Correlation neglect in asset prices*

Hongye Guo† Jessica A. Wachter‡

May 16, 2025

Abstract

The U.S. stock market's return during the first month of a quarter positively predicts the second month's return, which then negatively predicts the first month's return of the next quarter. The pattern arises from a model in which investors do not fully recognize that earnings announced in the second month of a quarter are inherently similar to those announced in the first month, thereby overreacting to such predictably repetitive earnings. The same pattern exists in the cross-section and time series of industry returns. Evidence from survey data lends support to the mechanism of correlation neglect.

Keywords: behavioral finance, correlation neglect, earnings announcements, efficient market hypothesis, stock return autocorrelation

JEL codes: G12, G14, G40

*The authors do not have conflicts to disclose We gratefully acknowledge research support from National Institutes of Health Grant MH55687.

†University of Hong Kong. Email: hoguo@hku.hk.

‡Department of Finance, The Wharton School, University of Pennsylvania and the National Bureau of Economic Research. Email: jwachter@wharton.upenn.edu. .

1 Introduction

Predicting stock market returns based on past returns is challenging. Research by Kendall and Hill (1953) and Fama (1965) from decades ago revealed that there is minimal serial correlation in stock returns. Fama (1970) introduced the concept of the weak form of the efficient market hypothesis, postulating that prices already incorporate all available information from past prices. This hypothesis appears particularly applicable to the U.S. aggregate market, as demonstrated by Poterba and Summers (1988), who found that monthly market returns in the U.S. exhibit a small, statistically insignificant positive autocorrelation over a 12-month period.[1] The lack of correlation between past and future market returns in the U.S. is not only a well-recognized statistical phenomenon, but also an emblem of market efficiency.

We document that the U.S. market return during the first month of a quarter positively predicts the second month's return, which then negatively predicts the first month's return of the next quarter. These first months of a quarter are special and important to investors because they contain fresh earnings news about the aggregate economy. This is owing to the nature of the earnings cycle in the United States. Take January as an example. At the end of the December, firms close their books for Q4, and they announce Q4 earnings in January, February, and March. January therefore contains the early earnings announcements and is the first time that investors learn about the economy's performance in Q4. The first months of the quarters are famously known as the "earnings seasons" among practitioners and receive heightened attention. Throughout this paper, we refer to them as "newsy" earnings months, as they produce fresh news about aggregate earnings.

The second month of each quarter contains about equally many earnings announcements as does a newsy month. The aggregate earnings announced in the second month are, however, predictably repetitive of those announced in the preceding newsy month because the

[1]Furthermore, Poterba and Summers (1988) noted a negative autocorrelation in the U.S. market over longer time horizons, while Lo and MacKinlay (1988) identified a positive autocorrelation in weekly returns for small U.S. stocks. Despite these variations, the consistently low month-to-month autocorrelation in U.S. market returns remains a crucial piece of evidence supporting the concept of market efficiency.

announced earnings are of the same fiscal quarter (Q4 in our example above). We call these second months "repetitive" earnings months, as they contain earnings announcements that produce repetitive aggregate earnings "news." Such repetition is not verbatim—in fact, different firms announce in the first and the second month of a quarter—but rather subtle: the earnings announced in the second month broadly resemble those in the first month, owing to their common aggregate component. The third month of a quarter contains substantially fewer earnings announcements and produce muted earnings news. The terminology allows us to compactly describe our return pattern in one sentence: a newsy month's return positively predicts the next repetitive month's return, which then negatively predicts the next newsy month's return.

One way to test our mechanism is to examine earnings expectations themselves. Specifically, our hypothesis states that imperfectly rational investors use announced earnings to forecast future earnings, but fail to fully account for the repetitive nature of the earnings in the second month of a quarter. This act of correlation neglect leads to overreaction in the repetitive months that are predictably in the same direction as the news in the preceding newsy month, giving rise to the newsy-to-repetitive return continuation. As the next newsy month comes, these overreactions are correctly, leading to the repetitive-to-newsy return reversal. If this is accurate, we would see it not only in the return data, but also in the expectations data.

We evaluate this prediction using survey data from IBES. We first show that sell-side analysts indeed exhibit correlation neglect, leading to different types of predictability in earnings growth revisions in the newsy and repetitive months. This evidence is uniquely useful because survey data are direct measures of market expectations, albeit imperfect. Moreover, we show that the predictable reversals in market returns do not occur until the earnings seasons begin. This alignment in timing suggests that it is indeed the earnings seasons that drive the return predictability, as opposed to other quarterly fluctuations unrelated to earnings announcements. As a further robustness chec, we find a similar result in the cross-section: Industry returns in excess of the market predict returns in repetitive months with a positive sign and in the next newsy month with a negative sign and in the first month of the following quarter with a positive sign.

We formalize the correlation neglect mechanism with a model in which there is variability in the second month regarding whether the information from the first is brought to investors' attention or not. A fully rational investor would not react to this "news," understanding that the new information was correlated with the old. An investor who believes he or she is seeing a genuinely new observation reacts; this reaction is predictable because the information is in fact the same as in the first quarter, and it also causes stock prices to rise or fall by more than is warranted.

These findings have broader implications. First, we provide evidence of return predictability that is robust, and economically and statistically significant, with monthly out-of-sample R^2 statistics of 4% , thus contributing to the debate on whether return predictability actually exists (Goyal and Welch 2008, Campbell and Thompson 2008,Goyal et al. 2024).[2] Second, we show that three important and commonly made assumptions about stock market returns are not innocuous: i) the expected stock market return is persistent, ii) the expected stock market return almost always exceeds the risk-free rate, and iii) the stock market return is covariance-stationary. Finally, we provide asset pricing and survey evidence for correlation neglect: investors seem to treat repeated information as new independent observations, even if they have substantial incentive to do otherwise.

Besides the very substantial literature on continuation and reversal, several recent papers make points that are related to ours. First, the "tug-of-war" pattern between continuation and reversal we document for monthly stock market returns is mirrored in intraday market returns: Lou et al. (2024) show that open-to-close stock market returns positively predict future close-to-open stock market returns, which then negatively predict future open-to-close returns.[3] The authors attribute their findings to clientele effects: the overreaction arises from news being repeatedly consumed by a second clientele (households). In our paper, the

[2] In contrast, Goyal et al. (2024) find out-of-sample R^2 statistics, should they be positive, tend to be in the tens of basis points, a finding that we replicate using our methodology.

[3] Relatedly, Lou et al. (2019) find that close-to-open (open-to-close) stock returns positively predict future close-to-open (open-to-close) returns, and negatively predict future open-to-close (close-to-open) returns.

overreaction aries from news itself being repeated in the second month of a quarter. Second, Fedyk and Hodson (2023) document that repetitive "news" on Bloomberg terminals generates overreactions in returns of related stocks, and that such overreaction is stronger when the repetition is more subtle, a result that, like ours, suggests correlation neglect. Finally, Enke and Zimmermann (2019) find that human subjects do not fully discount correlated news in experiments even when they are given a simple setup and clear instructions. We contribute to this literature by showing that the correlation neglect mechanism further exists in aggregate stock market returns and the cross section of industry returns.

2 Predicting aggregate stock market returns

We first characterize the U.S. earnings cycle, which originates from i) aligned fiscal quarters and ii) heterogeneous announcement lags. Panel A of Table 1 shows that the vast majority of the US firms close their books for a given fiscal quarter on the last day of a calendar quarter. Therefore, the fiscal quarters over which the earnings are recorded are effectively the calendar quarters. While the fiscal quarters are aligned, firms do not announce their quarterly earnings together. Almost all firms announce within the three months after the fiscal quarter end date, but there is much variation in their announcement lags. Panel B shows that among the on-cycle announcements, about 45% occur in the first month of a quarter, about 48% in the second month, and about 8% in the last month.

The quarterly cycles of earnings announcements implies that each of the three months in the quarter has a different status. The first month is the first time that investors learn about the *aggregate* economy's performance in the previous quarter. This economy-level learning is possible because corporate earnings have a common time component: the aggregate economic condition in fiscal quarter 4 (FQ4) of 2023 influences all firms, and consequently, if early announcers' FQ4 earnings are good, then FQ4 is more likely than not a good quarter for everyone. Earnings announced in the first month of a quarter therefore provide new information about aggregate earnings earned in the previous quarter. We refer to these first months of quarters as the newsy earnings months, or newsy months for convenience.

The second months, in contrast, see earnings that are predictably repetitive of those announced in the first months. This is again owing to the shared aggregate component of corporate earnings: earnings announced in February, like those announced in January, are earned in Q4 of the previous year and therefore inherently similar. We refer to the second months of quarters as the repetitive earnings months, or simply repetitive months. The third months of quarters also contain announcements of predictably repetitive earnings. However, as they contain substantially fewer on-cycle earnings announcements, the repetition is muted.

2.1 Baseline return pattern

We demonstrate our baseline predictability pattern in the U.S. stock market returns with the following monthly time-series regression:

$$mkt_t = \alpha + \beta mkt_{nr(t)} + \epsilon_t. \tag{1}$$

Here, mkt_t is the U.S. stock market return in excess of the risk free rate in month t, and $nr(t)$ is the most recent month before t that is newsy or repetitive. For example, if t is February, then $nr(t)$ is January, and if t is March or April, $nr(t)$ is February.

Panel A, Column 1 of Table 2 conducts regression 1 on the subsample in which the dependent variable month t is repetitive. The positive coefficient of 0.279 implies that a newsy month's return positively predicts the return in the subsequent repetitive month. Column 2 is on the subsample in which t is newsy. The negative coefficient of -0.279 shows that a repetitive month's return negatively predicts the return in the next newsy month. These values are economically sizable, as they are bounded by -1 and 1 in stationary time series.

Column 3 extracts the difference of the coefficients in Column 1 and 2 by conducting the following regression on the combined sample in which t is either newsy or repetitive:

$$mkt_t = \alpha + \beta_1 mkt_{nr(t)} + \beta_2 mkt_{nr(t)} \times I_t^n + \gamma I_t^n + \epsilon_t. \tag{2}$$

Here, I_t^n is a dummy variable indicating whether month t is newsy. The most important number in Table 2 is the coefficient on the interaction term, β_2. Its value of -0.557 captures the difference in Columns 1 and 2 and emblematizes the table. Economically, it demonstrates that monthly autocorrelation of the U.S. stock market returns varies strongly and predictably. Without the third months of quarters that contain muted earnings news, the first order autocorrelation of the monthly U.S. stock market returns alternates between strongly positive (newsy predicting repetitive) and strongly negative (repetitive predicting newsy). The associated t-statistic of -4.35 implies that the observed relation is unlikely to arise from pure noise or pure data-mining.[4] While this statistic is an useful diagnostic, the time-series return forecasting literature (e.g., Goyal and Welch (2008)) has critiqued and moved away from it. We examine more standard performance metrics such as out-of-sample R^2 and trading strategy behaviors in the next subsection.

Panel B, C, and D of Table 2 report results of the same regressions as in Panel A, except that they are on the first half, second half, and the post-WWII portion of the full sample, respectively. The pattern operates in each of these episodes and is not the sole result of a small number of observations, warranting further evaluation of its reliability.

2.2 Further evaluation

We now evaluate the baseline pattern from three additional aspects: i) out-of-sample forecasting performance, ii) trading strategy performance, and iii) and contribution to the existing literature.

[4] It further implies that an in-sample, market-neutral trading strategy that bets on both the continuation in Column 1 and the reversal in Column 2 approximately yields an annualized Sharpe ratio of $4.35/\sqrt{98} = 0.44$, where the number 98 is the number of years in the sample. This back-of-envelope calculation approximates the market excess returns with the regression residuals (which works in low-R^2 settings) and is merely a convenient heuristic. We formally construct a trading strategy in the next subsection.

Forecasting performance Goyal and Welch (2008) made an important critique on the literature studying stock market return predictability, which is that in a linear regression model, any proposed signals to forecast returns can only be used with estimated coefficients, and if such estimation is done on data available in real time, the signals often generate negative "out-of-sample" (OOS) R^2. This means that such signals, when employed to predict the stock market returns in real time, often fail to outperform the returns' historical average.

Table 3 presents the out-of-sample R^2 generated by our signal. Following Campbell and Thompson (2008), we extend the CRSP aggregate market return series to 1872 using GFD data, and evaluate the R^2 from 1926. The benchmark underlying these R^2s is the expanding-window mean excess return $\overline{mkt_t}$, the average stock market return in excess of the risk free rate up to month t. Our signal s_{t-1}, which forecasts mkt_t, is $mkt_{nr(t)} - \overline{mkt_{t-1}}$ for a repetitive t, $-(mkt_{nr(t)} - \overline{mkt_{t-1}})$ for a newsy t, and 0 otherwise.

Table 3 estimates the expanding-window OLS coefficients for our signal and compute the real-time forecasts for mkt_t in multiple ways. These methods are kept simple, with no constraint or shrinkage applied. Column 1 shows that even the most naive method—combining the signal values and the regression coefficients of historical returns on historical signals and a freely estimated constant—generates an R^2 of 4.20%. This number appears an order of magnitude higher than what are generated by other time-series signals for monthly US stock market returns surveyed in the comprehensive study of Goyal and Welch (2008), Rapach and Zhou (2022), and Goyal et al. (2024). This high R^2, along with those generated with methods 2 and 3, mitigates the concern that the signal implied by the time-series results in this paper could not have been successfully used to forecast stock market excess returns in real time.

Trading strategy Good forecasting performance does not necessarily translate into successful trading strategies.[5] We now use our signal to construct a beta-neutral market-timing strategy that rebalances positions in i) the stock market and ii) the risk-free bond and

[5]The converse of the statement is not true either.

examine its performance.

Overall, the strategy implements two rules with data available in real time. First, it places proportional weights on the market to the signal value, with a mean of zero. Second, the portfolio weight is scaled so that the strategy has a historical volatility of 5% per month. Specifically, for month t, the unscaled portfolio weight is simply s_{t-1}. The unscaled strategy volatility $\bar{\sigma}^s_{t-1}$ is the unscaled strategy's historical volatility up to $t-1$. The scaled portfolio weight is then $s_{t-1}\frac{0.05}{\bar{\sigma}^s_{t-1}}$. We further truncate the weight on the market at -3 and 3 to limit the impact of extreme observations.[6] The strategy's construction involves no look-ahead bias, and the scaling can be adjusted according to an investor's risk appetite.

Column 1 of Table 4 shows that this trading strategy yields a monthly return of 0.668% (i.e., 8.02% per annum). Columns 2-4 show that the CAPM alpha, Fama-French 3-factor alpha, and FF3 + Momentum alpha are 0.554%, 0.402%, and 0.715% per month respectively. Our time-series strategy adds value to these factors. Moreover, as it involves only positions on aggregate stock market and the risk-free bond, it is likely much cheaper to implement than the controlled cross sectional factors, which involve positions on individual stocks.

Relation to the existing literature Goyal et al. (2024) recently surveyed the literature on aggregate stock market return predictability and examined the signals' out-of-sample forecasting performance. Table 3 of their paper shows that among an expanded set of 34 predictors, the monthly in-sample R^2s are on average 0.32%, OOS R^2 are on average -0.44% after applying the regression constraints in Campbell and Thompson (2008). These results reaffirm the conclusion of Goyal and Welch (2008) that existing signals "predicted poorly both in-sample (IS) and out-of-sample (OOS)...seem unstable...and would not have helped an investor with access only to available information to profitably time the market."

We replicate and extend their results in Table B1. The vast majority of OOS R^2s in this table are negative, with the maximum being 0.60% even as multiple implementations of the regressions are attempted. This reaffirms the conclusion of Goyal and Welch (2008).

[6] Our results do not appreciably change if this step is dropped, although in practice, implausibly high leverage cannot be implemented and a similar constraint would exist explicitly or implicitly.

By documenting a predictor that generates an OOS R^2 that is an order of magnitude larger than the numbers in Table B1, we shift the literature's prior on an important question: How predictable is the U.S. monthly stock market returns?

Our paper also relates to Guo (2025), which also focuses on earnings seasons (newsy months in our language). Guo (2025) also shows evidence for subtle expectational errors around earnings data that lead to return predictability patterns. In his case, the error arises from a mechanism called parameter compression: investors assume that earnings announcements equally close in time are equally close in information content. However, earnings announcements in any given quarter are closer than those in separate quarters. More generally, the true relevance of a newsy month's earnings follows a step pattern, whereas the hypothesis in the paper is that it follows a straight line. This mechanism leads to simultaneous under- and overreaction in the newsy months. In contrast, in our paper, the mechanism is correlation neglect, and it leads to overreaction in the repetitive months.

The mechanisms are distinguishable in the data. Table 5 shows in monthly stock return regressions that they remain highly significant after controlling for one another. Besides the regressor $mkt_{nr(t)}$, namely the return in the previous month if the second month in the quarter (repetitive) or two months back if the first month in the quarter, the regressions also have the regressor $\sum_{j=1}^{4} mkt_{nm(t,j)}$, the key signal used in Guo (2025) that represents the average return over the last four newsy months. Both coefficients in column 3 are significantly negative, implying that correlation neglect retains its predictive power even when controlling for parameter compression.

Overall, our proposed signal featuring the tug-of-war between market return continuation and reversal appears to i) generate positive OOS R^2, ii) produce a profitable trading strategy, and iii) add value to the existing literature. It is therefore worthwhile to consider possible explanations for the pattern.

3 Explaining the return pattern

Before proposing an explanation, we note two telltale features of the data that inform us about the type of explanations that are likely to drive our return pattern.

First, the continuation and reversal arms appear tightly linked. We separate the return of trading strategy (as constructed in subsection 2.2) in each year into the component earned in the newsy months and that earned in the repetitive months. Table 6 regresses the two components on each other, both in the raw return unit and ranked to limit the influence of outliers. The significant positive coefficients imply that a year in which the continuation effect is strong is also likely to see strong reversal. This means that the continuation and reversal arms are likely driven by the same underlying economic force and should therefore be explained by one story rather than two.

Second, our signal implies expected market excess returns that are frequently negative. Figure 1 plots the expected stock market return implied by Method 2 of Table 3. This time series of expected stock market excess return is negative 24% of the months. This implies that an explanation based on risk alone is unlikely to explain the return pattern in full. In such a risk-based framework, the negative expected returns means that the stock market is safer than Treasury bills one month out of four. This feature is especially difficult to accommodate, as it is implausible that investors find the stock market to be safer than Treasury bills that often.

Given these two features, a parsimonious mechanism that jointly explain the continuation and reversal arms of the return pattern is correlation neglect: investors do not fully discount the inherently repetitive earnings announced in the second month of a quarter and thus overreact to them, creating the return continuation that gets corrected in the next newsy month. To formalize this intuition, we apply a model of correlation neglect, motivated by Enke and Zimmermann (2019) to asset pricing.[7]

[7]Our model differs from that of Enke and Zimmermann (2019) in several respects. In Appendix A, we show that directly applying the model of Enke and Zimmermann does not explain the effects in our paper, and give intuition as to why.

Assume that the economy is either in a high or low-productivity state. If the economy is in the high-productivity state, the aggregate market is worth 1 whereas it is worth 0 if it is in a low-productivity state. The high-productivity state occurs with probability p, which is unknown to investors. Investors have a prior distribution over p. The distribution is beta, such that they would estimate a probability p^* with a sample size of τ. To fix ideas and eliminate unnecessary notation, we set $\tau = 1$.

This model has three periods. Period 1 corresponds to the first month of a quarter, namely a newsy month. Period 2 is the repetitive month that follows. Period 3 is the first month of the following quarter. In period 1, investors receive a signal s, with $s \sim \text{Bernoulli}(p)$. In period 2, this signal is potentially repeated (as opposed to an independent signal being drawn from Bernoulli(p)). Whether or not the signal is repeated is determined by $x \sim \text{Bernoulli}(p_x)$. If $x = 1$, then s is brought again to investors' attention.

To focus on our main mechanism, we assume investors are risk-neutral and the riskfree rate is equal to zero. Under these conditions, the value of the market is equal to the posterior mean of p, which equals:

$$E[p|N \text{ out of } T] = \frac{N + p^*\tau + 1}{T + \tau + 2} \tag{3}$$

The investors' prior implies that the price prior to observing the signal in the first period is $P_0 = E[p] = (p^* + 1)/3$. After observing signal s, the mean shifts so that the price equals

$$P_1 = E[p|s] = \frac{s + p^* + 1}{4}.$$

For simplicity, we define returns as differences in prices rather than percent changes. The return $R_1 = P_1 - P_0$. Note that $E[R_1] = 0$.

Equilibrium with rational investors

There is no information and hence no change in the second period, $P_2 = P_1$ and $R_2 = 0$. The return in the third period R_3, conditional on both p and s is Bernoulli outcomes $1 - P_1$

with probability p and $-P_1$ with probability $1-p$. Note that

$$E[R_3|s] = E[E[R_3|p,s]|s] = E[[p|s] - P_1 = 0.$$

In addition,

$$E[R_1R_3] = E[E[R_1R_3|s]] = E[R_1E[R_3|s]] = 0$$

Because $E[R_1] = E[R_3] = 0$, this shows $\text{Cov}(R_1, R_3) = 0$. Because R_2 is a constant, it also has zero covariance with both random variables.

Equilibrium with correlation neglect

Investors form

$$P_2 = E^{\text{CN}}[p|x,s] = \begin{cases} \frac{2s+p^*+1}{5} & \text{for } x = 1 \\ \frac{s+p^*+1}{4} = P_1 & \text{for } x = 0 \end{cases}$$

Then

$$\begin{aligned} E[P_2] &= E[E[P_2|s]] \\ &= E\left[p_x\frac{2s+p^*+1}{5} + (1-p_x)P_1\right] \\ &= E\left[p_x\left(\frac{2s+p^*+1}{5} - P_1\right) + P_1\right] \\ &= p_xE\left[\frac{2s+p^*+1}{5} - \frac{s+p^*+1}{4}\right] + P_1 \\ &= P_1, \end{aligned}$$

where we use the fact that $E[s] = E[p] = (p^*+1)/3$. Therefore the investor is, on average, neither optimistic nor pessimistic.

However, unlike in the case with the rational agent, R_2 is positively correlated with R_1. Intuitively, the agent double-counts s in the second period, not realizing that it is the same signal as before. Because R_2 has mean zero, it suffices to show that $E[R_2R_1] > 0$. First, note that a useful intermediate result following from (4) is:

$$E[R_2|s] = p_x \left(\frac{3s - p^* - 1}{20} \right)$$

$$\begin{aligned} E[R_2 R_1] &= E[E[R_2 R_1|s]] \\ &= E[E[R_2|s] R_1] \\ &= E[E[R_2|s](P_1 - P_0)] \\ &= E\left[p_x \left(\frac{3s - p^* - 1}{20} \right) \left(\frac{3s - p^* - 1}{12} \right) \right] \\ &> 0 \end{aligned}$$

where the second line follows because R_1 is a constant in s.

Finally, we show that $E[R_3 R_2] < 0$, namely that returns in the second and third period are negatively correlated. Note that P_3 is identical in the equilibrium with correlation neglect and with rational investors. Also, in both economies, $E[P_3|s] = E[P_3|x, s] = P_1$. Finally, by definition, $P_3 - P_1 = P_3 - P_2 + P_2 - P_1$ and so $E[R_3 - R_2|x, s] = 0$. That is, the price in the first period is the optimal forecast. Therefore, any reaction to "news" in the second period will be an over-reaction that will on average be corrected in the final period:

$$\begin{aligned} E[R_3 R_2] &= E[E[R_3 R_2|x, s]] \\ &= E[E[R_3|x, s] R_2], \\ &= -E[R_2^2] < 0. \end{aligned}$$

where the second line follows because R_2 is a constant in x, s.

This model establishes continuation and reversal. It also makes several additional predictions. Our model is based on expectations driven by announcements. This leads to two additional hypotheses.

Prediction 1 (Survey data)**.** *Revisions in a repetitive month are likely to be in the same* *direction as those in the previous newsy month, and revisions in a newsy month are likely in the opposite direction to revisions in the previous repetitive month.*

Note that the expectation revision in a newsy month is $E^{\text{CN}}[p|s] - p^*$, which equals R_1. The revision in the repetitive month is $E^{\text{CN}}[p|s,x] - E^{\text{CN}}[p|s]$, which equals R_2. These are positively correlated as we show above. In the following newsy month, the expectation revision, if there is a positive outcome, is $1 - E^{\text{CN}}[p|s,x]$ and $-E^{\text{CN}}[p|s,x]$ otherwise. For the reasons described above, this is negatively correlated with the revision in the repetitive month.

Prediction 2 (Similarity of announcements)**.** *If continuation and over-reaction arise from announcements, the effect will be stronger when the announcement in the repetitive month is more similar to the previous newsy month.*

4 Testing the hypothesis

Having stated our hypothesis of correlation neglect, we further test it from four additional angles.

4.1 The pre-season periods

Our previous analysis primarily focused on monthly stock returns and utilized the concept of newsy months. However, the period with a significant number of earnings announcements does not immediately begin on the first day of these newsy months. Even the quickest announcers require some time to prepare their statements. Specifically, among companies with fiscal periods aligned with calendar quarters, only 0.27% of earnings announcements occur within the initial week (or the first five trading days) of the quarter. This percentage is significantly lower compared to the average weekly announcement rate of about 8%. In contrast, the second week of each quarter sees 2.93% of the announcements, which aligns more closely with the average weekly rate. Therefore, any market return reversals, if stemming from these earnings announcements, are not expected to occur in the first week of newsy months. Table 7 confirms this prediction. Columns 2 and 3 show a weak reversal in the first week of newsy months and a strong reversal in remaining periods, with the coefficients disproportionate to the length of the predicted returns (about one to three). This analysis suggests that our observed return pattern is indeed tied to the early earnings announcements. This piece of evidence speaks in favor of stories involving these earnings announcements and against those built upon other unrelated quarterly fluctuations.

4.2 Survey data

An influential branch of the behavioral finance literature uses survey data from IBES to measure investors' expectations of aggregate cash flow growth. In particular, Nagel and Xu (2022) and Bordalo et al. (2024) argue that the long-term growth (LTG) measure in IBES is what matters for aggregate asset price variations. As Prediction 1 states, our model traces the continuation and reversal effects to beliefs, which this branch of the literature has argued can be extracted using the LTG measure from the IBES Adjusted Details file.[8] Following Bordalo et al. (2024), we use medians across analysts to represent firm level consensus, and then aggregate firm level LTG revisions in each month using the market capitalization weight within the S&P 500 universe to obtain revisions in aggregate cash flow growth expectations. Our model predicts that investors overreact in the repetitive months in the sense that i) revisions in a repetitive month are likely in the same direction as those in the preceding newsy month and ii) revisions in a repetitive month are likely followed by correction in the opposite direction in the subsequent newsy month. These lead to two predictions on the time series of aggregate LTG revisions. First, revisions in a repetitive month are positively correlated with revisions in the previous newsy month. Second, revisions in a newsy month are negatively correlated with revisions in the previous repetitive month.

Column 1 of Table 8 tests prediction i) by running the following regression on the sample

[8] We do not use the Summary files, which contain middle-of-the-month consensuses struck on the Thursday after the 3rd Friday of the month, because we need revisions month-end to month-end.

in which t is repetitive:

$$Rev_t = \alpha + \beta Rev_{t-1} + \epsilon_t.$$

Here, Rev_t is the aggregate LTG revisions in month t, and as t is repetitive, $t-1$ is newsy. The positive and significant coefficient of 0.543 indicates that revisions in a repetitive month indeed tend to be in the same direction as those in the previous newsy month. However, this number is not in itself convincing. Survey data naturally incorporate latency,[9] resulting in a positive unconditional autocorrelation in revisions. Column 2 controls for this unconditional latency and extracts the overreaction effect using an interaction term between Rev_{t-1} and a dummy variable indicating whether t is repetitive. The resulting coefficient of 0.323 is again highly significant, indicating that the positive relation between the revisions in a repetitive month and the previous newsy month goes above and beyond what is implied by the data latency alone.

Column 3 tests prediction ii) by running the following regression on the sample in which t is newsy:

$$Rev_t = \alpha + \beta Rev_{t-2} + \epsilon_t.$$

Here, as t is newsy, $t-2$ is repetitive. We do not observe the negative coefficient predicted by our theory, but rather a near-zero coefficient of 0.038. However, this is simply owing to the confounding latency effect said above. Column 4 controls for this effect and extract the overreaction coefficient using the interaction term between Rev_{t-2} and a dummy variable indicating whether t is newsy. Here, we can clearly see a significantly negative coefficient of -0.174. This implies that reactions in the repetitive months are indeed overreactions that get corrected in the subsequent newsy month.

[9]One factor contributing to the latency in survey data is the requirement for analysts to produce detailed analyst reports (as examined in Barth et al. (2024) and Li et al. (2024)) in addition to their estimates. This increases the cost of generating estimates, thus causing delays in their release.

4.3 Earnings similarity and the strength of the return pattern

According to our mechanism, investors see a signal in the second month, which they believe is an independent draw from the distribution that produced the signal in the first month. In the data, however, this second signal are in fact earnings announcements from a different set of firms. These signals have a range of similarities with the first signal, which we capture with differences between aggregate return on equity in the two months. If the signal that investors see is highly dis-similar, then investors are not likely to interpret it as a draw from the same distribution. Thus one test of our underlying mechanism (see Prediction 2 is that over-reaction is likely to be especially strong when the earnings in the repetitive month are similar to those announced in the previous newsy month. We now test this prediction. Table 9 reports the key parameter β_2 on the interaction term from regression 2 conducted separately for episodes in which the earnings announced in an repetitive month and the preceding newsy month, as measured by ROE, are unusually similar, moderate, and unusually dissimilar. It shows that, when the earnings announced in the repetitive month more closely resemble those in the preceding newsy month, we indeed observe an amplified pattern of dynamic autocorrelation in the return of this repetitive month and the next newsy month. This confirms our model's prediction that investors overreact more to a repetitive month's earnings when they are more similar to those in the preceding newsy month.

4.4 The cross section of industries

4.4.1 Baseline return pattern

The main intuition behind the theory is as follows. If a group of stocks are 1) tightly and obviously connected in fundamentals, so that investors actively infer group-level information from its members' announcements and 2) sizable in number, so that the group as a whole announces progressively along the earnings cycle, then earnings announced in the second month of a quarter will be predictably similar to those announced in the first month. Failure to recognize this inherent similarity leads to predictable overreaction in the second month that later reverses.

This story applies naturally to all stocks in the US economy, but should additionally apply to industry-level returns. Two stocks randomly drawn from an industry are more connected with each other than two random stocks drawn from an economy. An industry is a smaller economy except more tightly connected. Investors clearly understand the tight intra-industry connection and actively learn industry-level information from the early earnings announcements (e.g., Foster (1981)). If the story is indeed about the early earnings announcements, the dynamic serial predictive relation in the aggregate market returns should also exist in industry returns *in excess of* the aggregate market returns.

Table 10 tests this prediction by conducting the following regression:

$$exret_{i,t} = \alpha + \beta exret_{i,nr(t)} + \epsilon_{i,t}. \tag{4}$$

Here $exret_{i,t}$ is the return of industry i in excess of the market return in month t. The returns are weighted by market capitalization, and the regression is weighted by industry level market capitalization scaled to sum to 1 in each cross section. Following Lou et al. (2019), we drop stocks with a smaller market capitalization than the 10th percentile on NYSE. We further drop industries with less than 10 constituents to ensure a sufficient number of announcers. Table 10 shows that the same tug-of-war pattern between continuation and reversal exists in industry excess returns in addition to aggregate market returns.

Two points are worth noting. First, this piece of cross sectional evidence is important in its own right. Table 11 shows that it can also be converted into a profitable trading strategy. Like the regressions in Table 10, this strategy uses market capitalization weights, making it highly implementable. Second, the consistency between aggregate and cross sectional evidence is not always seen in the return predictability literature. This piece of evidence thus speaks potently in favor of our hypothesis.

4.4.2 Heterogeneities across industries

Rich variation across industries offers us a good opportunity to further test our theory. As said in the previous subsection, our theory works on a group of stocks that are i) tightly connected in fundamentals and ii) sizable in numbers. These two elements are both necessary to set up the stage of intra-group learning, generate the newsy and repetitive signals for group-level cash flow in the first and second month of a quarter, and allow the possibility of correlation neglect. A natural implication of our theory is the return pattern should be stronger on industries that are more tightly connected and have more constituents.

Panel A of Table 12 test these predictions. We measure industry level earnings connectivity by first computing the covariance of normalized excess ROE (winsorized at -4 and 4) for each pair of stocks in the same industry-quarter, and then taking the within-industry average weighted by the average of the market capitalization of the pair of stocks. We then split each cross section is then split between high connectivity industries and low connectivity ones, and run the regression:

$$exret_{i,t} = \alpha + \beta_1 exret_{i,nr(t)} + \beta_2 exret_{i,nr(t)} \times I_t^n + \gamma I_t^n + \epsilon_{i,t}. \quad (5)$$

We clearly see that the effect of correlation neglect, as measured by the size of the coefficient β_2 on the interaction term, is much larger on the high connectivity industries. Panel B conducts a similar exercise for the number of firms in each industry, and show that the effect of correlation neglect is larger on the larger industries. Furthermore, as argued in Section 3, our theory predicts that the overreaction in the repetitive months is larger when the earnings in them more closely resemble those in the preceding newsy months. Panel C show that this is indeed the case in data.

To further demonstrate that the empirical pattern relates to the economic structure of real industries, we conduct a placebo test in Panel B of Table 12 by randomly generating fake "industries" and conduct regression 5 on the panel of these fake industries. The randomization approaches follow those in Chen et al. (2024). Panel B1 randomly assign each stock to 1 of the 39 (median number of industries across cross sections) "industries" with uniform probability. This approach generates random industries of roughly equal size (number of constituents). Panel B2 assign each stock to the industry of a randomly chosen stock in the same cross section without replacement. This approach preserves the size distribution of the industries in each cross section. Panel B3 first sorts each cross section into deciles by $exret_{s,nr(t)}$, where s is a stock, and then assigns each stock to the industry of a randomly chosen stock in the same cross section-decile. This approach preserves the size distribution of the industries in each cross section and further controls for $exret_{i,nr(t)}$. The 1st percentiles in Panel B of Table 12 are much less negative than the -0.164 observed in Panel A of Table 12. The results in Panel B3 are especially useful. It tells us that the observed return pattern in Table 12 depends critically on the economic structure of genuine industries and is not the sole result of the regressor $exret_{i,nr(t)}$ itself.

5 Alternative explanations

Predictable resolution of risks. The most important set of explanations for return predictability comes from predictable resolution of risks, which leads to high expected returns, as well as going through predictably low-risk periods, which leads to low expected returns. We have argued that this type of explanations is unlikely to account for our return pattern in its full scale owing to the frequency at which the expected stock market excess returns are negative. We now ignore this aspect of scale and discuss potential stories that can explain the pattern in direction.

The type of variation in risks that is necessary to explain my time-series results is that, i) after a good newsy month, the stock market is risky in the subsequent repetitive month, and ii) after a good repetitive month, the subsequent newsy month is safe. Individually, both points are not difficult to accommodate in a risk-based framework. In fact, point ii) naturally arises when a repetitive month resolve risks and take them away from the next newsy month. However, as we show in Table 6, we ideally should have one story that jointly explain both arms of the return pattern. This first requires the risk of the market to vary at a monthly frequency. Standard asset pricing models such as habit, long-run risk, disaster, and intermediary-based asset pricing all involve risks that should be rather persistent at a monthly frequency. So it does not appear hopeful to use standard macro risk factors to explain my predictability pattern.

Another candidate is the risk resolution associated with earnings announcements. Empirically, Savor and Wilson (2016) show that early announcers earn higher excess returns than later announcers (Table IV), which is indeed consistent with the overall importance of the earnings season. This, however, does not explain why the current earnings season being good makes the next repetitive month risky. Neither does it explain why a good repetitive month is followed by a safe newsy month.

Savor and Wilson (2013) document that the equity premium on macroeconomic announcement days is especially large and argue that this is because it resolves important macroeconomic risks. If this contributes to our pattern of return predictability, the return predictability would be realized on macroeconomic announcement days. A straightforward empirical investigation reveals that the predictability in fact works about equally well on the top five macroeconomic announcement dates on Bloomberg (i.e., FOMC, employment condition, GDP, PPI, and Institute for Supply Management manufacturing). The data suggest that the return predictability documented in this paper does not obviously relate to macroeconomic announcements.

Firms endogenously changing their announcing latency. A large accounting literature examines the relation between the timing of firms' earnings announcements and the news they convey. A clear empirical pattern that emerges is that early announcers tend to announce better news than late announcers (e.g., Kross 1981; Kross and Schroeder 1984; Chambers and Penman 1984; Johnson and So 2018; Noh et al. 2021). The literature has also extensively studied the reasons behind this pattern and has partially attributed it to firms' endogenous choices of the announcing lag. For instance, deHaan et al. (2015) argue that firms delay earnings announcements with bad results to avoid the early portion of the earnings cycle that receives heightened attention. Givoly and Palmon (1982) argue that they do so to buy time so that they can manipulate their accounting results.

However, it is not clear how this empirical regularity alone speaks to our results. Overall, it implies that early announcers announce better results than late announcers. If investors fail to anticipate that, then early announcers will have higher earnings surprises and higher announcement excess returns. It is not clear why it would cause the market return in the newsy month to positively correlate with that in the next repetitive month. If anything, exogenous variation in the intensity of this self-selection effect seems to lead to a negative correlation of early announcement returns and late announcement returns within the quarter, as bad announcers being moved out of the newsy months makes the newsy months look better, and the subsequent repetitive months look worse than they otherwise would. And even then, it is not clear why a pattern of return reversal would exist between a repetitive month and the upcoming newsy month.

6 Conclusion

Contrary to prior beliefs, monthly stock market returns in the United States can in fact be predicted with past returns. Specifically, the U.S. stock market's return during the first month of a quarter positively predicts the second month's return, which then negatively predicts the first month's return of the next quarter. The first months of a quarter are "newsy" because they contain fresh earnings news about the aggregate economy, whereas the second months of a quarter contain earnings announcements that produce predictably repetitive "news." We hypothesize that the return predictability pattern arises because investors use announced earnings to predict future earnings, but did not recognize that earnings in the second months of a quarter are inherently repetitive of those in the previous newsy month. Survey data support this hypothesis of correlation neglect, as does out-of-sample evidence across industries. These results challenge the efficient market hypothesis by documenting a strong and pervasive form of return predictability.

References

Adrian, T., Etula, E., and Muir, T. (2014). Financial intermediaries and the cross-section of asset returns. *Journal of Finance*, 69(6):2557–2596.

Baker, M. and Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *Journal of Finance*, 55(5):2219–2257.

Baker, S. R., Bloom, N., Davis, S. J., and Sammon, M. C. (2021). What triggers stock market jumps? Working Paper, Northwestern University.

Barth, M. E., Li, K., and McClure, C. (2024). Information supporting investor valuations: Evidence from a comparative content analysis of analyst reports and form 10-K. Working Paper, SSRN, https://papers.ssrn.com/abstract=4763765, Accessed 2024-09-01.

Bordalo, P., Gennaioli, N., Porta, R. L., and Shleifer, A. (2024). Belief overreaction and stock market puzzles. *Journal of Political Economy*, 132(5):1450–1484.

Campbell, J. and Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21(4):1509–1531.

Chambers, A. and Penman, S. (1984). Timeliness of reporting and the stock-price reaction to earnings announcements. *Journal of Accounting Research*, 22(1):21–47.

Chen, H., Dou, W. W., Guo, H., and Ji, Y. (2024). Industry Distress Anomaly.

deHaan, E., Shevlin, T., and Thornock, J. (2015). Market (in)attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics*, 60(1):36–55.

Enke, B. and Zimmermann, F. (2019). Correlation neglect in belief formation. *Review of Economic Studies*, 86(1):313–332.

Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business*, 38(1):34–105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fedyk, A. and Hodson, J. (2023). When can the market identify old news? *Journal of Financial Economics*, 149(1):92–113.

Foster, G. (1981). Intra-industry information transfers associated with earnings releases. *Journal of Accounting and Economics*, 3(3):201–232.

Gao, C. and Martin, I. W. R. (2021). Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment. *Journal of Finance*, 76(6):3211–3254.

Givoly, D. and Palmon, D. (1982). Timeliness of annual earnings announcements: Some empirical evidence. *Accounting Review*, 57(3):486–508.

Goyal, A. and Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508.

Goyal, A., Welch, I., and Zafirov, A. (2024). A comprehensive 2022 look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 37(11):3490–3557.

Guo, H. (2025). Earnings Extrapolation and Predictable Stock Market Returns. *The Review of Financial Studies*, page hhaf020.

He, Z., Kelly, B., and Manela, A. (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1):1–35.

Johnson, T. L. and So, E. C. (2018). Time will tell: Information in the timing of scheduled earnings news. *Journal of Financial and Quantitative Analysis*, 53(6):2431–2464.

Kendall, M. G. and Hill, A. B. (1953). The analysis of economic time-series, part I: Prices. *Journal of the Royal Statistical Society. Series A (General)*, 116(1):11–34.

Kross, W. (1981). Earnings and announcement time lags. *Journal of Business Research*, 9(3):267–281.

Kross, W. and Schroeder, D. A. (1984). An empirical investigation of the effect of quarterly earnings announcement timing on stock returns. *Journal of Accounting Research*, 22(1):153–176.

Lewellen, J. and Shanken, J. (2002). Learning, asset-pricing tests, and market efficiency. *Journal of Finance*, 57:1113–1145.

Li, K., Mai, F., Shen, R., Yang, C., and Zhang, T. (2024). Dissecting corporate culture using generative AI: Insights from analyst reports. Working Paper, SSRN, https://papers.ssrn.com/abstract=4558295, Accessed 2024-09-01.

Li, Y. and Wang, C. (2023). Slicing an asset to learn about its future: A new perspective on return and cash-flow forecasting. Working Paper, SSRN.

Lo, A. W. and MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1):41–66.

Lou, D., Polk, C., and Skouras, S. (2019). A tug of war: Overnight versus intraday expected returns. *Journal of Financial Economics*, 134(1):192–213.

Lou, D., Polk, C., and Skouras, S. (2024). The day destroys the night, night extends the day: A clientele perspective on equity premium variation. Working Paper, London School of Economics.

Martin, I. (2017). What is the expected return on the market? *Quarterly Journal of Economics*, 132(1):367–433.

Nagel, S. and Xu, Z. (2022). Asset pricing with fading memory. *Review of Financial Studies*, 35(5):2190–2245.

Noh, S., So, E. C., and Verdi, R. S. (2021). Calendar rotations: A new approach for studying the impact of timing using earnings announcements. *Journal of Financial Economics*, 140(3):865–893.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and Implications. *Journal of Financial Economics*, 22(1):27–59.

Rapach, D. E., Ringgenberg, M. C., and Zhou, G. (2016). Short interest and aggregate stock returns. *Journal of Financial Economics*, 121(1):46–65.

Rapach, D. E. and Zhou, G. (2022). Asset pricing: Time-series predictability. In *Oxford research encyclopedia of economics and finance*. Oxford University Press, Oxford.

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

Savor, P. and Wilson, M. (2016). Earnings announcements and systematic risk. *Journal of Finance*, 71(1):83–138.

Table 1: When do firms end their fiscal quarters and when do they announce earnings

	A: FQ-end month		B: Earnings announcement month	
	Count	Percent	Count	Percent
Group 1: Jan/Apr/Jul/Oct	82,257	8.6%	358,128	44.7%
Group 2: Feb/May/Aug/Nov	51,169	5.4%	382,162	47.7%
Group 3: Mar/Jun/Sep/Dec	819,534	86.0%	60,649	7.6%
Total	952,960	100.0%	800,939	100.0%

Panel A counts the company-fiscal quarters in Compustat by the months in which they end. The fiscal quarters always end on the last day of a month, and 86.0% means that 86.0% of the fiscal quarters end on the last day of March, June, September, or December. Panel B counts the company-fiscal quarters by the months in which the fiscal quarters' earnings are announced, requiring that i) the fiscal quarter ends in the Group 3 months and ii) the earnings announcement occurs within 92 days of the fiscal quarter end date. Data are quarterly from 1971 to 2023.

Table 2: Continuation and reversal in the U.S. aggregate market returns

	(1) Repetitive	(2) Newsy	(3) Difference
	A: Full sample		
$mkt_{nr(t)}$	0.279***	-0.279***	-0.557***
	[3.13]	[-3.02]	[-4.35]
N	386	386	772
	B: First half		
$mkt_{nr(t)}$	0.358***	-0.327**	-0.685***
	[2.71]	[-2.48]	[-3.67]
N	190	190	380
	C: Second half		
$mkt_{nr(t)}$	0.167***	-0.194**	-0.360***
	[2.77]	[-2.14]	[-3.31]
N	196	196	392
	D: Post-WWII		
$mkt_{nr(t)}$	0.142***	-0.163**	-0.305***
	[2.69]	[-2.31]	[-3.46]
N	304	304	608

This table presents estimated βs from the monthly time-series regression $mkt_t = \alpha + \beta mkt_{nr(t)} + \epsilon_t$. Here, mkt_t is the U.S. stock market return in excess of the risk free rate in month t, and $nr(t)$ is the most recent month before t that is newsy or repetitive. Column 1 is on the sample in which the dependent variable month t is a second month of a quarter (repetitive); column 2 is on the sample in which the dependent variable month t is a first month of a quarter (newsy). Column 3 shows there difference, extracted as β_2 from the regression $mkt_t = \alpha + \beta_1 mkt_{nr(t)} + \beta_2 mkt_{nr(t)} \times I_t^n + \gamma I_t^n + \epsilon_t$ on the combined sample in column 1 and 2, where I_t^n is a dummy variable taking the value of 1 if month t is newsy, and 0 otherwise. In Panel A, B, C, and D, data are monthly from 1926 to 2023, from 1926 to 1973, from 1974 to 2023, and 1947 to 2023, respectively. T-statistics computed with White standard errors are reported in square brackets.

Table 3: Out-of-sample R^2 in forecasting the U.S. stock market returns

Method	1	2	3
OOS R^2	4.20%	4.26%	4.73%

The R^2 in this table are calculated as $1 - \frac{\sum_{t=1}^{n}(r_t - \hat{r_t})^2}{\sum_{t=1}^{n}(r_t - \overline{r_t})^2}$, where $\overline{r_t}$ is the expanding window mean of past stock returns and $\hat{r_t}$ is the forecast being evaluated. The signal is $mkt_{nr(t)} - m\bar{kt}_{t-1}$ for a repetitive t, $-(mkt_{nr(t)} - m\bar{kt}_{t-1})$ for a newsy t, and 0 otherwise. For each month t, 2 coefficients—one for the signal and one for the constant term—are extracted from a simple expanding-window OLS regression of historical market excess returns on historical signal values and a constant. For method 1, the forecast in a given month is the estimated coefficient on the signal multiplied with the signal value plus the estimated constant. Method 2 replaces the estimated constant coefficient with $\overline{r_t}$. Method 3 replaces the estimated constant with the average signal (of the dividend/price, earnings/price, and book-to-market based numbers) proposed by Campbell and Thompson (2008) in the valuation constraint + growth specification with fixed coefficients. Data are monthly from 1926 to 2023.

Table 4: Time-series strategy performance

	(1) TSP_t	(2) TSP_t	(3) TSP_t	(4) TSP_t
MKT_t		0.169	0.050	-0.024
		[0.96]	[0.39]	[-0.23]
HML_t			0.575**	0.425**
			[2.17]	[2.22]
SMB_t			0.165	0.148
			[0.88]	[0.76]
MOM_t				-0.331**
				[-2.22]
α	0.668***	0.554***	0.402**	0.715***
	[3.51]	[3.24]	[2.32]	[3.73]
N	1,165	1,165	1,165	1,164

Column 1 shows results from the following monthly time-series regression: $TSP_t = \alpha + \epsilon_t$. Here, TSP_t is our time-series portfolio return in month t. This portfolio takes long or short positions in the aggregate market for a given month, and the weight is proportional to the previous-month-end signal value s_{t-1}, which is $mkt_{nr(t)} - \overline{mkt_{t-1}}$ for a repetitive t, $-(mkt_{nr(t)} - \overline{mkt_{t-1}})$ for a newsy t, and 0 otherwise. The market weight is then scaled so that the strategy has a historical volatility of 5% as of $t-1$, and then truncated at 3 and -3. Columns 2-4 add in the market, value, size, and momentum factor returns on the right-hand-side. In Column 1, the coefficient of the constant represents the average return of the time-series portfolio. In Columns 2-4, it represents its alphas with different factor models. Data are monthly from 1926 to 2023. Returns are all in percentage units. T-statistics computed with White standard errors are reported in the square brackets.

Table 5: Continuation and reversal in the U.S. aggregate market returns, with controls

	(1) Repetitive	(2) Newsy	(3) Difference
$mkt_{nr(t)}$	0.213**	-0.248***	-0.461***
	[2.22]	[-2.70]	[-3.46]
$\sum_{j=1}^{4} mkt_{nm(t,j)}$	0.329**	-0.251*	-0.580***
	[2.46]	[-1.85]	[-3.05]
N	383	383	766

This table presents estimated βs from the monthly time-series regression $mkt_t = \alpha + \beta_1 mkt_{nr(t)} + \beta_2 \sum_{j=1}^{4} mkt_{nm(t,j)} + \epsilon_t$. Here, mkt_t is the U.S. stock market return in excess of the risk free rate in month t, $nr(t)$ is the most recent month before t that is newsy or repetitive, and $nm(t,j)$ is the jth most recent newsy month before t. Column 1 is on the sample in which the dependent variable month t is a second month of a quarter (repetitive); column 2 is on the sample in which the dependent variable month t is a first month of a quarter (newsy). Column 3 shows there difference, extracted as β_2 from the regression of $mkt_t = \alpha + \beta_1 mkt_{nr(t)} + \beta_2 mkt_{nr(t)} \times I_t^n + \beta_3 \sum_{j=1}^{4} mkt_{nm(t,j)} + \beta_4 \sum_{j=1}^{4} mkt_{nm(t,j)} \times I_t^n + \gamma I_t^n + \epsilon_t$ on the combined sample in column 1 and 2, where I_t^n is a dummy variable taking the value of 1 if month t is newsy, and 0 otherwise. Data are monthly from 1926 to 2023. T-statistics computed with White standard errors are reported in square brackets.

Table 6: The link between the continuation and reversal effects

	(1) TSP_t^n	(2) TSP_t^r	(3) $Rank(TSP_t^n)$	(4) $Rank(TSP_t^r)$
TSP_t^r	0.957*** [15.16]			
TSP_t^n		0.893*** [6.87]		
$Rank(TSP_t^r)$			0.361*** [3.42]	
$Rank(TSP_t^n)$				0.361*** [3.40]
N	97	97	97	97

Column 1 shows results from the following annual time-series regression: $TSP_t^n = \alpha + \beta TSP_t^r + \epsilon_t$. Here, TSP_t is our time-series portfolio return in year t. This portfolio takes long or short positions in the aggregate market for a given month, and the weight is proportional to the previous-month-end signal value s_{t-1}, which is $mkt_{nr(t)} - \overline{mkt_{t-1}}$ for a repetitive t, $-(mkt_{nr(t)} - \overline{mkt_{t-1}})$ for a newsy t, and 0 otherwise. The market weight is then scaled so that the strategy has a historical volatility of 5% as of $t-1$, and then truncated at 3 and -3. TSP_t^n is the portfolio's return in the newsy months of year t, and TSP_t^r is the return in the repetitive months of year t. The $Rank(\cdot)$ function ranks the TSP_t^r and TSP_t^n across years, thus transforming them into integers between 1 and 97. T-statistics computed with White standard errors are reported in the square brackets.

Table 7: The first weeks of the newsy months

	(1) Newsy	(2) Week 1	(3) Ex week 1
$mkt_{nr(t)}$	-0.279***	-0.021	-0.256***
	[-3.02]	[-0.53]	[-2.76]
N	386	386	386

Column 1 presents estimated βs from the monthly time-series regression $mkt_t = \alpha + \beta mkt_{nr(t)} + \epsilon_t$. Here, mkt_t is the U.S. stock market return in excess of the risk free rate in month t, and $nr(t)$ is the most recent month before t that is newsy or repetitive. Column 1 is on the sample in which the dependent variable month t is a first month of a quarter (newsy). Column 2 replaces the dependent variable with mkt_t^{FW}, the market excess return during the first five trading days of month t. Column 3 replaces the dependent variable with mkt_t^{ExclFW}, the excess market return during t excluding the first five trading days. Data are monthly from 1926 to 2023. T-statistics computed with White standard errors are reported in square brackets.

Table 8: Evidence from survey data

	(1) Rev_t	(2) Rev_t	(3) Rev_t	(4) Rev_t
Rev_{t-1}	0.543***	0.220***		
	[4.91]	[2.94]		
I_t^r		0.078		
		[1.15]		
$Rev_{t-1} \times I_t^r$		0.323***		
		[3.42]		
Rev_{t-2}			0.038	0.211***
			[1.02]	[3.92]
I_t^n				-0.052
				[-1.05]
$Rev_{t-2} \times I_t^n$				-0.174***
				[-2.72]
const	-0.032	-0.110**	-0.143***	-0.091*
	[-0.81]	[-2.03]	[-3.04]	[-1.93]
N	170	509	170	509

Column 1 runs the following monthly time-series regression on the sample in which t is repetitive: $Rev_t = \alpha + \beta Rev_{t-1} + \epsilon_t$. Here, Rev_t is the aggregate revision in the IBES long-term growth (LTG) measure for month t within the S&P 500 universe using the market capitalization weight. Column 2 adds in a dummy variable indicating whether t is repetitive and its interaction with Rev_{t-1}, and expands the sample to all months. Columns 3 and 4 conduct analogous regressions to columns 1 and 2, but for newsy months rather than repetitive months. Data are monthly from 1982 to 2024. T-statistics computed with Newey-West standard errors are reported in square brackets.

Table 9: Aggregate earnings distance and the strength of the return pattern

	(1) Lowest 10%	(2) 10%-50%	(3) 50%-90%	(4) Highest 10%	(5) Low - High
$mkt_{nr(t)} \times I_t^n$	-0.769***	-0.398*	-0.228	-0.084	-0.684**
	[-4.33]	[-1.94]	[-1.51]	[-0.31]	[-2.12]
N	42	170	170	42	84

This table presents estimated β_2s from the monthly time-series regression $mkt_t = \alpha + \beta_1 mkt_{nr(t)} + \beta_2 mkt_{nr(t)} \times I_t^n + \gamma I_t^n + \epsilon_t$.. Here, mkt_t is the U.S. stock market return in excess of the risk free rate in month t, $nr(t)$ is the most recent month before t that is newsy or repetitive, and I_t^n is a dummy variable taking the value of 1 if month t is newsy, and 0 otherwise. Across columns, the regression sorts on the distance between the aggregate ROE announced in the most recent newsy months before t and the subsequent repetitive month. The aggregate ROE in a given month is defined as the total earnings announced in the month divided by the total book value of equity of these announcing firms. The ROE distance is computed as the absolute value of the difference between the newsy and the repetitive months' ROE relative to the median value of the difference across time. The distance increases from column 1 to 6 and are between the 0-10th, 10-50th, 50-90th, and 90-100th percentiles of the sample, respective. Data are monthly from 1971-2023. T-statistics computed with White standard errors are reported in square brackets.

Table 10: Continuation and reversal in the U.S. industry excess returns

	(1) Repetitive	(2) Newsy	(3) Difference
	A: Full sample		
$exret_{i,nr(t)}$	0.083***	-0.081**	-0.164***
	[2.68]	[-2.36]	[-3.55]
N	16,214	16,283	32,497
	B: First half		
$exret_{i,nr(t)}$	0.085*	-0.067	-0.152**
	[1.73]	[-1.36]	[-2.19]
	4,049	4,107	8,156
	C: Second half		
$exret_{i,nr(t)}$	0.082**	-0.089*	-0.171***
	[2.06]	[-1.94]	[-2.82]
	12,165	12,176	24,341
	D: Post-WWII		
$exret_{i,nr(t)}$	0.071**	-0.067*	-0.138***
	[2.17]	[-1.72]	[-2.71]
	15,044	15,102	30,146

This table presents estimated βs from the industry-monthly panel regression $exret_{i,t} = \alpha + \beta exret_{i,nr(t)} + \epsilon_{i,t}$. Here, $exret_{i,t}$ is the return of industry i in excess of market return in month t, and $nr(t)$ is the most recent month before t that is newsy or repetitive. Column 1 is on the sample in which the dependent variable month t is a second month of a quarter (repetitive); column 2 is on the sample in which the dependent variable month t is a first month of a quarter (newsy). Column 3 shows there difference, extracted as β_2 from the regression of $exret_{i,t} = \alpha + \beta_1 exret_{i,nr(t)} + \beta_2 exret_{i,nr(t)} \times I_t^n + \gamma I_t^n + \epsilon_t$ on the combined sample in column 1 and 2, where I_t^n is a dummy variable taking the value of 1 if month t is newsy, and 0 otherwise. The returns are weighted by market capitalization, and the regression is weighted by industry level market capitalization scaled to sum to 1 in each cross section. The regressions require at least 10 firms in an industry. In Panel A, B, C, and D, data are monthly from 1926 to 2023, from 1926 to 1973, from 1974 to 2023, and 1947 to 2023, respectively. T-statistics computed with standard errors clustered by month are reported in square brackets.

Table 11: Cross sectional strategy performance

	(1) CSP_t	(2) CSP_t	(3) CSP_t	(4) CSP_t
MKT_t		0.062	0.041	-0.044
		[0.92]	[0.68]	[-0.83]
HML_t			0.160	-0.017
			[1.20]	[-0.14]
SMB_t			-0.017	-0.035
			[-0.20]	[-0.44]
MOM_t				-0.381***
				[-3.58]
α	0.533***	0.497***	0.457***	0.827***
	[3.37]	[3.08]	[2.92]	[4.20]
N	1,154	1,153	1,153	1,152

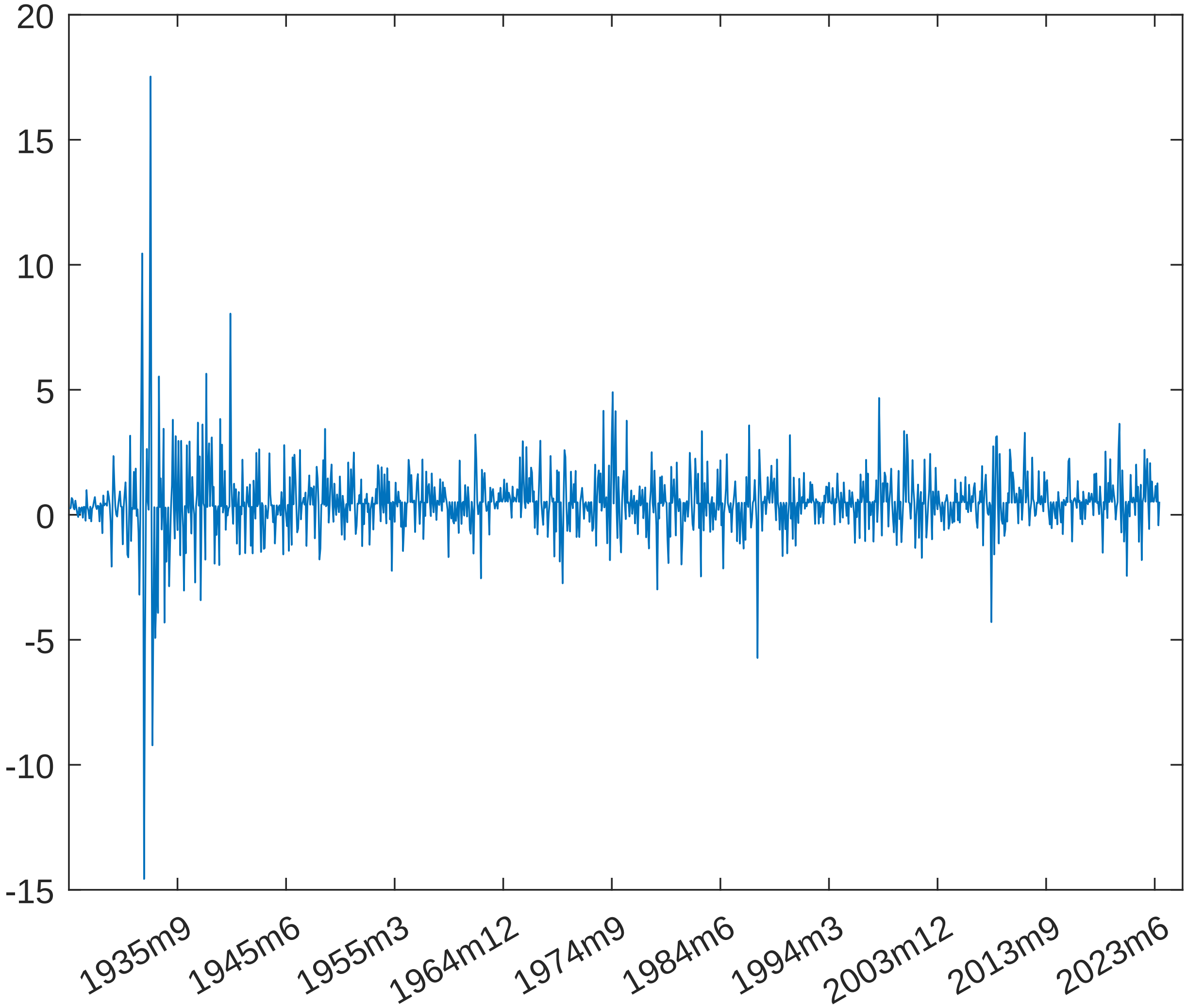
Column 1 shows results from the following monthly time-series regression: $CSP_t = \alpha + \epsilon_t$. Here, CSP_t is our cross-sectional portfolio return in month t. This portfolio takes long and short positions on industries in a given month, and the weight is constructed with the following steps: i) Compute signal value $s_{i,t-1}$, which is $exret_{i,nr(t-1)}$ for a repetitive t, $-exret_{i,nr(t-1)}$ for a newsy t, and 0 otherwise; ii) cross sectionally demean with market capitalization weight; iii) scale by standard deviation of the signal computed with data up to $t-1$ and Winsorize at [-3, 3] to limit the influence of extreme values; iv) interact with the industry's market capitalization divided by the total market capitalization in $t-1$. Step iv) is a simple way to amply positions in large stocks to lower trading costs. Columns 2-4 add in the market, value, size, and momentum factor returns on the right-hand-side. In Column 1, the coefficient of the constant represents the average return of the time-series portfolio. In Columns 2-4, it represents its alphas with respect to different factor models. Data are monthly from 1926 to 2023. Returns are all in percentage units. T-statistics computed with White standard errors are reported in the square brackets.

Table 12: Further analyses across industries

	A: Heterogeneity across industries			B: Random "industries"				
	Low	High	High-Low	1%	5%	50%	95%	99%
	A1: Size			B1: Uniform probability				
$exret_{i,nr(t)} \times I_t^n$	-0.062**	-0.141***	-0.079**	-0.049	-0.035	0.001	0.035	0.049
	[-2.31]	[-3.62]	[-2.57]					
	A2: Connectivity			B2: Preserve industry sizes				
$exret_{i,nr(t)} \times I_t^n$	-0.127**	-0.235***	-0.108**	-0.053	-0.036	-0.001	0.033	0.050
	[-2.14]	[-3.30]	[-2.00]					
	A3: Earnings similarity			B3: Control for $exret_{i,nr(t)}$				
$exret_{i,nr(t)} \times I_t^n$	-0.105	-0.255***	-0.150**	-0.066	-0.050	-0.015	0.021	0.036
	[-1.36]	[-4.14]	[-2.30]					

Panel A of this table presents estimated β_2s from the industry-monthly panel regression: $exret_{i,t} = \alpha + \beta_1 exret_{i,nr(t)} + \beta_2 exret_{i,nr(t)} \times I_t^n + \gamma I_t^n + \epsilon_t$. Here, $exret_{i,t}$ is the return of industry i in excess of market return in month t, $nr(t)$ is the most recent month before t that is newsy or repetitive, and I_t^n is a dummy variable taking the value of 1 if month t is newsy, and 0 otherwise. The returns are weighted by market capitalization, and the regression is weighted by industry level market capitalization scaled to sum to 1 in each cross section. The regressions require at least 10 firms in an industry. Panel A1 separates the regression by whether an industry has below- or above-average number of constituents in its cross section and shows the difference in β_2 in the 3rd column. Panel A2 conducts a similar exercise for within-industry earnings connectivity, measured by first computing the covariance of normalized excess ROE, winsorized at -4 and 4, for each pair of stocks in the same industry-quarter, and then taking the within-industry average weighted by the average of the market capitalization of the pair of stocks. Panel A3 does so for earnings similarity, computed as the absolute value of the distance between the industry's ROE announced in the newsy month before t and the subsequent repetitive month. Panel B conducts 3 placebo tests on randomly generated, fake "industries," and report the 1st, 5th, 50th, 95th, and 99th percentiles of the distribution of β_2 generated in 1000 simulations. Data are monthly from 1971-2023. T-statistics computed with standard errors clustered at monthly level are reported in square brackets.

Figure 1: Implied expected stock market excess returns over time



This figure plots the expected stock market return as in Method 2 of Table 3. The unit is percent per month.

Appendix

A Alternative model of correlation neglect

In this section, we consider an alternative model that is based more closely on Enke and Zimmermann (2019). As in the model in Section 3, assume there are three periods, corresponding to the newsy month, the repetitive month, and the first month of the next quarter when the value of the market portfolio $\tilde{P}$ is drawn from the normal distribution with mean μ. As in Enke and Zimmermann (2019), assume there are latent signals $s_j, j = 1, 2$, jointly normally distributed, independent, and with mean μ. Investors directly observe s_1 in period 1 and $\hat{s}_2 = 1/2(s_1 + s_2)$ in period 2. Investors need to estimate μ. Under the assumption of a flat prior on μ, $\hat{s}_2$ is the optimal forecast. Computing returns requires taking a stance on the full prior distribution. To focus attention on correlation neglect, we assume known variance of s_j and take the limit as the prior becomes uninformative.

Equilibrium with rational investors

In period 1, $P_1 = s_1$, and $R_1 = P_1 - \mu$. In period 2, $P_2 = \hat{s}_2$ and $R_2 = \frac{1}{2}(s_2 - s_1)$. Note that $E[R_1] = E[R_2] = 0$, and therefore $\mathrm{Cov}(R_1, R_2) = E[R_1 R_2] = \frac{1}{2}E[(s_1 - \mu)(s_2 - s_1)]$. Returns ex post appear to be predictable: $E[R_2 R_1] < 0$. However, as Lewellen and Shanken (2002) emphasize, the investor cannot profit from this opportunity in real time: $E[R_2|s_1] = 1/2(E[s_2|s_1] - s_1) = 0$. Finally,

$$\begin{aligned} E[R_3 R_2] &= \frac{1}{2}E[(\tilde{P} - \hat{s}_2)(s_2 - s_1)] \\ &= \frac{1}{2}\left(E[(\tilde{P} - \hat{s}_2)s_2] - E[(\tilde{P} - \hat{s}_2)s_1]\right) = 0 \end{aligned}$$

where the last equation follows from the symmetry in s_1 and s_2. It may seem surprising that this too is not negative. The reason is that the initial "over-reaction" from s_1 is "over-corrected" by s_2. In fact R_3 is negatively correlated with the overall return $R_2 - \mu$.

Equilibrium with correlation neglect

In period 1, again $P_1 = s_1$ and $R_1 = s_1 - \mu$. However, in period 2, investors form an inefficient forecast that over-weights s_1, $P_2 = \frac{1}{2}(s_1 + \hat{s}_2) = \frac{3}{4}s_1 + \frac{1}{4}s_2$, because they mistakenly believe that $\hat{s}_2$ is an independent draw. The period 2 return equals $R_2 = P_2 - P_1 = 1/4(s_2 - s_1)$. As before $E[R_1] = E[R_2] = 0$. Again, $\mathrm{Cov}(R_1, R_2) = E[R_1 R_2]$. Though prices in the second period differ in this economy as compared with the one above, the return R_2 differs only by a multiplicative constant. Therefore, by the logic for the rational investor, $\mathrm{Cov}(R_1, R_2) < 0$, which is counterfactual.

Furthermore, $\mathrm{Cov}(R_3, R_2) > 0$. It suffices to show $E[R_3R_2] > 0$. We have

$$\begin{aligned}
E[R_3R_2] &= \frac{1}{4}E[(\tilde{P} - \frac{3}{4}s_1 - \frac{1}{4}s_2)(s_2 - s_1)] \\
&= \frac{1}{4}E[(\tilde{P} - \hat{s}_2) - \frac{1}{4}s_1 + \frac{1}{4}s_2)))(s_2 - s_1)] \\
&= \frac{1}{4}E[(\tilde{P} - \hat{s}_2)(s_2 - s_1)] + \frac{1}{16}E[(s_2 - s_1)^2] \\
&= \frac{1}{16}E[(s_2 - s_1)^2] > 0
\end{aligned}$$

The fact that the first term is zero follows from the same reasoning as for the rational investor. The intuition for the positive correlation is that P_2 contains "too much" s_1 and "too little" s_2. Thus on average, R_3 moves in the same direction as R_2, which is to say toward s_2 and away from s_1. It might seem surprising that R_2 moves from s_2 and away from s_1 given that the point of correlation neglect is to have P_2 incorporate too much s_1. However, under this form of uncertainty it is impossible for it to incorporate more than is already in P_1. In effect, the information in s_1 is fully priced in at time 1.

This model differs from the one in the main text in two respects. Most obviously, investors here learn about the mean of a normal distribution, whereas in main modle, investors learn about a probability. When learning about the mean, receiving the same signal twice and incorporating it does not change the investor's estimate of the price (though it does shrink the variance). However, when one is learning about a probability, receiving a signal always causes the mean to update (see Equation 3), though eventually by a vanishingly small amount. In a model with normally distributed risks, over-reaction only occurs in the second moment, not in the first, as in the data. A second difference lies in the unique source of randomness in the second period. In this model, that source of randomness is an extra signal, and creates a negative correlation. In the model in the main text, the source of randomness is whether the first signal is brought to investors' attention or not. Despite these differences, the models have an important source of similarity in that incorrect inference arises from neglecting correlations.

B Timing-series signals and trading strategies

In the cross sectional return forecasting literature, it is a common practice to form long-short portfolios according to a signal's values across stocks. For example, Fama and French (1993) create the HML factor, which takes a long position in the 30% of stocks with the highest book-to-market ratios and shorting the 30% of stocks with the lowest book-to-market ratios. This approach has many benefits, among which are the following three. First, it does not require explicit forecasts for stock returns and estimates no parameter. Second, it works with any stock-level signal. Third, it creates a reasonable risk exposure that makes the portfolio returns comparable across signals, as the equity exposures of the long and short arms of the portfolio are the same 100% and -100%.

In the time-series literature, we do not yet have a commonly used approach to create analogous portfolios. Such portfolios are an useful diagnostic tool that complements the standard metric of out-of-sample R^2 put forth by Goyal and Welch (2008). We develop such a tool in this section. Overall, for a generic signal s, we construct a beta-neutral market-timing strategy that rebalances positions in i) aggregate stock market and ii) the risk-free bond. The strategy implements two rules with data available in real time. First, it places proportional weights on the market to the signal value, with a target mean of zero. Second, the portfolio weight is scaled so that the strategy has a target volatility of, say, 5% per month.

Let s_t be the signal value at the end of period t. Then, the unscaled portfolio weight on the market is s_{t-1}, and the unscaled strategy volatility $\bar{\sigma}^s_{t-1}$ is the unscaled strategy's historical volatility up to $t-1$. The scaled portfolio weight is $s_{t-1}\frac{0.05}{\bar{\sigma}^s_{t-1}}$. We further truncate the weight on the market at -3 and 3 to avoid unrealistic strategies.[10] The strategy's construction involves no look-ahead bias, and the scaling can be adjusted according to an investor's risk appetite.

Columns Return5 and Sharpe5 of Table B1 tabulates means and Sharpe ratios of the portfolios for existing signals in the literature forecasting aggregate stock market returns. Four points are worth noting. First, likely owing to a consistent risk appetite targeting i) a historical strategy volatility of 5% and ii) a maximum market weight of 300% and -300%, the returns are Sharpe ratios of the diagnostic portfolios are rarely catastrophic. This sharply contrasts the message conveyed by the OOS R^2s, which can be several or even tens of percentage points negative, even with the constraints from Campbell and Thompson (2008).

Second, the returns and Sharpe ratios have positive mean and median across the 23 signals. The average returns have a mean and median of 0.07% per month, and the Sharpe ratios have a mean and median of 0.05 and 0.04 per annum. This relates to the first point, and again contrasts the message in the OOS R^2s, which, with a 5 year burn-in period, have a mean of -4.42% and a median of -1.49%, and a mean of -1.14% and a median of -0.22%

[10]Since overly high leverages are unattainable in practice, such truncation is always (though perhaps implicitly) present when trading in real world.

with the CT constraints.

Third, the two metrics based on our diagnostic portfolios are (perhaps surprisingly) dissimilar to the OOS R^2s. Table B2 show that with the same burn-in period of 5 years, the returns and Sharpe ratios have a near-zero correlation of 0.03 and 0.12 with the original OOS R^2 in Goyal and Welch (2008), and a negative correlation of -0.50 and -0.32 with the OOS R^2 in Campbell and Thompson (2008). In contrast, the portfolio-based metrics positively correlate with the in-sample R^2, though at 0.5, this correlation is far from perfect.

Fourth, related to the third point, it is entirely possible for a signal to have very low R^2 but high diagnostic portfolio returns and Sharpe ratio. The signal rtz proposed in Rapach et al. (2016) is such an example. It has low OOS R^2s around -10% but the highest diagnostic portfolio average return of 0.87% per month, and the highest Sharpe ratio of 0.43 per annum.

These four differences arise from a key economic difference between the two sets of metrics. The R^2 computation involves explicit return forecasts and therefore the estimation of an regression coefficient on the signal in real time. This coefficient is an representation of the signal's historical performance: A time-series regression of the market excess return mkt_t on the signal s_{t-1} and a constant yield a coefficient of $\beta^s = \sum_t \frac{(s_{t-1}-\bar{s})}{\sum_t (s_{t-1}-\bar{s})^2} mkt_t = \sum_t w_{t-1} mkt_t$ on s. Here, $\bar{s}$ is the in-sample mean of the signal s. Note that β^s is the mean excess return of a portfolio that rebalances between positions in the aggregate stock market (which yields an excess return of mkt_t) and the risk-free bond (which yields an excess return of 0). The month-t portfolio weight on the stock market $w_{t-1} = \frac{(s_{t-1}-\bar{s})}{\sum_t (s_{t-1}-\bar{s})^2}$ is i) linear in the signal's value s_{t-1}, ii) on average 0, so that the portfolio is on average market neutral, and iii) scaled by $\sum_t (s_{t-1} - \bar{s})^2$, so that the portfolio has unit exposure to s (i.e., $\sum_t \frac{(s_{t-1}-\bar{s})}{\sum_t (s_{t-1}-\bar{s})^2} s_{t-1} = 1$).[11]

Using this coefficient to form a quantitative forecast of the market excess return implicitly assumes that performance of the signal's coefficient portfolio is so stable that its past performance reliably predicts its future performance at any point in history. This is a somewhat paradoxical assumption in the first place, as it likely implies too high a T-statistic and a Sharpe ratio—as the coefficient portfolio's return volatility approaches 0, the T-statistic in the regression and the portfolio's Sharpe ratio both approach infinity. In particular, extreme performance at the beginning of the sample, whether good or bad, tend to lead to forecasts that are overly or insufficiently disperse, degrading the signal's OOS R^2.

Our diagnostic portfolio, on the other hand, requires no quantitative prediction of the signal's performance.[12] The size of the bet on the market is based on a constant risk appetite

[11]If s does not have an economically interpretable scale itself, the scaling in step iii) has no economic interpretation either. Our diagnostic portfolio is designed to mimic this coefficient portfolio, except that we replace the in-sample mean with the expanding mean for implementability, adjust the scaling in iii) for economic interpretability, and impose a position constraint that is inevitable in reality.

[12]If one is evaluating signals *individually*, it is not obvious why such quantitative prediction is necessary

(5% in our example) and does not change with the past returns of the signal's coefficient portfolio. This is consistent with the construction of cross sectional factors (e.g., HML, SMB, MOM, etc.), in which the stock weights also do not change with the past factor returns. Removing the intermediate step of predicting the signal's performance removes potential errors in that step. Therefore, the diagnostic portfolio-based performance metrics appear better than the OOS R^2s overall. A well-behaved diagnostic portfolio does depend on the stability of the signal's mean and volatility over time, and such stability appears much better supported by the data.

To more directly see why the diagnostic portfolio based metrics can be so different from the OOS R^2, we plot in Figure B1 the univariate expanding window regression coefficients of the aggregate stock market excess return on the signals and a constant. The signals are scaled to have a standard deviation of 1 to make the coefficients comparable. As explained above, degradation in OOS R^2 from in-sample R^2 is in large part owing to these coefficients not being a constant. As the sample size grows over time, the coefficients become increasingly stable, resembling the full sample coefficients which are the last observation in each series. Hence, the largest deviations occur at the beginning of the sample. From this figure, it is immediately obvious that csp, aem, rrz, vrp, gm, and svix are the 6 signals with large and persistent deviations in the coefficients. They are precisely the 6 signals with the lowest OOS R^2.

It is worth emphasizing that we do not imply that the metric of OOS R^2s used in Goyal and Welch (2008) is inappropriate. On the contrary, we believe it is a good, consistent, and general metric. We do want to point out that there are more to this metric. Importantly, poor OOS R^2s do not imply that the corresponding signal cannot be successfully employed in real time trading. The diagnostic portfolio we propose is a simple approach that both complements the OOS R^2s and fills an obvious gap between the practices of two literatures studying the time-series and the cross-sections of stock returns.

in the first place.

Table B1: R^2 in forecasting monthly stock market returns using other signals

	IS	CTOOS10	CTOOS5	OOS10	OOS5	Return5	Sharpe5	Begin	End
dy	0.68%	-0.28%	-0.28%	-0.48%	-0.48%	-0.19%	-0.08	1871m2	2019m12
ep	0.88%	0.27%	0.27%	0.28%	0.28%	0.16%	0.09	1871m1	2019m12
bm	1.04%	0.16%	0.07%	-3.02%	-2.93%	-0.14%	-0.09	1921m3	2019m12
svar	0.25%	0.03%	0.03%	-1.49%	-1.49%	0.06%	0.04	1885m2	2019m12
csp	0.90%	-0.31%	0.24%	-5.81%	-6.12%	-0.16%	-0.10	1937m5	2002m12
ntis	0.75%	-0.18%	0.02%	-0.13%	-0.78%	0.10%**	0.24	1926m12	2019m12
eqis	0.29%	-0.43%	-1.62%	-0.40%	-2.11%	-0.06%	-0.11	1926m10	2008m4
tbl	0.23%	-0.80%	-0.61%	-1.70%	-1.61%	0.08%	0.05	1920m1	2019m12
lty	0.22%	-1.33%	-1.08%	-1.64%	-1.39%	0.14%	0.07	1919m1	2019m12
dfy	0.49%	-0.22%	-0.22%	-2.45%	-2.24%	-0.24%	-0.16	1919m1	2019m12
dfr	0.37%	-0.30%	-1.72%	-0.21%	-1.69%	0.14%	0.12	1926m1	2019m12
infl	0.35%	-0.05%	-0.05%	-0.03%	-0.03%	0.10%	0.10	1913m2	2019m12
ik	1.09%	0.30%	0.30%	-0.03%	-0.01%	0.37%*	0.21	1947m3	2019m12
cay	0.37%	-0.32%	-0.31%	-0.39%	-0.97%	0.25%	0.16	1947m4	2019m12
tms	0.29%	-0.10%	-0.20%	-1.04%	-1.32%	0.14%	0.09	1920m1	2019m12
aem	0.31%	-2.12%	-2.01%	-40.25%	-38.18%	-0.02%	-0.01	1947m4	2009m4
hkm	0.60%	-0.09%	0.00%	-0.28%	-0.30%	-0.03%	-0.02	1970m1	2012m12
rrz	1.45%	-10.18%	-12.30%	-9.83%	-12.00%	0.87**	0.43	1973m1	2014m12
kp	1.05%	0.31%	0.23%	0.36%	0.27%	0.19%***	0.37	1930m1	2019m11
svix	0.21%	-5.04%	-6.21%	-7.26%	-8.67%	-0.27%	-0.12	1996m1	2012m1
gm	0.78%	0.60%	0.36%	-0.52%	-5.20%	0.07%	0.04	1996m1	2012m1
vrp	0.10%	-0.67%	-0.53%	-17.98%	-14.02%	-0.09%	-0.06	1989m12	2019m12
sent	0.39%	-0.10%	-0.64%	-0.01%	-0.67%	0.02%	0.02	1965m7	2019m12

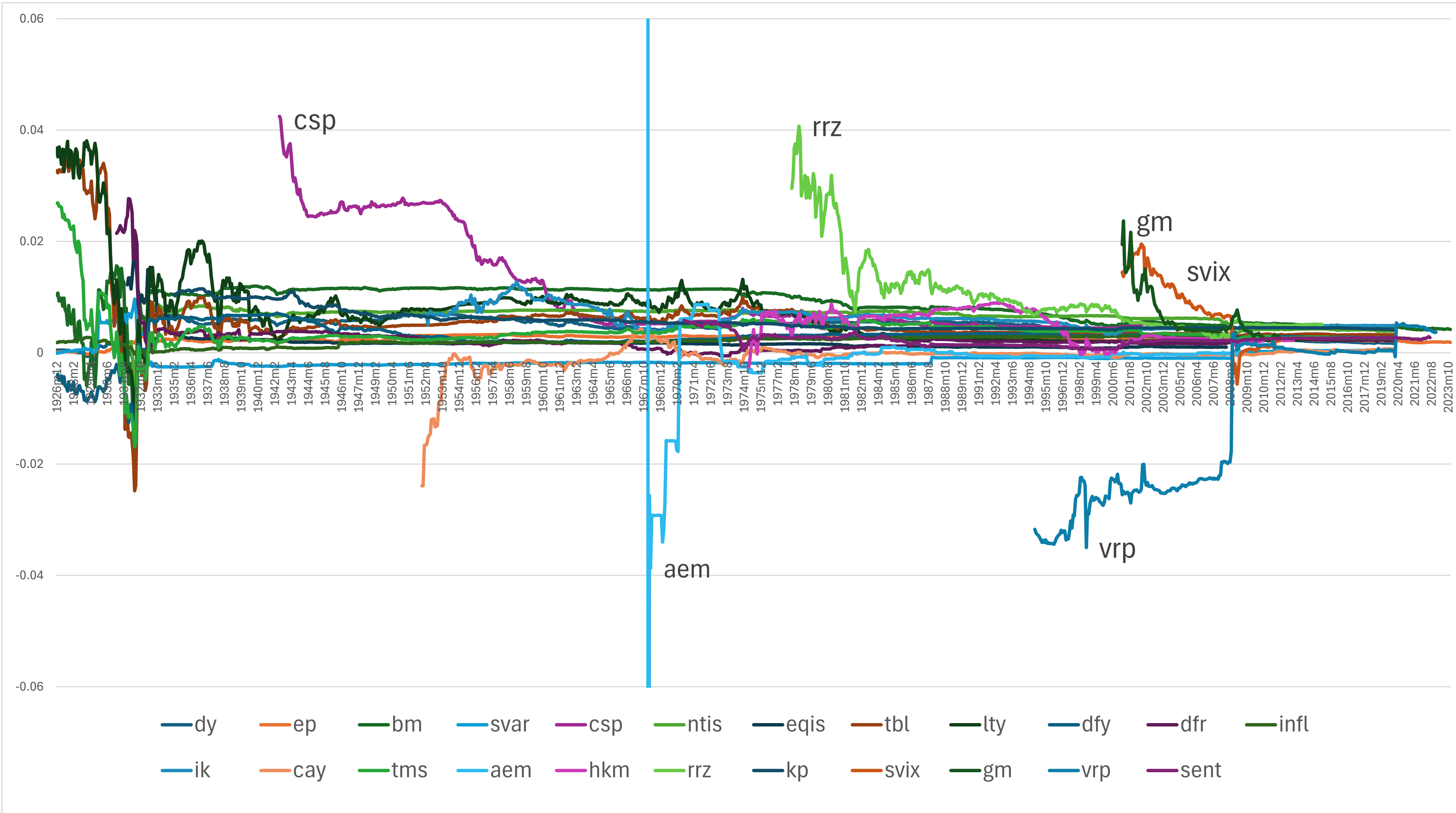
This table displays R^2 and the sample start/end month of various monthly signals that predict monthly stock market returns in the US. The "IS" column shows the in-sample R^2s, the next 4 columns show the out-of-sample R^2s as in Goyal and Welch (2008), and the next 2 columns show the monthly returns and annualized Sharpe ratios of the diagnostic portfolios. Columns with the 5 and 10 suffixes evaluate the R^2s 5 and 10 years after the date's inception (but not after 1926m1). Columns with the "CT" prefix i) truncates the estimated coefficient on the signal at 0 if it has a different sign from that estimated on the full sample and ii) truncates the forecast at the risk-free rate if it exceeds the risk-free rate, as in Campbell and Thompson (2008). All predictors are sourced from Amit Goyal's website (and follow his naming convention) except for the following: eqis is the new equity issuance share from Baker and Wurgler (2000), aem is the financial factor from Adrian et al. (2014), hkm is the leverage factor from He et al. (2017), rrz is the short interest from Rapach et al. (2016), svix is from Martin (2017), gm is the implied cash flow growth from Gao and Martin (2021), sent is the orthogonalized investor sentiment from Baker et al. (2021). I thank Li and Wang (2023) for sharing the data they use, and all the mentioned authors for making their data available online.

Table B2: Correlation among evaluation metrics for aggregate stock market return predictors

	Ret5	Sharpe5	IS	OOS5	CTOOS5
Ret5	1.00				
Sharpe5	0.90	1.00			
IS	0.50	0.51	1.00		
OOS5	0.03	0.12	0.12	1.00	
CTOOS5	-0.50	-0.32	-0.26	0.34	1.00

This table tabulates the correlation among different measures across the 23 signals in Table B1.

Figure B1: The expanding-window averages of the signals' coefficient portfolio returns



This figure plots the expanding window regression coefficients of the aggregate stock market excess returns on the 23 signals in Table B1.