89(2):307–327, 2008. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2007.07.008. URL https://www.sciencedirect.com/science/article/pii/S0304405X08000901.

Rory Ernst, Thomas Gilbert, and Christopher Hrdlicka. More than 100% of the equity premium: How much is really earned on macroeconomic announcement days? *Working paper*, 2019.

Eugene F. Fama and Kenneth R. French. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1):23–49, 1989.

Amit Goyal and Ivo Welch. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508, 03 2007.

Hui Guo. On the out-of-sample predictability of stock market returns. *The Journal of Business*, 79(2):645–670, 2006.

Campbell R. Harvey, Yan Liu, and Heqing Zhu. ... and the cross-section of expected returns. *The Review of Financial Studies*, 29(1):5–68, 10 2015.

Jun Li and Jianfeng Yu. Investor attention, psychological anchors, and stock return predictability. *Journal of Financial Economics*, 104(2):401–419, 2012. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2011.04.003. URL https://www.sciencedirect.com/science/article/pii/S0304405X11002121. Special Issue on Investor Sentiment.

Robert H. Litzenberger and Krishna Ramaswamy. The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics*, 7(2):163–195, 1979.

Stig V. Møller and Jesper Rangvid. End-of-the-year economic growth and time-varying expected returns. *Journal of Financial Economics*, 115(1):136–154, 2015. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2014.08.006. URL https://www.sciencedirect.com/science/article/pii/S0304405X14001810.

Joshua M. Pollet and Mungo Wilson. Average correlation and stock market returns. *Journal of Financial Economics*, 96(3):364–380, 2010. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2010.02.011. URL https://www.sciencedirect.com/science/article/pii/S0304405X10000437.

Pavel Savor and Mungo Wilson. How much do investors care about macroeconomic risk? Evidence from scheduled economic announcements. *Journal of Financial and Quantitative Analysis*, 48(2):343–375, 2013.

Pavel Savor and Mungo Wilson. Asset pricing: A tale of two days. *Journal of Financial Economics*, 113(2):171–201, 2014.

Pavel Savor, Mungo Wilson, and Martin Puhl. Uncertainty premia for small and large risks. *Working Paper*, 2015.

Robert J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review*, 71(3):421–436, 1981.

Figures

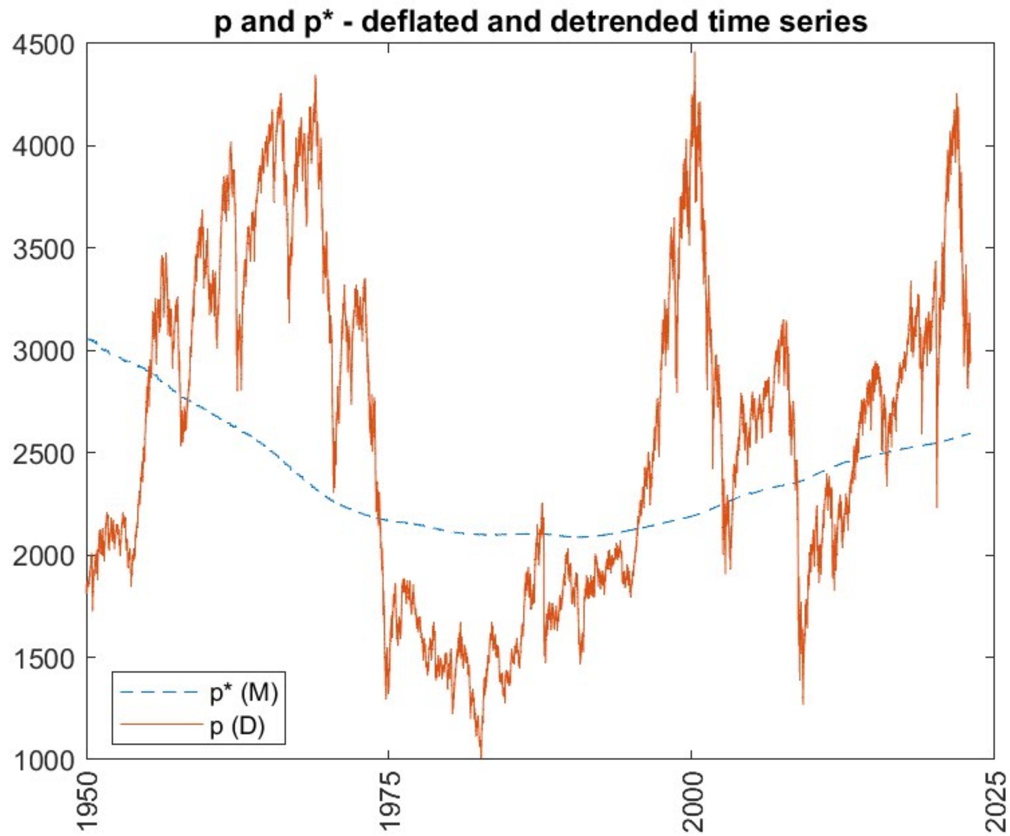

Figure 1: **Deflated and detrended market price (p) and fundamental value (p*)**

This figure shows the behaviour of detrended real prices (p_t) and corresponding ex-post rational prices (p_t^*) for the CRSP NYSE, Amex, Nasdaq value-weighted index between 1950 and 2022. It corresponds to figures (1) and (2) in Shiller (1981).

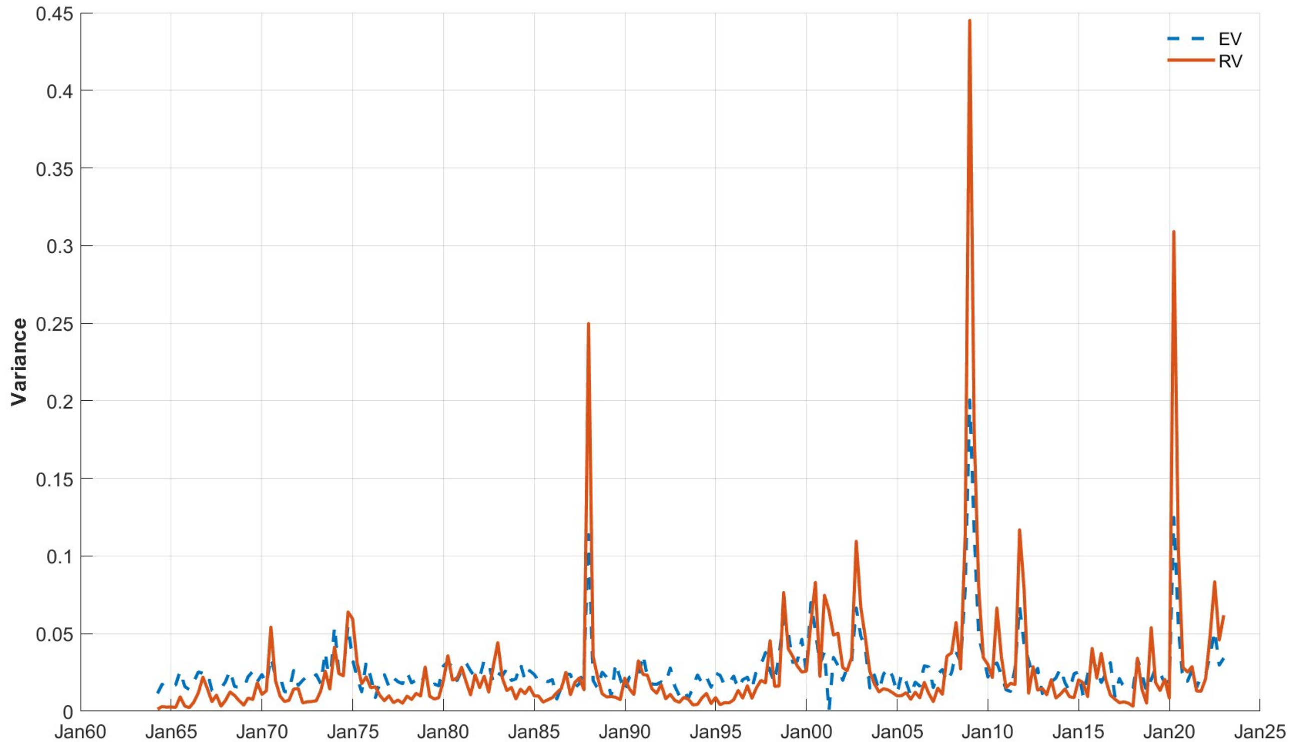

Figure 2: **Expected and Realized Variance computed using wider CRSP index**

The figure plots the realized variance of quarterly log excess market returns (RV) and its one-quarter-ahead forecast (EV) between 1964 and 2022. EV is a linear combination of RV, A-day, and N-day log excess returns.

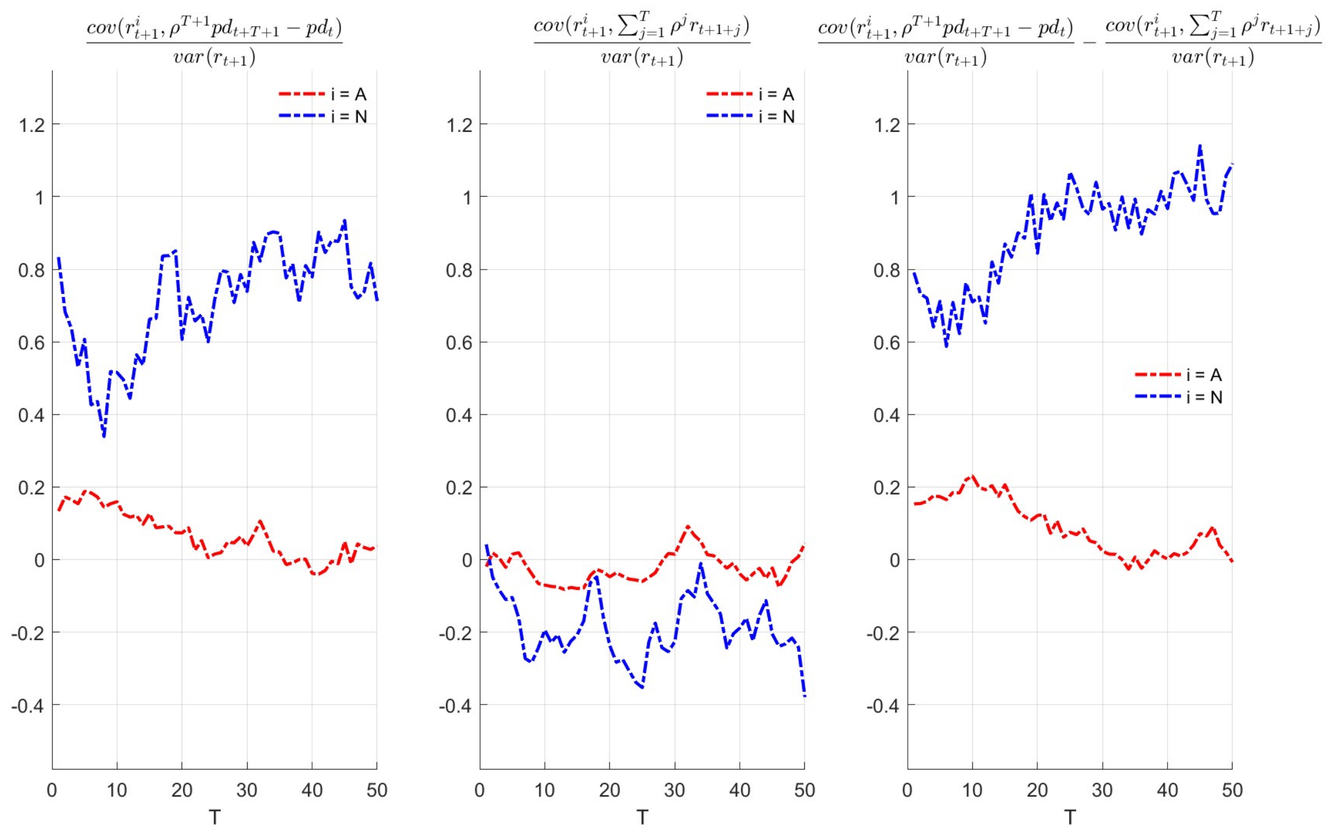

Figure 3: **Excess variance of the CRSP index – A- and N-day components**

The figure shows the contribution of quarterly A-day and N-day returns to the excess volatility of aggregate quarterly returns and its components as defined in equations 9 and 10 for various values of T. The returns and dividends correspond to the CRSP value-weighted index. Returns and dividends data cover the period between 1964 and 2022.

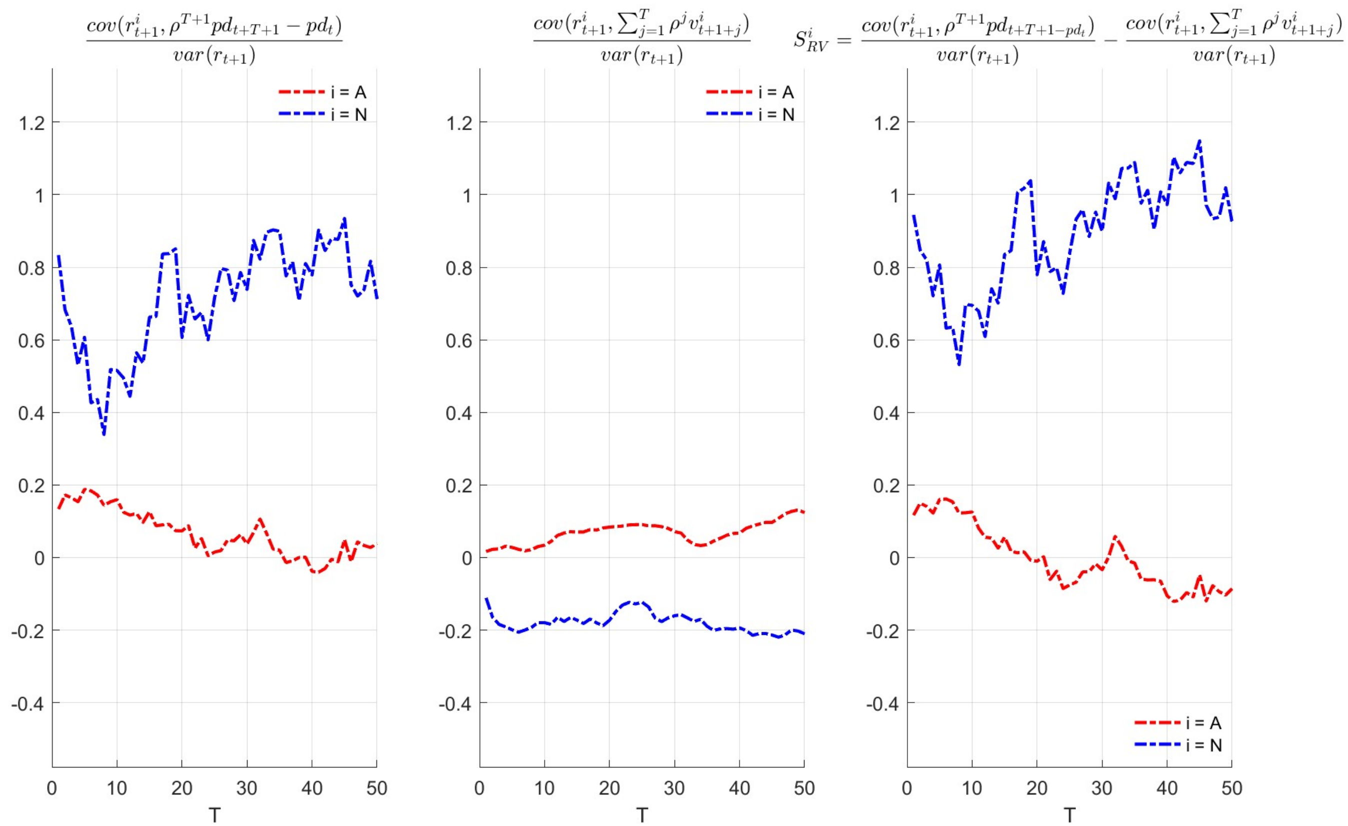

Figure 4: **Residual variance of the CRSP index – A-day and N-day components**

The figure shows the contribution of quarterly A-day and N-day returns to the residual volatility of aggregate quarterly returns and its components as defined in equations 12 and 13 for various values of T. The returns and dividends correspond to the CRSP value-weighted index. Returns and dividends data cover the period between 1964 and 2022.

Tables

Table 1: Variables definition

This table presents definitions of variables used to forecast stock market returns on A-days and N-days and the source of the data used. Most variables and content of this table are courtesy of Amit Goyal and co-authors.

OurName	**Descr**	**Authors**	**Freq**
GWZ_Q_pce	consumption/trend	Atanasov, Møller, Priestley	Quarterly
GWZ_M_vp	variance premium	Baekert, Hoerova	Monthly
GWZ_M_impvar	implied σ^2	Bakshi, Panayotov, Skoulakis	Monthly
GWZ_M_vrp	σ^2 risk premium	Bollerslev, Tauchen, Zhou	Monthly
GWZ_Q_govik	public sector investmt	Belo, Yu	Quarterly
GWZ_M_lzrt	9 illiq measures	Chen, Eaton, Paye	Monthly
GWZ_S_skew	skewness	Colacito, Ghysels, Meng, Siwasarit	Semiannual
GWZ_Q_crdstd	credit standards	Chava, Gallmeyer, Park	Quarterly
GWZ_M_ogap	prdctn-output gap	Cooper, Priestley	Monthly
GWZ_M_wtexas	oil price changes	Driesprong, Jacobsen, Maat	Monthly
GWZ_A_accrul	accruals	Hirshleifer, Hou, Teoh	Annual
GWZ_A_cfacc	accruals (CFO)	Hirshleifer, Hou, Teoh	Annual
GWZ_M_sntm	distilled sentiment	Huang, Jiang, Tu, Zhou	Monthly

Continued on next page

Table 1: – continued from previous page

OurName	Descr	Authors	Freq
GWZ_M_ndrbl	new order-ship durables	Jones, Tuzel	Monthly
GWZ_M_skvw	avg stock skewness	Jondeau, Zhang, Zhu	Monthly
GWZ_M_tail	x-sect tail risk	Kelly, Jiang	Monthly
GWZ_M_fbm	b/m x-sect factor	Kelly, Pruitt	Monthly
GWZ_M_dtoy	to Dow 52-week high	Li, Yu	Monthly
GWZ_M_dtoat	to Dow all-time high	Li, Yu	Monthly
GWZ_M_ygap	stock-bond yield gap	Maio	Monthly
GWZ_M_rdsp	stock return dispersion	Maio	Monthly
GWZ_M_svix	scaled risk-neutral vix	Martin	Monthly
GWZ_A_gpce	yearend econ growth	Møller, Rangvid	Annual
GWZ_A_gip	yearend econ growth	Møller, Rangvid	Annual
GWZ_M_tchi	14 technical indicators	Neely, Rapach, Tu, Zhou	Monthly
epbound_M3	low. bound on 3m exp. r. premium	Martin	Monthly
GWZ_A_house	housing/consumption	Piazzesi, Schneider, Tuzel	Annual
epbound_M6	low. bound on 6m exp. r. premium	Martin	Monthly
GWZ_M_avgcor	acvg corr stock returns	Pollett, Wilson	Monthly

Continued on next page

Table 1: – continued from previous page

OurName	Descr	Authors	Freq
epbound_M12	low. bound on 12m exp. r. premium	Martin	Monthly
GWZ_M_shtint	short interest	Rapach, Ringgenberg, Zhou	Monthly
GWZ_M_disag	analyst disagreement	Yu	Monthly
logPD	dividend price ratio	Campbell, Shiller	Monthly
logSP500e12p	earnings price ratio	Campbell, Shiller	Monthly
logSP500d12e12	dividend payout	Campbell, Shiller	Monthly
svar	σ^2	Guo	Monthly
bm	b/m	Kothari, Shanken	Monthly
ntis	net equity issuance	Boudoukh, Michaely, Richardson, Roberts	Monthly
BW_eqis	pct equity issuance	Baker, Wurgler	Annual
tbl	t-bill	Campbell	Monthly
lty	long govt yield	Fama, French	Monthly
ltr	long govt return	Fama, French	Monthly
tms	term spread	Fama, French	Monthly
dfy	default yield spread	Fama, French	Monthly
dfr	default return spread	Fama, French	Monthly

Continued on next page

Table 1: – continued from previous page

OurName	Descr	Authors	Freq
inflscsp	inflation	Fama, Schwert	Monthly
ik	invstmt/capital	Cochrane	Quarterly
cayGW	cnsm, wlth, incm	Lettau, Ludvigson	Quarterly

Table 2: Summary Statistics – 1953 through 2021

The table reports summary statistics of variables analysed. Panel A provides summary statistics of quarterly returns, quarterly returns accrued on A-days, and quarterly returns accrued on N-days. Panel B (C) reports summary statistics of variables that have been found to be A-day (N-day) returns' predictors in univariate linear regressions using quarterly data. Panel D (E) reports summary statistics of variables which have been found to predict returns accrued on both (neither) A-days and (nor) N-days. For each of the variables we report the length of its time series and the first and last month for which the data is available. We then report the summary statistics of the relevant time series: mean, standard deviation, minimum value, maximum value, first order autocorrelation, and skewness

Variable	N	Start	End	Mean	SD	Min	Max	Autocorr	skew
Panel A: Returns									
r	276	19530331	20211231	0.02	0.08	−0.30	0.22	0.04	−0.77
rA	276	19530331	20211231	0.01	0.03	−0.18	0.13	−0.02	−0.64
rNA	276	19530331	20211231	0.01	0.07	−0.28	0.18	0.10	−0.75
Panel B: A-day return predictors									
dfy	276	19530331	20211231	0.01	0.00	0.00	0.03	0.87	1.87
svar	276	19530331	20211231	0.01	0.01	0.00	0.11	0.38	6.74
GWZ_M_wtexas	276	19530331	20211231	0.01	0.08	−0.54	0.45	0.03	−0.41
GWZ_M_dtoy	276	19530331	20211231	0.93	0.08	0.58	1.00	0.64	−1.75

Continued on next page

Table 2: – continued from previous page

Variable	N	Start	End	Mean	SD	Min	Max	Autocorr	skew
Panel C: N-day return predictors									
logPD	275	19530630	20211231	4.96	0.39	4.09	5.95	0.94	0.24
lty	276	19530331	20211231	0.06	0.03	0.01	0.15	0.98	0.74
tbl	276	19530331	20211231	0.04	0.03	0.00	0.15	0.95	0.88
tms	276	19530331	20211231	0.02	0.01	−0.04	0.05	0.84	−0.15
ltr	276	19530331	20211231	0.02	0.05	−0.15	0.24	−0.05	0.89
ik	276	19530331	20211231	0.04	0.00	0.03	0.04	0.97	0.44
GWZ_M_ogap	276	19530331	20211231	0.00	0.07	−0.16	0.14	0.95	0.08
GWZ_Q_pce	273	19531231	20211231	0.00	0.04	−0.11	0.08	0.94	−0.20
GWZ_S_skew	275	19530630	20211231	−0.21	0.61	−1.29	1.28	0.50	0.43
GWZ_A_gpce	273	19531231	20211231	0.00	0.00	−0.01	0.02	0.78	0.03
GWZ_A_gip	273	19531231	20211231	0.01	0.02	−0.05	0.05	0.70	−0.90
Panel D: Predictors of returns on both types of days									
GWZ_M_dtoat	276	19530331	20211231	0.90	0.10	0.54	1.00	0.79	−1.15

Continued on next page

Table 2: – continued from previous page

Variable	N	Start	End	Mean	SD	Min	Max	Autocorr	skew
GWZ_M_avgcor	276	19530331	20211231	0.27	0.11	0.03	0.71	0.53	1.00
Panel E: Variables which do not predict returns									
cayGW	276	19530331	20211231	0.00	0.04	−0.28	0.05	0.86	−3.33
ntis	276	19530331	20211231	0.01	0.02	−0.05	0.05	0.94	−0.76
dfr	276	19530331	20211231	0.00	0.03	−0.15	0.16	−0.13	−0.52
bm	276	19530331	20211231	0.50	0.25	0.13	1.20	0.98	0.75
logSP500d12e12	276	19530331	20211231	−0.74	0.30	−1.24	1.38	0.89	2.81
logSP500e12p	276	19530331	20211231	−2.84	0.42	−4.81	−1.90	0.94	−0.72
GWZ_M_lzrt	276	19530331	20211231	−1.76	0.35	−4.69	−1.20	0.71	−4.40
GWZ_M_skvw	276	19530331	20211231	0.03	0.05	−0.38	0.16	−0.01	−2.56
GWZ_M_tail	276	19530331	20211231	0.42	0.05	0.30	0.53	0.90	−0.66
GWZ_M_fbm	276	19530331	20211231	0.17	0.11	−0.10	0.62	0.85	1.10
GWZ_M_ygap	275	19530630	20211231	−2.90	0.41	−4.84	−2.02	0.93	−0.83
GWZ_M_rdsp	276	19530331	20211231	0.03	0.01	0.01	0.12	0.66	3.60
GWZ_M_tchi	276	19530331	20211231	−0.02	1.45	−2.68	1.06	0.60	−0.94

Continued on next page

Table 2: – continued from previous page

Variable	N	Start	End	Mean	SD	Min	Max	Autocorr	skew
GWZ_Q_govik	276	19530331	20211231	0.03	0.01	0.03	0.06	0.97	1.09
GWZ_A_house	273	19531231	20211231	−0.25	0.01	−0.26	−0.22	0.95	0.37

Table 3: Univariate regressions - predicting quarterly returns

This table reports the results of univariate regressions of quarterly returns on predictor variables. All regressions are of the following type:

$$r^i_{t+1} = \alpha + \beta x_t, \quad i = (A\&N, A, N),$$

where r^i_{t+1} is the quarterly return (columns (2) through (4)), quarterly return accrued on A-days (columns (5) through (7)), or quarterly return accrued on N-days (columns (8) through (10)). In the interest of readibility we only report the values of the β coefficient, it's t-statistics, and the R^2 for each of the regressions. Panel A (B) reports regression results for variables that have been found to be A-day (N-day) returns' predictors. Panel C (D) reports regression results for variables which have been found to predict returns accrued on both (neither) A-days and (nor) N-days. Variables are summarized in Table 1. Period covered: 1953 – 2021.

	Quarterly returns			A-day quarterly returns			N-day quarterly returns		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variable	β	$t(\beta)$	R^2	β	$t(\beta)$	R^2	β	$t(\beta)$	R^2

Continued on next page

Table 3: – continued from previous page

Variable	Quarterly returns β	Quarterly returns $t(\beta)$	Quarterly returns R^2	A-day quarterly returns β	A-day quarterly returns $t(\beta)$	A-day quarterly returns R^2	N-day quarterly returns β	N-day quarterly returns $t(\beta)$	N-day quarterly returns R^2
Panel A: Announcement day return predictors									
dfy	1.36	[1.18]	0.1%	**0.79**	**[1.90]**	0.9%	0.57	[0.55]	−0.3 %
svar	**0.90**	**[1.94]**	1.0%	**0.58**	**[3.54]**	4.0%	0.32	[0.76]	−0.2 %
GWZ_M_wtexas	0.02	[0.33]	−0.3 %	**0.04**	**[1.77]**	0.8%	−0.02	[−0.35]	−0.3 %
GWZ_M_dtoy	**-0.14**	**[-2.05]**	1.2%	**-0.06**	**[-2.42]**	1.7%	−0.08	[−1.32]	0.3%
Panel B: Non-announcement day return predictors									
logPD	−0.02	[−1.42]	0.4%	0.00	[0.45]	−0.3 %	**-0.02**	**[-1.78]**	0.8%
lty	**-0.30**	**[-1.74]**	0.7%	0.01	[0.19]	−0.4 %	**-0.31**	**[-2.03]**	1.1%
tbl	**-0.38**	**[-2.39]**	1.7%	−0.01	[−0.22]	−0.3 %	**-0.36**	**[-2.59]**	2.0%
tms	**0.63**	**[1.78]**	0.8%	0.11	[0.87]	−0.1 %	**0.52**	**[1.64]**	0.6%
ltr	**0.22**	**[2.40]**	1.7%	0.01	[0.30]	−0.3 %	**0.21**	**[2.57]**	2.0%
ik	**-5.31**	**[-3.34]**	3.6%	0.17	[0.29]	−0.3 %	**-5.48**	**[-3.89]**	4.9%
GWZ_M_ogap	**-0.31**	**[-4.25]**	5.9%	0.00	[−0.17]	−0.4 %	**-0.30**	**[-4.73]**	7.2%
GWZ_Q_pce	**-0.47**	**[-3.49]**	4.0%	−0.06	[−1.11]	0.1%	**-0.41**	**[-3.45]**	3.9%

Continued on next page

Table 3: – continued from previous page

	Quarterly returns			A-day quarterly returns			N-day quarterly returns		
Variable	β	$t(\beta)$	R^2	β	$t(\beta)$	R^2	β	$t(\beta)$	R^2
GWZ_S_skew	−0.01	[−1.39]	0.3%	0.00	[0.70]	−0.2 %	**-0.01**	**[-1.84]**	0.9%
GWZ_A_gpce	**-3.62**	**[-3.57]**	4.2%	−0.28	[−0.76]	−0.2 %	**-3.34**	**[-3.70]**	4.5%
GWZ_A_gip	**-0.91**	**[-3.49]**	4.0%	−0.03	[−0.32]	−0.3 %	**-0.88**	**[-3.80]**	4.7%
Panel C: Announcement and non-announcement day return predictors									
GWZ_M_dtoat	**-0.15**	**[-2.87]**	2.6%	**-0.03**	**[-1.77]**	0.8%	**-0.11**	**[-2.49]**	1.9%
GWZ_M_avgcor	**0.18**	**[4.17]**	5.6%	**0.05**	**[3.00]**	2.8%	**0.13**	**[3.41]**	3.7%
Panel D: Variables not predicting either announcement or non-announcement returns									
cayGW	−0.03	[−0.23]	−0.35 %	0.04	[0.92]	−0.06 %	−0.08	[−0.63]	−0.22 %
ntis	−0.25	[−0.99]	−0.01 %	−0.04	[−0.39]	−0.31 %	−0.22	[−0.95]	−0.04 %
dfr	0.08	[0.40]	−0.31 %	−0.04	[−0.54]	−0.26 %	0.12	[0.67]	−0.20 %
bm	0.01	[0.26]	−0.34 %	0.00	[−0.59]	−0.24 %	0.01	[0.53]	−0.26 %
logSP500d12e12	0.02	[1.25]	0.20%	0.01	[1.62]	0.59%	0.01	[0.74]	−0.16 %
logSP500e12p	0.00	[0.26]	−0.34 %	−0.01	[−1.34]	0.29%	0.01	[0.83]	−0.11 %
GWZ_M_lzrt	−0.01	[−0.51]	−0.27 %	−0.01	[−0.91]	−0.06 %	0.00	[−0.20]	−0.35 %
GWZ_M_skvw	−0.06	[−0.61]	−0.23 %	−0.01	[−0.14]	−0.36 %	−0.06	[−0.62]	−0.22 %

Continued on next page

Table 3: – continued from previous page

Variable	Quarterly returns			A-day quarterly returns			N-day quarterly returns		
	β	$t(\beta)$	R^2	β	$t(\beta)$	R^2	β	$t(\beta)$	R^2
GWZ_M_tail	0.13	[1.41]	0.36%	0.00	[0.12]	−0.36 %	0.13	[1.53]	0.49%
GWZ_M_fbm	0.02	[0.38]	−0.31 %	0.02	[1.06]	0.04%	0.00	[−0.01]	−0.37 %
GWZ_M_ygap	0.01	[0.44]	−0.30 %	−0.01	[−1.37]	0.32%	0.01	[1.05]	0.04%
GWZ_M_rdsp	−0.06	[−0.15]	−0.36 %	0.18	[1.21]	0.17%	−0.24	[−0.66]	−0.20 %
GWZ_M_tchi	0.00	[0.46]	−0.29 %	0.00	[−0.67]	−0.20 %	0.00	[0.79]	−0.14 %
GWZ_Q_govik	0.24	[0.33]	−0.33 %	−0.16	[−0.59]	−0.24 %	0.40	[0.61]	−0.23 %
GWZ_A_house	0.60	[1.12]	0.10%	−0.11	[−0.56]	−0.25 %	0.71	[1.49]	0.45%

Table 4: **Forecasting changes in fundamental value using real ex-post rational price P^*_{t+1} as its proxy**

This table reports the results of regressing the change in real ex-post rational price $\widetilde{P}^*_{t+1}$ on lagged returns accrued on A-days and N-days at various frequencies. The left-hand side of the table presents results of estimating equation (4): $Y_{t+1} = \beta_0 + \beta_1 r^A_t + \beta_2 r^N_t$. The right-hand side of the table presents results of estimating equation (5): $Y_{t+1} = \delta_0 + \delta_1(r^A_t + r^N_t) + \delta_2(r^A_t - r^N_t)$. In both cases $Y_{t+1} = \widetilde{P}^*_{t+1}$. For each of the equations two time periods are considered: 1953 – 2022 and 1953 – 2010. This is to account for the fact that starting from January 2011 the terminal value of P^*_T is becoming non-negligible share of P^*_t. t-statistics are reported in the second row below each coefficient value. Since Y_{t+1} are constructed using non-overlapping windows, standard unadjusted t-statistics are reported.

Eq.(4): $Y_{t+1} = \beta_0 + \beta_1 r^A_t + \beta_2 r^N_t$					Eq.(5): $Y_{t+1} = \delta_0 + \delta_1(r^A_t + r^N_t) + \delta_2(r^A_t - r^N_t)$				
β_0	β_1	β_2	$Adj.R^2$	N	δ_0	δ_1	δ_2	$Adj.R^2$	N
Panel A: Monthly data 1953 – 2022					Panel A: Monthly data 1953 – 2022				
0.003	0.002	-0.002	0.00%	839	**0.003**	0.000	0.002	0.00%	839
[39.96]	[0.60]	[-1.31]			**[39.96]**	[0.03]	[1.05]		
Panel B: Monthly data 1953 – 2010					Panel B: Monthly data 1953 – 2010				
0.003	0.004	-0.003	0.08%	695	**0.003**	0.001	0.003	0.08%	695
[31.49]	[0.85]	[-1.37]			**[31.49]**	[0.26]	[1.30]		
Panel C: Quarterly data 1953 – 2022					Panel C: Quarterly data 1953 – 2022				
0.008	0.007	0.000	-0.32%	279	**0.008**	0.003	0.004	-0.32%	279
[40.62]	[1.05]	[-0.13]			**[40.62]**	[0.98]	[0.98]		
Panel D: Quarterly data 1953 – 2010					Panel D: Quarterly data 1953 – 2010				
0.008	**0.014**	-0.001	0.61%	231	**0.008**	0.006	**0.008**	0.61%	231
[34.61]	**[1.81]**	[-0.53]			**[34.61]**	[1.56]	**[1.82]**		
Panel E: Annual data 1953 – 2022					Panel E: Annual data 1953 – 2022				
0.033	0.032	-0.005	-0.67%	69	**0.033**	0.013	0.018	-0.67%	69
[18.34]	[1.17]	[-0.46]			**[18.34]**	[0.93]	[1.24]		
Panel F: Annual data 1953 – 2010					Panel F: Annual data 1953 – 2010				
0.030	**0.046**	-0.012	3.72%	57	**0.030**	0.017	**0.029**	3.72%	57
[16.02]	**[1.70]**	[-1.16]			**[16.02]**	[1.19]	**[1.99]**		

Table 5: **Forecasting changes in fundamental value using *detrended* real ex-post rational price p^*_{t+1} as its proxy**

This table reports the results of regressing the change in *detrended* real ex-post rational price $\widetilde{p}^*_{t+1}$ on lagged returns on A-days and N-days at various frequencies. The left-hand side of the table presents results of estimating equation (4): $Y_{t+1} = \beta_0 + \beta_1 r^A_t + \beta_2 r^N_t$. The right-hand side of the table presents results of estimating equation (5): $Y_{t+1} = \delta_0 + \delta_1(r^A_t + r^N_t) + \delta_2(r^A_t - r^N_t)$. In both cases $Y_{t+1} = \widetilde{p}^*_{t+1}$. For each equation two time periods are considered: 1953 – 2022 and 1953 – 2010. This is to account for the fact that starting from January 2011 the terminal value of p^*_T is becoming non-negligible share of p^*_t. t-statistics are reported in the second row below each coefficient value. Since Y_{t+1} are calculated using non-rolling windows, standard unadjusted t-statistics are reported.

Eq.(4): $Y_{t+1} = \beta_0 + \beta_1 r^A_t + \beta_2 r^N_t$					Eq.(5): $Y_{t+1} = \delta_0 + \delta_1(r^A_t + r^N_t) + \delta_2(r^A_t - r^N_t)$				
β_0	β_1	β_2	$Adj.R^2$	N	δ_0	δ_1	δ_2	$Adj.R^2$	N
Panel A: Monthly data 1953 – 2022					Panel A: Monthly data 1953 – 2022				
0.000	0.002	-0.002	0.05%	839	**0.000**	0.000	0.002	0.05%	839
[-2.66]	[0.66]	[-1.46]			**[-2.66]**	[0.03]	[1.16]		
Panel B: Monthly data 1953 – 2010					Panel B: Monthly data 1953 – 2010				
0.000	0.003	-0.002	0.07%	695	**0.000**	0.001	0.003	0.07%	695
[-4.34]	[0.83]	[-1.36]			**[-4.34]**	[0.24]	[1.28]		
Panel C: Quarterly data 1953 – 2022					Panel C: Quarterly data 1953 – 2022				
0.000	0.006	0.000	-0.21%	279	**0.000**	0.003	0.003	-0.21%	279
[-3.25]	[1.19]	[-0.15]			**[-3.25]**	[1.11]	[1.11]		
Panel D: Quarterly data 1953 – 2010					Panel D: Quarterly data 1953 – 2010				
-0.001	**0.010**	-0.001	0.56%	231	**-0.001**	0.005	**0.006**	0.56%	231
[-5.47]	**[1.80]**	[-0.46]			**[-5.47]**	[1.57]	**[1.78]**		
Panel E: Annual data 1953 – 2022					Panel E: Annual data 1953 – 2022				
-0.002	0.025	-0.005	0.11%	69	**-0.002**	0.010	0.015	0.11%	69
[-1.66]	[1.29]	[-0.67]			**[-1.66]**	[0.98]	[1.43]		
Panel F: Annual data 1953 – 2010					Panel F: Annual data 1953 – 2010				
-0.004	**0.035**	-0.009	3.69%	57	**-0.004**	0.013	**0.022**	3.69%	57
[-2.81]	**[1.67]**	[-1.20]			**[-2.81]**	[1.14]	**[1.97]**		