Variance Dynamics Across the Information Cycle

Liuren Wu* Yaofei Xu†

Abstract

This paper uses historical variance term structures to characterize the different phases of a company's information cycle, and shows that the up and down phases of the information cycle experience different durations, with the up phase fast but the down phase slow. The paper proposes a conditional variance dynamics to capture the dynamics variation across the information cycle, and develops a two-dimensional conditional pooling estimation approach that balances the needs for reducing estimation errors and accommodating conditional dynamics variation. As an application, the paper finds that classic asset pricing relations can vary strongly across the information cycle.

JEL Classification: C13; C21; C51; C52; C53; C55; C58; G12; G13

Keywords: Historical variance term structure; information cycle; conditional variance dynamics; conditional pooling estimation; conditional asset pricing.

We thank Linda Allen, Torben Anderson, Aaron Barkley, Gerald Hanweck, Jian Hua, Jonathan Lewellen, Vance Martin, Francisco Peñaranda, Kevin Robik, Yong Song, Jun Wang, Tomasz Wozniak, Yildiray Yildirim, Yuzhao Zhang, and seminar participants at the 2022 African Meeting of the Econometric Society, the 2022 Asian Meeting of the Econometric Society, Baruch College, and the University of Melbourne for their comments and suggestions. Liuren Wu gratefully acknowledges the support by a grant from the City University of New York PSC-CUNY Research Award Program.

*Zicklin School of Business, Baruch College, The City University of New York, One Bernard Baruch Way, New York, NY 10010, USA. Tel: 646-312-3509; E-mail: liuren.wu@baruch.cuny.edu.

†School of Mathematics and Physics, Research Institute of Quantitative Finance, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China. E-mail: yaofei.xu@xjtlu.edu.cn.

1. Introduction

Stock return variance reflects the intensity of the information flow about the underlying company. The information intensity increases with both intensified trading activity (Andersen (1996), Ané and Geman (2000), He and Velu (2014), Lamoureux and Lastrapes (1990), and Tauchen and Pitts (1983)) and increased market pricing sensitivity to the order flow (Berger, Chaboud, and Hjalmarsson (2009)). Carr and Wu (2004) propose to capture the effect of stochastic variance through the concept of stochastic time change. The time change maps the fixed calendar clock to a stochastically varying business clock. The speed at which the business clock runs defines the variance rate and reflects the intensity of the underlying business activity. The well-documented "volatility clustering" behavior (Mandelbrot (1963), Engle (1982), Ding and Granger (1996)) suggests that the information intensity variation experiences persistent up and down cycles.

In this paper, we propose to use historical variance term structures to capture the local trend of the variance variation and to characterize the different phases of the information cycle. We construct the term structure as the log difference between historical variance estimators at short and long horizons, and we use two segments of the historical variance term structure, from one to three months and from three to 12 months, respectively, to capture the phase variation of the information cycle. Across the US stock universe over the past century, we find that not only do the historical variance term structures vary strongly over time for a given company, but they also show strong cross-sectional variations at any given date across the different companies. The information cycles of different companies are anything but synchronous.

More importantly, we document that the up and down phases of the information cycle tend to experience different durations, with the up phase fast and dramatic but the down phase slow and protracted, a phenomenon we characterize as "elevators up, stairs down." As a result of the

duration variation, the autoregressive volatility clustering behavior of the variance process varies strongly across different phases of the information cycle. In particular, we show that future realized variance can be reliably predicted by near-term historical variance estimators during the protracted downward moving phase, but the prediction becomes highly unreliable during the fast upward moving phase.

Historically, the literature has proposed a long list of variance forecasting specifications based on its autoregressive behaviors.[1] Our findings suggest that the autoregressive behavior is not fixed either over time or across companies, but can vary strongly across the different phases of the information cycle. Based on this observation, we proposes a new model of stock return variance forecasting dynamics that allows conditional dynamics variation across the information cycle. Specifically, we forecast the future realized variance of a company with its own historical variance estimators over different horizons via a centered specification on the natural logarithm of the variance estimators. The specification is similar in structure to the heterogeneous autoregressive (HAR) time-series model of Corsi (2009), but with modifications tailored to our particular application of predicting stock return variance over a large and ever changing stock universe. Most importantly, we allow the loading coefficients of the autoregressive specification to vary conditionally as a function of its historical variance term structure, capturing the dynamics variation across the information cycle.

To estimate the conditionally varying variance dynamics, we develop a two-dimensional conditional pooling estimation approach. The approach achieves strong identification through both cross-sectional and time-series pooling; meanwhile, it accommodates the conditional dynamics variation as a function of the information cycle through conditional weighting. The approach in-

[1] Prominent examples include the classic ARCH specification of Engle (1982), the GARCH specification of Bollerslev (1986), the EGARCH specification of Nelson (1991), the asymmetric GARCH of Glosten, Jagannathan, and Runkle (1993), and the stochastic variance rate model of Heston (1993).

volves two steps. First, at each date, we perform conditionally weighted cross-sectional forecasting regressions on the conditional dynamics across the large stock universe. The stock universe is not only very large, but also changes constantly over time. The median life span of a stock in our sample universe is merely 57 months. Name-by-name time-series estimation becomes unreliable for a large portion of the names with short histories. By starting with a cross-sectional identification of the conditional dynamics, the approach benefits from the large cross section while circumventing the difficulties for names with short histories.

To perform the conditionally weighted cross-sectional regression estimation, we use two segments of the historical variance term structures, from one to three months and from three to 12 months, respectively, as the conditioning variables. We generate the conditional weighting with a bivariate Gaussian kernel on the rank differences of the term structures. Estimation shows that the conditionally weighted loading coefficient estimates on the short-term historical variance estimator tend to be highly positive when a company's information cycle is in the protracted downward moving phase, but the conditional estimates decline and can even become negative when the information cycle is in the uprising phase. When the historical variance has been increasing rapidly, it becomes highly uncertain on what happens next to the company. The near-term historical variance estimator becomes an unreliable forecast of future variance. The specification needs to rely more on the longer-term estimators to generate a more conservative and less informative variance forecast.

In the second step, to generate out-of-sample variance forecasts at a given date, we perform conditional weighting on the historical coefficient estimates from the cross-sectional regressions. By averaging the historical variance term structure across the stock universe at each date, we construct the market variance term structure to capture the information cycle of the aggregate market. We use the market variance term structure as the conditioning variables for the time-

series conditional weighting of the coefficient estimates. The conditional weighting assigns higher weights to historical time periods with similar market conditions in generating the forecasting dynamics.

The conditional dynamics specification captures the variance dynamics variation across the information cycle. The conditional weighting estimation approach reduces estimation errors via two-dimensional pooling while also capturing dynamics variations both across names at different phases of their information cycle and over time as the aggregate market condition changes. A historical analysis of nearly a century worth of the US stock market data highlights the importance of both cross-sectional and time-series pooling in reducing estimation errors. Meanwhile, the analysis identifies large time-series and cross-sectional variations in the stock variance dynamics across the information cycle. An out-of-sample forecasting exercise shows that removing pooling along either the time-series or the cross-sectional dimension can lead to complete breakdown in the out-of-sample forecasts due to increased estimation errors. On the other hand, allowing variance dynamics variation across both dimensions through the conditional pooling approach can significantly increase the out-of-sample forecasting R^2 compared to a panel estimation that ignores dynamics variations in the two dimensions.

In principle, the variance dynamics can also vary across firms due to other characteristic differences. As a robustness check, we examine a long list of firm characteristics as alternative choices of conditioning variables. The list includes the stock return beta, company size, book-to-market equity value ratio, momentum, and industry classification. While some of these characteristics can indeed differentiate the variance dynamics to some degree, none of them can differentiate the dynamics nearly as strongly as the historical variance term structures we have constructed to capture the information cycle.

Given the strong variance dynamics variation across the information cycle, an important question to ask is whether and how capital asset pricing relations vary across the information cycle. As a preliminary exploration of this important topic, we estimate classic asset pricing relations conditionally with the historical variance term structure as the conditioning variable. We find that both the size effect and the momentum effect become significantly more positive for companies at the upward moving phase of the information cycle than for companies at the downward moving phase. The average beta effect on stock returns is negative in our sample and becomes increasingly so for companies at the upward moving phase. By comparison, the value effect is reasonably stable across the information cycle. High book-to-market companies generate higher average returns across all phases of the information cycle.

The rest of the paper is organized as follows. The next section documents the strong variance dynamics variation across the different phases of the information cycle. Section 3 proposes the conditionally varying variance forecasting dynamics and the two-dimensional conditional pooling estimation approach. Section 4 examines the conditional asset pricing variations across the information cycle. Section 5 concludes.

2. Variance dynamics variation across the information cycle

We perform the analysis using US daily stock return data from the Center for Research in Security Prices (CRSP) from December 1925 to December 2022.

2.1. Data sampling and stock universe selection

We sample the data once a month in the middle of the month. At each month, we choose the Wednesday 30 days from the third Friday of the next month as the sampling date. When the chosen Wednesday is a holiday, we use the previous business day as the sampling date. The general findings do not depend on the particular sampling choice. The third Friday of next month is the common expiration date for the front-month exchange-traded options on individual stocks in the US. Therefore, under this particular sampling approach, the one-month variance forecasts can be directly applied to the valuation of front-month 30-day stock options.

On each sampling date, we determine the stock universe based on the following filtering criteria. First, the underlying security is a common stock traded in one of the main exchanges (AMEX, NYSE, and NASDAQ). Second, the stock has daily returns available over the past 252 business days and the next month so that we can construct the historical variance estimators over different historical horizons up to one year and the future realized variance estimators over the next month. Third, the stock price level both at the sampling date and the average over the past year are no smaller than \$5 so that the return calculations are not unduly affected by bid-ask bounces.

The filtering generates a total of 2,373,320 company-month observations, spanning 20,462 companies over 1,152 months (96 years) from December 1926 to November 2022. Figure 1 plots the number of selected stocks at each month in panel A. The size of the cross section varies from 191 names in June 1932 to 5,053 names in October 1997, with an average of 2,060 names per sampling date. The number of selected stocks is small in the early years, but increases with time, over 500 since the 1940s, over 1,000 since the mid 1960s, and over 2,000 since the mid 1970s.

[Figure 1 about here.]

Panel B of Figure 1 plots the number of selected sampling months for each of the 20,462 selected stocks, sorted from low to high and represented in log scale. The data sample includes a large number of stocks, but the length of the time series varies greatly, from as few as one month (for 403 stocks) to as many as 1,141 months (for one stock), with an average of 116 months and a median of 57 months.

The large variation in the length of the time series per each stock dictates that name-by-name time-series analysis or variance dynamics estimation cannot be relied upon for a large portion of the stock universe with short time-series lengths. Therefore, the analysis in this paper switches from the traditional time-series perspective per each name to a more cross-sectional focus on the large stock universe at each sampling date.

2.2. Capture information intensity variation with variance estimators

Stock return variance reflects the intensity of the information flow about the underlying company. The literature has tried to link the information intensity variation to trading activities and market pricing sensitivities to the order flow. In this paper, we construct variance estimators from daily stock returns and use these variance estimators to quantify the information intensity variation of a company.

At each sampling month t, for each selected name i, we construct the annualized realized variance estimator over the next month $RV_{t+1,i}$ using daily simple-compounding stock returns $r_{t+d\Delta,i}$,

$$\mathrm{RV}_{t+1,i} = \frac{252}{n_t} \sum_{d=1}^{n_t} r_{t+d\Delta,i}^2 \,, \tag{1}$$

where the time t is measured in months, d denotes the business days, $\Delta = 1/21$ denotes the daily

time interval measured in business days, and n_t denotes the number of business days over the next month. For stock indexes and macro time series, the literature often uses log returns to construct the variance estimators. For individual stocks, however, the prices have a non-zero probability of dropping to zero upon default. Continuous compounding returns become explosive in such cases. We therefore use simple compounding return to construct the variance estimators.

We also construct historical variance estimators over different horizons h,

$$\mathrm{HV}_{t,h,i} = \frac{252}{h/\Delta} \sum_{d=0}^{\frac{h}{\Delta}-1} r^2_{t-d\Delta,i}, \tag{2}$$

where we choose the historical horizons from $h = 1$ to 3, 6, 9, and 12 months. Except for the front month, we choose quarterly segments as quarters represent natural time intervals for company financial reports and business analysis.

At each month, we compute the cross-sectional summary statistics of the historical variance estimators, including their cross-sectional percentile values, as well as the cross-sectional distributional skewness and excess kurtosis of both the variance estimators and the natural logarithm of the variance estimators. Table 1 reports the time-series averages of the cross-sectional summary statistics. Panel A reports the average percentile values in volatility percentage points. The average percentile values highlight the extremely large cross-sectional variation of the average volatility levels. For the one-month historical variance estimators, the average volatility level can vary from 10% at the 1st percentile to over 100% at the 99th percentile. For longer-term historical variance estimators, the cross-sectional range of variation declines, but only to a small degree.

[Table 1 about here.]

Panel B reports the time-series averages of the cross-sectional skewness and excess kurtosis

estimates on the variance estimators. The cross-sectional distributions of the variance estimators are highly positively skewed with extremely large excess kurtosis. The average skewness estimates range from 7.1 for the 12-month variance estimator to 9.7 for the one-month estimator, and the average excess kurtosis estimates range from 186.8 to 260.2.

Panel C shows that the cross-sectional distributions become much closer to a normal distribution by taking natural logarithm on the variance estimators. The average skewness and excess kurtosis estimates on the logarithm of the variance estimators are all close to zero.

2.3. Volatility clustering in the cross section

It is well known that the variance dynamics show autoregressive volatility clustering behavior. Future realized variance on a stock can be predicted by its historical variance estimators based on this autoregressive behavior. This section documents the volatility clustering behavior in the cross section of the large stock universe. In particular, Table 2 estimates the cross-sectional forecasting correlation between the historical variance estimators at different horizons (h) and the realized variance over next month. Entries report the time-series statistics of the cross-sectional forecasting correlation estimates, including the sample average (Mean), standard deviation (SD), minimum (Min), and maximum (Max).

[Table 2 about here.]

Panel A estimates the forecasting correlations between the historical and future realized variance estimators. The average forecasting correlation ranges from 37% with the one-month historical variance estimator to 45% with 6-12 month historical variance estimators. The highly positive average estimates suggest that the well-documented volatility clustering behavior in the time series

is also present in the cross-sectional setting of the stock universe. Nevertheless, the cross-sectional estimates show large time-series variation, with time-series standard deviation estimates at 16–18%. In particular, the minimum cross-sectional forecasting correlation can drop to zero.

Panel B estimates the forecasting correlation between the natural logarithm of the future and historical variance estimators. The forecasting correlation estimates in the log variance space have much higher sample averages at 65%–70% and with much smaller time-series standard deviation at 7%. The correlation estimates become much more uniform over time. Even in the worst days, the minimum correlation estimates can still be as high as 42%.

By taking natural logarithm, the forecasting correlation in panel B amounts to comparing proportional variance variation rather than variance level variation across names. As the variance level differs greatly across names, the proportional variation becomes much more comparable than the level variation, especially in the cross section. A 10-percentage point volatility variation represents a huge shock for a name with an average volatility level of 10%, but it represents only a very small variation for a name with an average volatility level of over 100%. By taking natural logarithms and comparing proportional changes, panel B captures a much more comparable and more stable forecasting relation in the cross section.

In panel C, we use the 12-month historical variance estimator $\mathrm{HV}_{t,12,i}$ as a benchmark predictor of the future realized variance, and examine the additional contribution of the shorter-term historical variance estimators. Specifically, we divide both the future realized variance and the historical variance estimators by the 12-month historical variance estimator benchmark, and compute the forecasting correlations between the natural logarithm of the variance ratios,

$$\mathcal{S}_{t+1,12,i} \equiv \left(\mathrm{RV}_{t+1,i}/\mathrm{HV}_{t,12,i}\right), \quad \mathcal{S}_{t,h,12,i} \equiv \ln\left(\mathrm{HV}_{t,h,i}/\mathrm{HV}_{t,12,i}\right). \tag{3}$$

The historical log variance ratio $\mathcal{S}_{t,h,12,i}$ can be regarded as the proportional historical variance trend estimator between the 12-month benchmark and the more recent horizon h, whereas the realized log variance ratio $\mathcal{S}_{t+1,12,i}$ extends the trend to the next month. The average forecasting correlation estimates on the log variance ratios in panel C are smaller than those in Panel B on the log variance, but they remain strongly positive, suggesting that the shorter horizon variance estimators can make additional forecasting contributions to the future realized variance in addition to the 12-month historical variance estimator benchmark. In particular, if we think of $\mathcal{S}_{t,h,12,i}$ and $\mathcal{S}_{t+1,12,i}$ as historical and future variance trend estimators, the positive forecasting correlation estimates represent another measure of the volatility clustering behavior and suggest that historical trends of variance movement tends to carry forward to the next month.

2.4. Capture the information cycle with historical variance term structures

With the variance estimators capturing the intensity of the information flow, the volatility clustering behavior, both over the time series of a single name and across the cross section of the stock universe, suggests that the information intensity variation experiences persistent up and down cycles. In this section, we propose to construct historical variance term structures to capture the local trend of the information intensity variation and the different phases of the information cycle.

At each sampling date t and for each name i, we construct two historical variance term structure segments,

$$\mathcal{S}_{t,1,3,i} \equiv \ln\left(\mathrm{HV}_{t,1,i}/\mathrm{HV}_{t,3,i}\right), \quad \mathcal{S}_{t,3,12,i} \equiv \ln\left(\mathrm{HV}_{t,3,i}/\mathrm{HV}_{t,12,i}\right), \tag{4}$$

where $\mathcal{S}_{t,1,3,i}$ measures the most recent trend of the variance movement and $\mathcal{S}_{t,3,12,i}$ measures the more distant trend of variance movement. In a forward-looking term structure such as that on the yield curve or the option implied volatility, the longer horizon looks further into the future and

the term structure is constructed as the slope between the longer horizon and the shorter horizon to capture the expected future direction of movement. By contrast, the horizon in our historical variance estimators is backward looking. Accordingly, we define the historical variance term structure as the proportional slope between the shorter horizon capturing the more recent history and the longer horizon capturing the more distant history. Under this construction, a positive term structure represents upward historical variance movement and a negative term structure reflects downward historical variance movement.

We use two segments, $\mathcal{S}_{t,1,3,i}$ and $\mathcal{S}_{t,3,12,i}$, to capture the shape variation of the variance term structure. To examine whether the two segments capture different information about the information cycle, we measure the cross-sectional correlation between the two term structure segments. The time-series average of the cross-sectional correlation estimates is very small at -3.9%, suggesting that the short and long ends of the variance term structure can show different movements and patterns. The two segments together present a more complete picture on the transition of the information cycle.

From the historical variance term structures on the individual stocks, we also construct the market variance term structure as the cross-sectional averages of the historical variance term structures across all names in the stock universe at each date,

$$\mathcal{S}_{t,1,3,m} = \frac{1}{N_t}\sum_{i=1}^{N_t}\mathcal{S}_{t,1,3,i}, \quad \mathcal{S}_{t,3,12,m} = \frac{1}{N_t}\sum_{i=1}^{N_t}\mathcal{S}_{t,3,12,i}, \tag{5}$$

with N_t denoting the time-varying size of the cross section at date t. While the individual term structures capture the information cycle variation of a single company, the market term structure captures the information cycle of the aggregate market. The two segments of the market variance term structure also show very low time-series correlation at 3.7%.

To examine the cross-sectional variation of the information cycle across the stock universe, we compute the cross-sectional percentile values of the historical variance term structure. Table 3 reports the time-series averages of the cross-sectional percentile values in panel A. Both segments of the variance term structures show large cross-sectional variations. On average, the short tend of the term structure $\mathcal{S}_{t,1,3,i}$ can vary strongly from highly downward sloping at -170.7% at the first percentile to highly upward sloping at 80.7% at the 99th percentile. The long end of the term structure $\mathcal{S}_{t,3,12,i}$ shows similarly large average cross-sectional variations, from -130.9% at the first percentile to 80% at the 99th percentile. The large cross-sectional variation of the variance term structures suggests that the information cycles of different companies are anything but synchronous. At each date, there are always companies whose stock return variance has been going up and others whose variance has been moving down.

For both segments, the median term structures average negatively at -7.7% and -8.4%, respectively. In the last row of the table, we also compute the proportion of companies with positive variance term structures (PTS), which average at 44.3% for $\mathcal{S}_{t,1,3,i}$ and 42.7% for $\mathcal{S}_{t,3,12,i}$. Both statistics suggest that a larger proportion of company-month observations show downward sloping variance term structures than upward sloping term structures. The statistics reflect the asymmetric nature of the variance dynamics and the duration difference between the up and down information cycles. The duration tends to be shorter for the upward moving phase but longer for the downward moving phase.

Panel B of Table 3 reports the time-series percentile values of the market variance term structures. The aggregate market variance term structure shows large time-series variation. The short end of the market term structure $\mathcal{S}_{t,1,3,m}$ varies from -69.4% at the first percentile to 57.8% at the 99th percentile. The long end of the market term structure $\mathcal{S}_{t,3,12,m}$ varies even more from -84.9% at the first percentile to 71.5% at the 99th percentile. Therefore, not only different companies ex-

perience different information cycles, the market as a whole also experiences an aggregate cycle of variance movements.

The duration differences between the upward moving phase and downward moving phase is even larger on the market information cycle. Only 20.1% of the short market term structure $\mathcal{S}_{t,1,3,m}$ and 28.2% of the long market term structure $\mathcal{S}_{t,3,12,m}$ are upward sloping, suggesting that for the aggregate market, the return variance tends to spike up even faster while calming down much more slowly.

2.5. Dynamics variation across different phases of the information cycle

Not only does the information intensity variation show persistent up and down cycles, but the different phases of the information cycle also experience different durations. The duration tends to be shorter for the upward moving phase but longer for the downward moving phase. As a manifestation, stock return variance tends to spike up quickly and then calm down slowly, a feature we label as "elevators up, stairs down."

The duration variation suggests that the variance forecasting relation can vary across different phases of the information cycle. To examine how the variance forecasting relation varies across the information cycle, we use the two segments of the variance term structure as conditioning variables, and classify each variable into three cross-sectional quantiles at each date, with the low quantile capturing companies at the down trend phase and the high quantile capturing companies at the up trend phase. For companies at each phase of the information cycle, we estimate the cross-sectional forecasting correlations between the historical log variance ratios $\mathcal{S}_{t,h,12,i}$ at different horizons and the one-month ahead log variance ratio $\mathcal{S}_{t+1,12,i}$. Table 4 reports the time-series averages of the cross-sectional forecasting correlation estimates conditional on the different phases

of the information cycle. The first row of the table repeats the average unconditional forecasting correlation estimates from Panel C of Table 2 as a benchmark.

[Table 4 about here.]

When both the long and the short ends of the variance term structure are downward sloping, the information cycle is in a stable and protracted downward moving phase. For companies in this downward moving phase of their information cycle, the average conditional forecasting correlation estimates are positive across all horizons. Compared to the unconditional estimates, the conditional estimates are stronger with historical log variance ratios at shorter horizons but weaker at longer horizons. The estimates suggest that during this phase, future realized variance can be forecasted more accurately with the near-term historical variance estimators.

When the long-term trend remains downward moving, but the near-term trend becomes upward moving, the information cycle is in a transition period from the downward moving phase to the upward moving phase. During this transition, forecasting future realized variance becomes harder, and the average forecasting correlation estimates at shorter horizons become lower.

As the information cycle propagates further, both the long and short ends of the historical variance term structure become upward sloping. The companies are in a highly unstable state. Historical log variance ratios lose all their forecasting powers. The average conditional forecasting correlation estimates are all close to zero or even negative. In such an unstable phase, one can no longer rely on autoregressive behaviors to generate future variance forecasts, but can only rely on a long-term variance estimator as a conservative, uninformative fall-back forecast.

Finally, when the long end of the term structure remains upward sloping but the short end of the term structure starts to come down, the short-term variance estimator becomes informative

again about the future realized variance. The average forecasting correlation with $\mathcal{S}_{t,1,12,i}$ becomes highly positive, but the average forecasting correlations with the longer end of the term structure remain small and negative.

As a company moves through its information cycle, the predictability of historical variance estimators on its future realized variance varies, and so does the structure of the underlying forecasting relation. As such, both the forecasting power and the forecasting relation can differ greatly across different phases of their information cycle.

3. Predict stock variance across the information cycle

The strong dynamics variation across the information cycle has important implications for variance forecasting. In particular, classic autoregressive variance forecasting specifications ignoring the dynamics variation across the information cycle are unlikely to generate accurate variance forecasts across all market conditions. In this section, we propose a conditionally varying variance dynamics to capture the dynamics variation across the information cycle. To estimate the conditional variance dynamics, we develop a two-dimensional conditional pooling estimation procedure that balances the need for reducing estimation errors and the need for capturing the dynamics variation across different phases of the information cycle.

3.1. Conditionally varying variance dynamics across the information cycle

To capture the dynamics variation across different phases of the information cycle, we propose the following conditionally varying variance dynamics,

$$\mathcal{S}_{t+1,12,i} = \sum_h \beta_h(\mathbf{x}_{t,i})\, \mathcal{S}_{t,h,12,i} + e_{t+1,i}\,. \tag{6}$$

The conditional variance dynamics specification in (6) predicts the one-month ahead realized log variance ratio $\mathcal{S}_{t+1,12,i}$ of a company i with its historical log variance ratios $\mathcal{S}_{t,h,12,i}$ over different historical horizons h. Mostly importantly, the specification does not impose constant loading coefficients either over time or across different companies, but allow the coefficients to vary conditionally with the company's information cycle as captured by the two segments of its historical variance term structure, $\mathbf{x}_{t,i} = [\mathcal{S}_{t,1,3,i}, \mathcal{S}_{3,12,i}]$.

Beside the conditionally varying loading coefficients $\beta_h(\mathbf{x}_{t,i})$, the general structure of the dynamics specification in (6) is in line with recent advancements in variance forecasting specifications, e.g., Corsi (2009), Ghysels, Santa-Clara, and Valkanov (2006) and Bollerslev, Hood, Huss, and Pedersen (2018), where future realized variance is directly forecasted with historical variance estimators across different historical horizons. The unconditional analog of the specification is similar to the heterogeneous autoregressive (HAR) time-series model of Corsi (2009), but with modifications tailored to our particular application. First, we replace the additive specification of the HAR model with a multiplicative specification because proportional variance variations are much more comparable across the large universe of stocks, thus facilitating cross-sectional pooling in the dynamics estimation. Second, the HAR model has a constant term to capture the reversion to a long-run mean variance level. Our specification in (6) is centralized with the 12-month historical variance estimator so that the future realized variance of each name is predicted by its own

historical variance estimators. The additive HAR specification forecasts the future variance with a weighted average of historical variance estimators and a constant. The centralized multiplicative specification in (6) predicts the natural logarithm of future variance of each name i as a weighted average of the natural logarithm of its own historical variance estimators,

$$\mathbb{E}_t\left[\ln RV_{t+1,i}\right] = \left(1 - \sum_h \beta_h(\mathbf{x}_{t,i})\right) \ln \mathrm{HV}_{t,12,i} + \sum_h \beta_h(\mathbf{x}_{t,i}) \ln \mathrm{HV}_{t,h,i} \ . \tag{7}$$

In their panel estimation of the variance dynamics across several equal-length macro time series, Bollerslev, Hood, Huss, and Pedersen (2018) specify a centralized additive HAR model with the recursive full-sample mean variance estimator of each series as the center. An recursively expanding window in our case, however, would lead to different lengths for companies with different lengths of histories, making the center less comparable across names. The choice of 12 months as the long-run benchmark balances several considerations. The length is long enough to capture the average variance level of a stock. The 12-month historical variance estimator has been a common default choice as a variance predictor in many applications, e.g., Goyal and Saretto (2009), Frazzini and Pedersen (2014), and Bai and Wu (2016). Meanwhile, the choice is short enough to include not only companies that have been around for many years but also new companies that have been traded on the exchange for just 12 months.

While the conditional variance dynamics in (6) can vary both over time and across stocks as a function of the company's information cycle, some pooling is necessary to achieve identification. The next subsection proposes a two-dimensional conditional pooling estimation procedure that balances the need for capturing the conditional dynamics variation and the need for robust identification.

3.2. Two-dimensional conditional pooling estimation

The accuracy of variance prediction depends on both the validity of the dynamics specification and the accuracy of the model parameter estimation. While the dynamics can vary both over time and across different securities, estimation inevitably involves some pooling across either or both dimensions in an effort to reduce estimation errors. Traditionally, variance forecasting tends to focus on time-series dynamics specification and rolling-window estimation per each name. This paper faces the unique challenge of predicting stock return variance in a large and ever-changing universe. The large cross section makes the name-by-name estimation cumbersome, and the ever-changing universe dictates that the historical length can vary greatly across names, making the time-series estimation unreliable for a large portion of the names with short histories.

To overcome this challenge, we propose to replace the classic time-series perspective with a more cross-sectional orientation, and develop a two-dimensional conditional pooling estimation procedure that balances the need for reducing estimation errors with the need for allowing both cross-sectional and time-series dynamics variation. The procedure starts with a conditionally weighted cross-sectional regression estimation of the specified dynamics in (6) over the large and ever-changing stock universe at each date. We then perform conditionally weighted averaging on the cross-sectional regression coefficient estimates over a rolling historical window to generate the forecasting dynamics. We discuss the details of each step in the following subsections.

3.2.1. The average unconditional variance dynamics

Before jumping into conditional estimation, we first examine the average variance dynamics behavior through unconditional estimation ignoring the information cycle variation. We start by imposing common loading coefficients across all stocks in the universe at each date and perform

unconditional cross-sectional ordinary least square regression estimation of the forecasting relation. We then perform unconditional time-series averaging on the cross-sectional regression estimates to obtain the average variance dynamics. The average coefficient estimates obtained from such a two-step procedure are similar to those obtained from a panel regression (e.g., Bollerslev, Hood, Huss, and Pedersen (2018)), but the two-step procedure is more flexible and better suited for estimation on the large and ever-changing stock universe.

Table 5 reports the time-series averages of the unconditional loading coefficient estimates and the Newey and West (1987) t-statistics on the averages (in parentheses), computed with an optimal lag chosen according to Andrews (1991). We consider four model specifications with increasingly more historical log variance ratios as predictors. Across all model specifications, the average loading coefficient β_1 on the log variance ratio $\mathcal{S}_{t,1,12}$ is the largest. As the horizon h increases, the average loading coefficient estimate declines. In model IV, the average loading β_9 on the log variance ratio $\mathcal{S}_{t,9,12}$ becomes very small and is no longer statistically significant.

[Table 5 about here.]

The last column of Table 5 reports the time-series averages of the cross-sectional regression R^2 estimates. By using the 12-month historical variance as the benchmark to define the log variance ratio and to centralize the cross-sectional variance forecasting relation, equation (6) effectively uses the 12-month historical variance estimator as the default forecast for the future realized variance under the null hypothesis that the loading coefficients on the historical log variance ratios are zero: $\beta_{t,h} = 0$. The R^2 estimate of the regression measures the outperformance of the model in predicting future realized variance against this benchmark prediction. The absolute magnitudes of the average R^2 estimates depend on the performance of the benchmark prediction. The better the benchmark prediction, the smaller the outperformance and hence the smaller the average R^2

estimates. With the 12-month historical variance estimator as the benchmark, the full-sample outperformance estimates average from 17% for model I to 22.1% for model IV.

In brackets under the average R^2 of each model j reports the Diebold and Mariano (1995) (DM) t-statistics on the outperformance of the best performing model IV against the model j, $\delta_{t,j} = R^2_{t,\mathrm{IV}} - R^2_{t,j}$. The DM test statistics show that in sample, the outperformances of model IV against the more parsimonious model specifications I to III are all statistically significant.

3.2.2. Conditional dynamics variation across different companies

The unconditional cross-sectional regression estimation in Table 5 assumes that the proportional variance dynamics are identical across the large universe of companies at any given date. This commonality assumption allows us to impose cross-sectional pooling over the large stock universe to reduce estimation errors, but it completely ignores the strong dynamics variation across different phases of the information cycle.

To capture the dynamics variation across the information cycle while also enjoying the benefit of cross-sectional pooling, we develop a conditional cross-sectional pooling estimation procedure to estimate the conditional variance dynamics in (6). Specifically, at each date t, we discretize the cross-sectional ranking of the conditional variables by transforming $\mathbf{x}_t$ into quantiles $\mathbf{q}$. We use the convolution of the quantiles to create a fixed grid of target conditioning quantile levels. For each target quantile grid δ_i, we estimate the conditional relation by applying conditional weighting on each name j in the universe based on its quantile conditioning variable differences from the target quantile grid,

$$w_{j,i}(\delta) = \mathcal{K}\left(\|\mathbf{q}_j - \delta_i\|\right). \tag{8}$$

with $\mathcal{K}(\cdot)$ denoting a kernel function operating on the quantile differences.

In this setting, the dynamics estimation for each quantile of names uses the whole cross section to reduce estimation error. The dynamics differentiation across different conditioning quantiles comes from the different conditional weighting on the estimation. The choices of the conditioning variables determine the sources of dynamics differentiation. The choice of the kernel function and bandwidth balances the tradeoff between cross-sectional smoothing and dynamics differentiation.

The approach is general, but its success depends crucially on choosing conditioning variables that can truly differentiate the dynamics. The sources of differentiation can vary across markets and applications. In our particular application to stock return variance prediction, we have identified the historical variance term structure as the key conditioning variable that can differentiate the variance dynamics across different phases of the information cycle. We use two segments of the historical variance term structure $\mathbf{x}_t$ as conditioning variables to capture the phase variation of the information cycle.

We apply a bivariate Gaussian kernel on the quantile values of the conditioning variables to generate the conditional weighting,

$$w_{j,i}(\delta) = \mathcal{K}\left(\|\mathbf{q}_j - \delta_i\|\right) = \exp\left(\frac{-(\mathbf{q}_j - \delta_i)^\top \Sigma_q^{-1}(\mathbf{q}_j - \delta_i)}{2\xi}\right), \tag{9}$$

where Σ_q denotes the sample covariance matrix for the bivariate quantile series $\mathbf{q}$ and ξ denotes the bandwidth, set to be a multiple (m) of the default choice (Simonoff 1996),

$$\xi = m\left(\frac{4}{3N_t}\right)^{(1/5)}. \tag{10}$$

We convert each variable into decile ranks for a convolution of 100 quantile grids. A more refined quantile grid can enhance the cross-sectional dynamics differentiation, but at the expense of in-

creased computational burden. We choose a bandwidth multiple of $m = 2$ for our main analysis. A robustness analysis in a later section shows that the conditional regression generates stable out-of-sample performance as long as the bandwidth choice is not too small compared to the default choice.

To highlight how the variance dynamics vary across different phases of the information cycle, Table 6 takes model II as an example and reports the time-series averages of the conditional coefficient estimates $\beta_1(\mathbf{x})$ and $\beta_3(\mathbf{x})$. The first column and the first row report the average decile values of the two conditioning variables, respectively. The average short end of the term structure $\mathcal{S}_{1,3}$, reported in the first column, varies from -118% at the 1st decile to 61% at the 10th decile. The average long end of the term structure $\mathcal{S}_{3,12}$, reported in the first row, varies from -91% at the 1st decile to 56% at the 10th decile.

[Table 6 about here.]

Under model II, the two conditional loading coefficients $\beta_1(\mathbf{x})$ and $\beta_3(\mathbf{x})$ measure the conditional weight contributions from the two historical variance estimators HV_1 and HV_3, respectively, with $1-\beta_1(\mathbf{x})-\beta_3(\mathbf{x})$ being the conditional weight on the 12-month estimator HV_{12}. Panel A reports the time-series averages of the conditional estimates on $\beta_1(\mathbf{x})$. The variation of the coefficient estimates across the conditioning grid shows a stronger dependence on the short end of the term structure $\mathcal{S}_{1,3}$. The average loading estimates are highly positive when $\mathcal{S}_{1,3}$ is downward sloping, but decline and even become negative as the term structure flattens and eventually becomes upward sloping.

The long end of the term structure $\mathcal{S}_{3,12}$ does not have a monotonic impact on the loading estimates, but it changes the sensitivity of the loading estimates to the short end of the term structure. The sensitivity increases as the long end of the term structure increases and moves from downward

to upward sloping. Eventually, when both the short and the long end of the term structure become highly upward sloping at the bottom right corner of the panel, the information cycle enters into an unstable upward moving phase, the average loading estimate on HV_1 becomes strongly negative. Compared to the average unconditional estimate of 0.277 in Table 5, the average conditional estimate can become as high as 0.48 during the transition phase from an upward sloping term structure to a downward sloping term structure at the top right corner of the panel. It can also become as negative as -0.29 during the unstable upward moving phase at the bottom right corner of the panel.

Panel B reports the time-series averages of the conditional estimates on $\beta_3(\mathbf{x})$. The variation of $\beta_3(\mathbf{x})$ shows a stronger dependence on the long end of term structure $\mathcal{S}_{3,12}$. The average estimates are strongly positive when the long end of the term structure is downward sloping, but can also become negative when the long end of the term structure becomes upward sloping. The effect of the short end of the term structure $\mathcal{S}_{1,3}$ is in the opposite direction: The loading estimate increases as the short end of the term structure increases. Compared to the average unconditional estimate of 0.225 in Table 5, the average conditional estimate can become as high as 0.61 at the bottom left corner of the panel during the transition phase from a downward moving cycle to an upward moving cycle, but it can also become negative at the top right corner during the transition phase from an upward moving cycle to a downward moving cycle.

The average conditional estimates for $\beta_1(\mathbf{x})$ and $\beta_3(\mathbf{x})$ show a remarkable range of variation. The deviations from their corresponding average unconditional estimates (β_1 and β_3) are not only very large, but also highly statistically significant. The Newey and West (1987) t-statistics on the average differences range from -18 to 16 for $\beta_1(\mathbf{x}) - \beta_1$ and from -15 to 19 for $\beta_3(\mathbf{x}) - \beta_3$. The variance dynamics can vary strongly across different phases of the information cycle. The conditional weighting estimation captures the cross-sectional dynamics variations across names at different phases of their information cycle while also enjoying the benefit of cross-sectional

pooling over the large stock universe to reduce the estimation errors.

3.2.3. Conditional dynamics variation across the market information cycle

When making out-of-sample variance forecasts at a given date t, we can achieve time-series pooling in a second step by averaging the historical coefficient estimates from the conditionally weighted cross-sectional regressions. Furthermore, the market as a whole also has an information cycle that propagates at different speeds at different phases. To capture this time variation, we propose to apply conditional weighting to the time-series averaging of the historical coefficient estimates, analogous to the conditional weighting in the cross-sectional regression estimation at each date.

Specifically, we characterize the market condition using the two segments of the market variance term structure, $\mathbf{m}_t = (\mathcal{S}_{t,1,3,m}, \mathcal{S}_{t,3,12,m})$. We construct the conditional weighting based on the ranking difference between the historical market conditions and the conditions at the forecasting date t,

$$w_{s,t} = \mathcal{K}(\|r(\mathbf{m}_s) - r(\mathbf{m}_t)\|)\,, \tag{11}$$

for $s = t-L, \cdots, t-1$, with L denoting the length of the historical rolling sample. Similar to the cross-sectional conditional weighting, we apply a bivariate Gaussian kernel with a bandwidth multiple of $m = 2$. Different from the cross-sectional weighting, we do not need to group the market conditions into quantile groups and can instead directly measure the ranking difference between the historical market conditions and the conditions at the forecasting date.

To examine how the average variance dynamics vary over time with market conditions, we take model II as an example and perform unconditional cross-sectional regressions on the variance forecasting relation to obtain the average variance dynamics estimates at each date. We then gen-

erate conditionally weighted time-series averages on the historical coefficient estimates over the whole sample according to the conditioning weighting defined in (11) at different percentile values of the two market variance term structure segments as conditioning variables. Figure 2 plots the conditional dependence structures of the two loading coefficients.

[Figure 2 about here.]

The conditional averages of $\beta_1(\mathbf{m})$ vary mainly with the short end of the market variance term structure $\mathcal{S}_{t,1,3,m}$. Panel A of Figure 2 plots the conditional averages of $\beta_1(\mathbf{m})$ as a function of the percentile values of $\mathcal{S}_{t,1,3,m}$. The three lines are at three different percentile values of the long end of the market term structure $\mathcal{S}_{t,3,12,m}$ at 10% (dash-dotted line), 50% (solid line), and 90% (dashed line), respectively. The three lines are all strongly downward sloping, as the conditional averages decline with the short end of the market variance term structure $\mathcal{S}_{t,1,3,m}$. The full-sample unconditional average of β_1 is 0.28. The conditional averages can be much higher when the short end of the market term structure is downward sloping but can also become lower when the short end of the market term structure becomes upward sloping.

The conditional averages of $\beta_3(\mathbf{m})$ vary mainly with the long end of the market variance term structure $\mathcal{S}_{t,3,12,m}$. Panel B of Figure 2 highlights this dependence structure by plotting the conditional averages against percentile values of $\mathcal{S}_{t,3,12,m}$. The three lines are at three different percentile values of the short end of the market term structure $\mathcal{S}_{t,1,3,m}$ at 10% (dash-dotted line), 50% (solid line), and 90% (dashed line), respectively. All three lines are downward sloping as the conditional averages decline rapidly with increasing market variance term structure at the long end. The full-sample unconditional average of β_3 is 0.22. The conditional estimates can be twice as large when the long end of the term structure is downward sloping but can also drop below zero when the long end of the term structure becomes highly upward sloping.

With the conditional time-series weighting, we can use the historical average of the forecasting coefficients to reduce estimation errors while also applying different weighting to different historical time periods based on the affinity of the phases of the market information cycle to that at the forecasting time.

3.3. Out-of-sample forecasting performance comparison

To gauge the contribution of the information cycle variation to the out-of-sample variance forecasting performance, we perform an out-of-sample forecasting exercise starting in December 1936, 10 years from the beginning of the sample. At each date t, we perform conditional averaging on the historical conditional cross-sectional coefficient estimates over the past 10 years, and apply the average coefficients on the date-t historical variance estimators to generate the out-of-sample forecasts on the realized variance over the next month. The 10-year rolling window represents a sufficiently long time period to perform the time-series pooling. We have repeated the analysis with an expanding window. The results are similar.

Table 7 summarizes the out-of-sample forecasting R^2 estimates for four model specifications and four estimation approaches for each specification. A positive R^2 estimate indicates out-of-sample outperformance against the benchmark prediction with the 12-month historical variance estimator. The four model specifications from I to IV have progressively more historical variance segments as predictors. For each model specification, the four estimation approaches perform either equal weighted (EW) or conditionally weighted (CW) pooling along the cross-sectional (CS) and time-series (TS) dimensions. The equal-weighted estimation amounts to ignoring the information cycle variation and assuming constant loading coefficients either across the stock universe in the cross section or over different time periods in the time-series aggregation. In brackets are the

DM t-statistics on the outperformance of the best performer against the particular estimate.

[Table 7 about here.]

The first estimation approach performs unconditional equal-weighted pooling along both the cross-sectional and the time-series dimensions, similar to a panel estimation ignoring the information cycle variation both across different companies and over different market conditions. The approach creates a benchmark for the contribution of allowing information cycle variation along each of the two dimensions. The out-of-sample R^2 estimates are positive for all four model specifications. Out of sample, adding one term structure segment ($\mathcal{S}_{1,12}$) with simple unconditional pooling estimation can enhance the forecasting performance over the 12-month historical variance estimator benchmark by 13%. Adding one more term structure segment ($\mathcal{S}_{3,12}$) can further enhance the out-of-sample R^2 to 13.7%. Adding additional segments no longer enhances the out-of-sample forecasting performance as the increased estimation errors cancel out the information contribution from the additional historical variance segments.

The second estimation approach retains the unconditional cross-sectional regression but performs conditionally weighted averaging on the time series of the unconditional coefficients. The approach ignores the information cycle difference across different companies at any given date but accommodates aggregate dynamics variation across the market information cycle. Across the different model specifications, the conditional time-series weighting based on the aggregate market conditions can enhance the out-of-sample forecasting R^2 by half to one percentage point.

The third estimation approach applies conditional weighting to the cross-sectional regression but maintains equal weighting to the time series of the cross-sectional conditional coefficient estimates. The approach captures the information cycle variation across different companies at a given date but ignores the aggregate market information cycle variation. Compared to the contribution

of the time-series conditional weighting in the second approach, applying conditional weighting to the cross-sectional regression generates much larger performance improvement. The forecasting R^2 increases to 17% for model I and to 18.4% for models II to IV. The much larger improvement suggests that the cross-sectional variation of the variance dynamics is much larger than the aggregate dynamics variation over time.

The last estimation approach applies conditional weighting to both dimensions and generates the best out-of-sample forecasting performance. The out-of-sample forecasting R^2 is the highest for model II at 19.2%. The DM t-statistics in the brackets are all computed against this best performing model-estimation combination. Under the same two-dimensional conditional weighting estimation approach, model II significantly outperforms model I, highlighting the statistical significance of adding two term structure segments in the forecasting relations. On the other hand, although adding additional components generates better in-sample performance for models III and IV, the added segments no longer improve the out-of-sample performance.

The out-of-sample performance comparison in Table 7 highlights the importance of allowing both cross-sectional and time-series variance dynamics variation across different phases of the information cycle. Our proposed conditional variance dynamics in (6) allows us to capture dynamics variations along both dimensions. Nevertheless, the importance of allowing cross-sectional and time-series dynamics variation does not, in any way, diminish the important role played by data pooling along the two dimensions to reduce estimation errors. A name-by-name time-series estimation would allow maximum and free variation across names, whereas using the cross-sectional regression on a single date to generate forecasts would allow maximum variation over time. Neither approach, however, is likely to generate stable out-of-sample forecasting performance because of the large estimation errors produced from these approaches. Applying conditional pooling along both dimensions with conservative bandwidth choices enjoys the benefits of both worlds. The con-

ditional pooling estimation reduces the estimation error via pooling over both the cross section and the time series while also allows the variance dynamics to vary across the information cycle.

3.4. Robustness analysis

In our conditional pooling estimation approach, the source of dynamics differentiation comes from the choice of the conditioning variables, whereas the balance between allowing dynamics variation and reducing estimation errors is achieved through the bandwidth choice on the bivariate Gaussian kernels in two weighting functions (8) and (11). As a robustness analysis, this section first gauges the effect of bandwidth choice on the out-of-sample forecasting performance, and then examines the effectiveness of other commonly constructed firm characteristics as alternative conditioning variables to differentiate the variance dynamics.

3.4.1. Balance differentiation with identification via bandwidth choice

The two-dimensional conditional pooling approach balances the need for pooling the data along the two dimensions to reduce estimation errors and the need for capturing the dynamics variation across different phases of the information cycle. The balance is achieved via the bandwidth choice. A smaller bandwidth allows more variation, but at the expense of reducing the degree of pooling and potentially increasing the estimation errors of the parameter estimates. A larger bandwidth choice generates more stability in the parameter estimates and smaller out-of-sample performance deterioration relative to in-sample performance, but at the expense of smoothing out true dynamics variations and generating inferior in-sample and out-of-sample performance.

Our main analysis starts with the default bandwidth choice and applies a multiple of two to increase the stability in out-of-sample forecasting applications. In this section, we repeat the out-

of-sample forecasting exercise by applying different bandwidth multiples. Table 8 reports the out-of-sample forecasting R^2 estimates from this robustness analysis. Panel A fixes the time-series conditional weighting with the bandwidth multiple of two while changing the bandwidth multiple on the conditionally weighted cross-sectional regressions. The out-of-sample performance is reasonably stable when the bandwidth multiple increases from one to five, and all are much better than the unconditional cross-sectional regression, which we denote in the last row of the panel as effectively having a cross-sectional smoothing bandwidth of infinity.

[Table 8 about here.]

On the other hand, when the bandwidth multiple decreases to half of the default choice, the out-of-sample forecasting performance deteriorates significantly. When we further reduce the bandwidth multiple, the conditionally weighted regressions start to experience numerical singularity issues, leading to extreme coefficient estimates and degenerating out-of-sample performance.

Panel B of Table 8 fixes the cross-sectional conditional weighting with a robust bandwidth multiple of two and varies the bandwidth choice on the time-series conditional weighting of the cross-sectional coefficients. The conditionally weighted cross-sectional regression with a bandwidth multiple of two generates stable coefficient estimates overall, making the out-of-sample performance less sensitive to the conditional weighting along the time series dimension. When the bandwidth multiple increases from one to five, the out-of-sample performance remains stable and in all cases better than the case in the last row without conditional weighting over the time series.

When the time-series conditional weighting bandwidth reduces to a fraction of the default choice, the out-of-sample forecasting performance also starts to deteriorate, similar to the observations for the cross-sectional weighting. The out-of-sample R^2 estimates become close to zero when we set the multiple to 1% of the default choice. The performance deterioration suggests that

even with the large cross section, a single cross-sectional regression is not sufficient to generate stable parameter estimates. Time-series pooling remains important in reducing estimation errors and generating robust out-of-sample performance.

3.4.2. Other firm characteristics as conditioning variables

A major contribution of this paper is the identification of the historical variance term structure as the key conditioning variable that differentiates the variance dynamics across different phases of the information cycle. Other types of firm characteristics may also differentiate the variance dynamics across companies and can serve as conditioning variables. In particular, the literature has identified a list of firm characteristics that differentiate the stock return behaviors. The Carhart (1997) four-factor return model represents such a classic example. The model explains the cross-sectional stock return variation with four firm characteristics: stock return beta, company size, book-to-market (BM) equity ratio, and momentum. Starting with Fama and French (1992) and Jegadeesh and Titman (2001), the literature has repeatedly examined the role of these firm characteristics in explaining the cross-sectional stock return variation. In addition, industry classification has often been used in return factor models to capture the higher return co-movements between firms within the same industry. This section examines the effectiveness of these firm characteristics in differentiating the variance dynamics as conditioning variables.

We construct these characteristics with the following procedure: We estimate the stock return beta by regressing the past one year history of the daily stock returns against the return series on the value-weighed market portfolio obtained from Professor French's on-line data library. We represent the size of a company by the logarithm of its market capitalization. We construct the BM ratio by obtaining the company's the total book common equity from the most recently available quar-

terly financial statement from Compustat and dividing it by the company's market capitalization. We represent momentum with the cumulative stock return over the past 12 month excluding the most recent month. In matching the financial statement with the market pricing information, we assume that the statement becomes available 45 days after the quarter. For industry classification, we rely on the two-digit GICS classification, which classifies the companies into 10 sectors. A real estate sector was recently separated out from the financial sector, but this newly created sector has very few observations during the early sample period. We recombine this sector with the financial sector in our classification.

To construct the above firm characteristics, we need the availability of the quarterly financial statement information and the GICS industry classification for each company. The financial statement data becomes widely available starting in 1970. We construct the joint sample from February 1970 to November 2022 for a total of 634 months. The number of names at each month ranges from 844 in February 1970 to 4,871 in October 1997, for an average of 2,704 names per month.

We take model II as an example and repeat the conditionally weighted cross-sectional regression on the model based on the different conditioning variables. For the four firm characteristics, we convert each variable into decile ranks and generate conditional weighting with a univariate Gaussian kernel on the decile rank difference and a bandwidth multiple of two. For each target sector, we generate conditional weighting based on whether the companies are in the sector or not. Companies in the target sector are assigned a distance of zero. Companies outside the sector are assigned a distance of one. We apply the same Gaussian kernel to this industry distance measure with the same bandwidth multiple of two. This leads to a weight of one for companies within the target sector and a very small weight close to zero for companies outside the target sector.

Table 9 reports the time-series averages of the conditional loading coefficient estimates across

each conditioning decile, with $\beta_1(x)$ in panel A and $\beta_3(x)$ in panel B. Each row is for one conditioning variable (x). Each column is for one conditioning decile. The last two columns report the average 10-1 decile difference and the absolute Newey and West (1987) t-value on the average difference (in parentheses). For comparison, we also perform univariate conditional estimation using the two segments of the historical variance term structures.

[Table 9 about here.]

The short end of the variance term structure, $\mathcal{S}_{1,3}$, generates the largest differentiation on the $\beta_1(x)$ coefficients. The average conditional estimates decline from 0.49 at the first decile to -0.42 at the 10th decile, almost completely flipping the predictive direction. The 10-1 decile difference is extremely large at -0.91, with a very high absolute t-value on the difference at 13.6. The cross-decile variations of the $\beta_1(x)$ estimates are much smaller on other conditioning variables. Among the four return factors, momentum generates the most significant differentiation, with an average 10-1 difference of 0.16 and a high t-value of 14.2. The high t-value suggests that the cross-decile difference on momentum is very stable over time, but the absolute magnitude of the average difference is much smaller than the average difference with $\mathcal{S}_{1,3}$ as the conditioning variable.

The last row reports the average $\beta_1(x)$ estimates for each GICS sector, sorted based on the GICS number from energy (10) to utility (55). There is no natural ranking across the 10 GICS sectors. The last two columns report the average difference between the maximum and minimum average estimates. The average conditional estimates from the different sectors are very close to one another. The maximum to minimum difference is merely 0.08. Therefore, although stock returns within the same sector tend to have higher correlations than stock returns from different sectors, sector classification does not differentiate the variance dynamics.

Panel B reports the conditional estimates on $\beta_3(x)$. The average conditional estimates increase

strongly with the short end of the variance term structure $\mathcal{S}_{1,3}$, but decline strongly with the long end of the term structure $\mathcal{S}_{3,12}$. Both the 10-1 average absolute differences and the absolute t-values are large. By comparison, the other conditioning variables generate much smaller conditional variations. Neither the four return factor characteristics nor the industry classification are as effective as the variance term structures in differentiating the variance dynamics.

Table 10 reports in panel A the time-series averages of the in-sample cross-sectional R^2 estimates from the different conditional regressions. For comparison, we also report the in-sample R^2 estimates from the unconditional regression, as well as the univariate and bivariate conditional regressions based on the two segments of the variance term structure, all on the same data sample. The unconditional regression generates an average R^2 of 19.74%. The univariate conditional regressions using the two segments of the variance term structure generate average R^2 estimates of 29.03% and 27.83%, respectively, both about 10 percentage points higher than the unconditional benchmark. Combining the two segments in a bivariate conditional regression further increases the average R^2 to 32.64%.

[Table 10 about here.]

The univariate conditional regressions based on the four return factors and industry classification all generate higher average in-sample R^2 estimates than the unconditional regression benchmark, but the in-sample improvements are all small, and are all much lower than those from the conditional regressions with the variance term structure. In sample, neither the four return factors nor the industry classification can differentiate the variance dynamics as effectively as the historical variance term structure.

For each cross-sectional regression, we also perform an out-of-sample forecasting exercise where we use the 10-year rolling average of the cross-sectional regression estimates to predict

the variance for the next month. Table 10 reports the out-of-sample R^2 estimates in panel B. The unconditional regression benchmark generates an out-of-sample R^2 of 16.42%. Conditioning on the four return factors or industry classification generates out-of-sample R^2 estimates from 16.41 to 15.78%, not much different from the unconditional regression. By contrast, using either segment of the historical variance term structure can significantly raise the out-of-sample R^2 estimates to 21.51% and 19.32%, respectively. Combining the two segments of the variance term structure into a bivariate conditioning regression further increases the out-of-sample R^2 to 22.53%.

This conditional analysis highlights the uniquely important role of the information cycle, as captured by the historical variance term structure, in differentiating the variance dynamics.

4. Conditional asset pricing across the information cycle

This paper identifies the historical variance term structure as a key conditioning variable that can differentiate the variance dynamics across different phases of the information cycle of a company. An important research question is whether and how classic capital asset pricing relations vary across the information cycle. This section offers a first look at the conditional asset pricing implication.

We start with the classic Carhart (1997) four-factor model as an example and examine how the classic factor return relations vary across the information cycle. The model explains the cross-sectional stock return variation with four firm characteristics: beta, company size, book-to-market (BM) equity ratio, and momentum. We have constructed these firm characteristics in the previous section. We use the same data sample for the analysis from February 1970 to November 2022.

At each month t, we start by estimating the unconditional cross-sectional forecasting factor

return relation,

$$\mathbf{r}_{t+1} = \eta_{t+1} + F_t \varphi_{t+1} + \mathbf{e}_{t+1}, \tag{12}$$

where $\mathbf{r}_{t+1}$ denotes the $(N_t \times 1)$ vector of stock returns over the next month and F_t denotes the $(N_t \times 4)$ matrix of the firm characteristics. Following industry standard practice, we perform cross-sectional winsorization on each characteristic at the (1-99%) range and cross-sectionally standardize the winsorized sample by deducting the cross-sectional mean and then dividing the demeaned characteristics by its cross-sectional standard deviation. The slope coefficient φ_{t+1} from the cross-sectional regression represents the next-month return on the corresponding factor portfolio that has a unit exposure to the particular firm characteristic while being neutral to all other characteristics. The time-series averages of the slope coefficients can be regarded as the average risk premium estimates on the factor exposures.

To examine how the asset pricing relation varies across the information cycle, we expand on the Carhart (1997) model and propose a conditional return factor model that allows the factor return relations to vary conditionally with the information cycle $x_t = \mathcal{S}_{t,1,12,i}$,

$$\mathbf{r}_{t+1} = \eta_{t+1}(x_t) + F_t\, \varphi_{t+1}(x_t) + \mathbf{e}_{t+1}. \tag{13}$$

To ease exposition, instead of dividing the term structure into two segments, we use the 1-12 month variance term structure $\mathcal{S}_{t,1,12,i}$ as the single conditioning variable. We perform conditionally weighted cross-sectional relation estimation using the historical variance term structure as the conditioning variable. The procedure is analogous to the conditional estimation of the cross-sectional variance forecasting relation. We convert the conditioning variable $\mathcal{S}_{t,1,12,i}$ into decile ranks and repeat the conditional estimation for each decile group. We use a univariate Gaussian kernel and a bandwidth multiple of two to construct the conditional weight.

Table 11 summarizes the estimation results. Panel A reports the average unconditional factor portfolio returns from the unconditional cross-sectional regression estimation of the classic factor structure in (12). The absolute magnitudes of the Newey and West (1987) t-statistics on the average returns are reported in parentheses. The average return on the beta portfolio is significantly negative, whereas the average returns on the size, BM, and momentum portfolios are significantly positive. These unconditional estimates are qualitatively similar to literature findings.

[Table 11 about here.]

Panel B reports the average conditional return behaviors of the factor portfolios from (13) at each decile of the historical variance term structure, with the lower deciles denoting downward sloping variance term structures and the higher deciles denoting upward sloping term structures. The average returns on the beta portfolio become increasingly negative as the variance term structure becomes increasingly upward sloping. The average return at the 10th decile more than doubles the average return at the first decile in absolute magnitude. Classic asset pricing theory predicts the expected excess return on a security to increase with its beta, but the literature has found evidence similar to ours against this theoretical prediction. Researchers have attributed this finding to capital constraints (Frazzini and Pedersen (2014)), investor penchant for lottery payoffs (Barberis and Huang (2008)), and optionality (Wu and Xu (2023)). Our conditional analysis shows that investor demand for high beta names are particularly high when the companies are in the upward moving phase of the information cycle and are facing greater uncertainty about its future path. The last row of the table reports the average returns and t-values of the 10-1 decile spread portfolio. The average decile spread portfolio return on the beta portfolio is significantly negative.

The size portfolio generates positive returns on average, suggesting that larger firms earn higher average returns under our model and sample setting. This size effect also varies strongly with the

historical variance term structure, weaker when the term structure is at the downward sloping lower deciles and stronger when the term structure is at the upward sloping higher deciles. The average return on the 10-1 decile spread portfolio is positive and highly statistically significant. If larger firms make higher average returns, it is more so when companies are at the highly uncertain stages of their information cycle.

The book-to-market portfolio generates positive returns on average. The average returns on the conditional portfolios are stable, and do not vary much with the information cycle. If one regards the BM ratio as a relative value measure (Lakonishok, Shleifer, and Vishny (1994) and La Porta, Lakonishok, Shleifer, and Vishny (1997)), the value effect works well across all phases of the information cycle. If on the other hand one regards the BM ratio as a risk factor and attributes the return prediction as a risk premium for bearing the risk (Fama and French (1995, 1997, 2006), Penman, Richardson, and Tuna (2007), Griffin and Lemmon (2002), and Vassalou and Xing (2004)), the risk premium shows particular stability across the information cycle.

Out of the four factor portfolios, the unconditional momentum portfolio generates the highest average returns with the highest t value. Conditionally, the average portfolio return increases strongly with increasing variance term structure. The average return on the 10th decile portfolio more than doubles the average return on the first decile portfolio. The 10-1 decile spread portfolio generates a highly positive average return, with the strongest statistical significance. In aggregate, Daniel and Moskowitz (2016) find that the momentum effect is stronger when the market volatility is low, but can crash during crisis times with volatile market conditions. Our findings in the cross section are in a somewhat opposite direction: On average, the momentum effect can actually be much stronger for companies with upward moving volatility trends.

Conditional asset pricing is a big topic. What our exercise highlights is the important role

the historical variance term structure can play in identifying the information cycle of different companies. Companies at different phases of their information cycle not only experience different variance dynamics, but can also experience different asset pricing relations.

5. Concluding remarks

Stock return variance reflects the intensity of the information flow about the underlying company. The well-documented volatility clustering behavior shows that the information intensity variation exhibits persistent up and down cycles. In this paper, we use historical variance estimators to proxy the information intensity for each company, and use historical variance term structures to identify the different phases of the information cycle. Across the US stock universe over the past century, we find that the information cycles of different companies are anything but synchronous. At any given date, the historical variance term structure can be highly upward sloping for some companies, but downward sloping for others. More importantly, we find that the up and down phases of the information cycle tend to have different durations, with the up phase fast and dramatic but the down phase slow and protracted, a phenomenon we characterize as "elevators up, stairs down."

We examine the implications of this phenomenon in two areas. The first is in the prediction of stock return variance over a large stock universe. As a result of the duration variation, the autoregressive volatility clustering behavior of the variance process varies strongly across different phases of the information cycle. We propose a conditional variance dynamics that predicts the future realized variance of a stock with its historical variance estimators, with the loading coefficients varying conditionally as a function of the company's information cycle as captured by its historical variance term structures. We propose a two-dimensional conditional pooling estimation approach that balances between the need for reducing estimation errors and the need for accommo-

dating the dynamics variation across the information cycle. An out-of-sample forecasting exercise highlights the effectiveness of the conditional pooling estimation approach and the importance for accommodating dynamics variations across the information cycle.

Second, we examine the asset pricing implication of the information cycle. In particular, we examine how classic asset pricing relations vary across the information cycle. We expand on classic return factor models and propose a new conditionally varying return factor model where the factor portfolio returns are allowed to vary conditionally as a function of the information cycle. We estimate the conditional factor return relation with a conditionally weighted cross-sectional regression. The conditional estimation shows that the average returns on the beta portfolio are significantly negative and increasingly so as a company's historical variance term structures become more upward sloping. By contrast, the average returns on the size and momentum portfolios are not only significantly positive, but also increase strongly as a company's historical variance term structure becomes more upward sloping. One exception is the book-to-market equity portfolio, which generates significantly positive average returns that are largely stable across different phases of the information cycle.

References

Andersen, T. G. 1996. Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility. *Journal of Finance* 51(1):169–204.

Andrews, D. W. K. 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* 59(3):817–858.

Ané, T, and H. Geman. 2000. Order Flow, Transaction Clock and Normality of Asset Returns. *Journal of Finance* 55(5):2259–2284.

Bai, J, and L. Wu. 2016. Anchoring Corporate Credit Default Swap Spreads to Firm Fundamentals. *Journal of Financial and Quantitative Analysis* 51(5):1521–1543.

Barberis, N, and M. Huang. 2008. Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *American Economic Review* 98(5):2066–2100.

Berger, D., A. Chaboud, and E. Hjalmarsson. 2009. What Drives Volatility Persistence in the Foreign Exchange Market?. *Journal of Financial Economics* 94(2):192–213.

Bollerslev, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31(3):307–327.

Bollerslev, T., B. Hood., J. Huss, and L. H. Pedersen. 2018. Risk Everywhere: Modeling and Managing Volatility. *Review of Financial Studies* 31(7):2729–2773.

Carhart, M. M. 1997. On Persistence in Mutual Fund Performance. *Journal of Finance* 52(1):57–82.

Carr, P, and L. Wu. 2004. Time-Changed Lévy Processes and Option Pricing. *Journal of Financial Economics* 71(1):113–141.

Corsi, F. 2009. A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics* 7(2):174–196.

Daniel, K, and T. J. Moskowitz. 2016. Momentum Crashes. *Journal of Financial Economics* 122:221–247.

Diebold, F. X, and R. S. Mariano. 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13(3):253–263.

Ding, Z, and C. W. J. Granger. 1996. Modeling Volatility Persistence of Speculative Returns: A new Approach. *Journal of Econometrics* 73(1):185–215.

Engle, R. F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica* 50(4):987–1008.

Fama, E. F, and K. French. 1992. Cross-section of Expected Stock Returns. *Journal of Finance* 47(2):427–465.

Fama, E. F, and K. R. French. 1995. Size and Book-to-Market Factors in Earnings and Returns. *Journal of Finance* 50(1):131–155.

Fama, E. F, and K. R. French. 2006. The Value Premium and the CAPM. *Journal of Finance* 61(5):2163–2185.

Frazzini, A, and L. H. Pedersen. 2014. Betting Against Beta. *Journal of Financial Economics* 111(1):1–25.

Ghysels, E., P. Santa-Clara, and R. Valkanov. 2006. Predicting Volatility: Getting the Most Out of Return Data Sampled at Different Frequencies. *Journal of Econometrics* 131(1-2):59–95.

Glosten, L. R., R. Jagannathan, and D. E. Runkle. 1993. On the Relation Between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks. *Journal of Finance* 48(5):1779–1801.

Goyal, A, and A. Saretto. 2009. Cross-Ssection of Option Returns and Volatility. *Journal of Financial Economics* 94(2):310–326.

Griffin, J. M, and M. L. Lemmon. 2002. Book-to-Maket Equity Ratio, Distress Risk, and Stock Returns. *Journal of Finance* 57(5):2317–2336.

He, X, and R. Velu. 2014. Volume and Volatility in a Common-Factor Mixture of Distributions Model. *Journal of Financial and Quantitative Analysis* 49(1):33–49.

Heston, S. L. 1993. Closed-Form Solution for Options with Stochastic Volatility, with Application to Bond and Currency Options. *Review of Financial Studies* 6(2):327–343.

Jegadeesh, N, and S. Titman. 2001. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *Journal of Finance* 56(2):699–720.

Lakonishok, J., A. Shleifer, and R. W. Vishny. 1994. Contrarian Investment, Extrapolation, and Risk. *Journal of Finance* 49(5):1541–1578.

Lamoureux, C. G, and W. D. Lastrapes. 1990. Heteroskedasticity in Stock Return Data: Volume versus GARCH Effects. *Journal of Finance* 45(1):221–229.

Mandelbrot, B. B. 1963. New Methods in Statistical Economics. *Journal of Political Economy* 71(5):421–440.

Nelson, D. B. 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59(2):347–370.

Newey, W. K, and K. D. West. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3):703–708.

Penman, S. H., S. A. Richardson, and I. Tuna. 2007. The Book-to-Price Effect in Stock Returns: Accounting for Leverage. *Journal of Accounting Research* 45(2):427–467.

La Porta, R., J. Lakonishok., A. Shleifer, and R. Vishny. 1997. Good News for Value Stocks: Further Evidence on Market Efficiency. *Journal of Finance* 52(2):859–874.

Simonoff, J. S. 1996. *Smoothing Methods in Statistics*. Springer-Verlag New York.

Tauchen, G. E, and M. Pitts. 1983. The Price Variability-Volume Relationship on Speculative Markets. *Econometrica* 51(2):485–505.

Vassalou, M, and Y. Xing. 2004. Default Risk in Equity Returns. *Journal of Finance* 59(2):831–868.

Wu, L, and Y. Xu. 2023. Separate Risk from Optionality. Working paper, Baruch College.

Table 1

Summary statistics of variance estimators

Entries report the time-series averages of the cross-sectional summary statistics of the historical variance estimators over different horizons, including the percentile values in volatility percentage points at 1, 50, and 99th percentiles (panel A), and the skewness and excess kurtosis estimates on the cross-sectional distribution of the variance estimators (panel B) and the skewness and excess kurtosis on the cross-sectional distribution of the natural logarithm of the variance estimators (panel C).

	A. Volatility percentiles			B. Variance distribution		C. Log variance distribution	
	1	50	99	skewness	kurtosis	skewness	kurtosis
HV_1	10.1	35.5	108.4	9.7	260.2	-0.3	1.2
HV_3	12.6	37.3	99.4	7.8	207.3	-0.2	0.9
HV_6	13.8	38.0	95.4	7.2	191.9	-0.2	0.7
HV_9	14.4	38.3	94.4	7.1	190.5	-0.1	0.6
HV_{12}	14.8	38.6	94.3	7.1	186.8	-0.1	0.5

Table 2

Cross-sectional forecasting correlation between historical and future variance estimators

Entries report the time-series summary statistics of the cross-sectional forecasting correlation between the historical variance estimators at different horizons (h) and the future realized variance over next month. The three panels estimate the forecasting correlations in the variance space (panel A), the natural logarithm of variance (panel B), and the natural logarithm of the variance ratio to the 12-month historical variance estimator (panel C). The statistics include the sample average (Mean), standard deviation (SD), minimum (Min), and maximum (Max).

	A. Variance				B. Log variance				C. Log variance ratio			
h	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
1	0.37	0.16	-0.00	0.98	0.65	0.07	0.42	0.92	0.27	0.09	-0.14	0.64
3	0.43	0.17	0.00	0.88	0.69	0.07	0.42	0.92	0.24	0.09	-0.27	0.57
6	0.45	0.18	0.00	0.96	0.70	0.07	0.43	0.91	0.17	0.09	-0.34	0.52
9	0.45	0.18	0.00	0.95	0.70	0.07	0.42	0.90	0.11	0.08	-0.30	0.46
12	0.45	0.18	0.00	0.89	0.70	0.07	0.42	0.89				

Table 3

Cross-sectional and time-series variation of historical variance term structures

Entries in panel A report the time-series averages of the cross-sectional statistics of the two historical variance term structure segments ($\mathcal{S}_{t,1,3,i}, \mathcal{S}_{t,3,12,i}$). Panel B takes the cross-sectional average at each date to construct the market variance term structure ($\mathcal{S}_{t,1,3,m}, \mathcal{S}_{t,3,12,m}$) and reports their time-series summary statistics. The statistics include the percentile values at 1, 15, 50, 75, 90, and 99th percentiles. The last row reports the proportion of positive term structures (PTS).

	A. Average cross-sectional variation		B. Market time-series variation	
	$\mathcal{S}_{t,1,3,i}$	$\mathcal{S}_{t,3,12,i}$	$\mathcal{S}_{t,1,3,m}$	$\mathcal{S}_{t,3,12,m}$
1	-1.707	-1.309	-0.694	-0.849
10	-0.758	-0.592	-0.373	-0.398
25	-0.403	-0.333	-0.236	-0.238
50	-0.077	-0.084	-0.139	-0.111
75	0.207	0.152	-0.033	0.028
90	0.441	0.370	0.105	0.172
99	0.807	0.800	0.578	0.715
PTS	0.443	0.427	0.201	0.282

Table 4

Forecasting correlation across different phases of the information cycle

We use two segments of the historical variance term structure $\mathcal{S}_{t,3,12,i}$ and $\mathcal{S}_{t,1,3,i}$ as conditioning variables to capture the different phases of the information cycle. For each conditioning variable at each date, we classify the stock universe into three quantiles and label the first and last quantile as the "down" and "up" phase of the information cycle, respectively. We measure the cross-sectional forecasting correlation between the historical variance term structure $\mathcal{S}_{t,h,12}$ at different horizons h and the one-month ahead realized variance term structure $\mathcal{S}_{t+1,12}$ conditional on each quantile of the two conditioning variables. Entries report the time-series averages of the conditional cross-sectional correlation estimates. The first row reports the unconditional forecasting correlation estimates as a reference.

Conditioning variables		Forecasting correlations between $\mathcal{S}_{t+1,12}$ and $\mathcal{S}_{t,h,12}$			
$\mathcal{S}_{3,12}$	$\mathcal{S}_{1,3}$	$h=1$	3	6	9
unconditional		0.268	0.235	0.171	0.105
down	down	0.338	0.321	0.161	0.083
down	up	0.262	0.296	0.161	0.085
up	up	-0.054	-0.011	0.010	0.005
up	down	0.305	-0.084	-0.040	-0.027

Table 5

The average behavior of cross-sectional variance forecasting dynamics

At each date, we perform unconditional ordinary least square regressions on the cross-sectional variance forecasting specifications with increasingly more historical log variance ratios as predictors from model I to model IV. Entries report the sample averages of the coefficients β_h and the percentage R^2 estimates for each model. In parentheses under the coefficient estimates are the Newey-West t-statistics for the sample average. In brackets under each R^2 estimate is the DM t-statistic of the R^2 difference between the best performing model IV and the model in question.

Model	Loading coefficients, β_h				R^2
	$h=1$	3	6	9	(%)
I.	0.390				17.0
	(31.5)				[22.6]
II.	0.277	0.225			19.7
	(31.4)	(16.5)			[17.0]
III.	0.270	0.156	0.119		21.1
	(33.9)	(13.0)	(6.16)		[13.2]
IV.	0.264	0.156	0.075	0.010	22.1
	(34.6)	(13.9)	(4.1)	(0.4)	—

Table 6

Conditional variance dynamics variation with historical variance term structures

Entries report the time-series averages of the conditionally weighted estimates of the conditional loading coefficients $\beta_1(\mathbf{x})$ and $\beta_3(\mathbf{x})$ for model II at each of the 10×10 grids of the bivariate conditional variables of $(\mathcal{S}_{1,3} \times \mathcal{S}_{3,12})$. The first column reports the average values of the conditioning variable$\mathcal{S}_{1,3}$ at each decile. The first row reports the average values of the conditioning variable $\mathcal{S}_{3,12}$ at each decile level.

$\mathcal{S}_{3,12}$	-0.91	-0.49	-0.33	-0.22	-0.13	-0.04	0.05	0.15	0.28	0.56
$\mathcal{S}_{1,3}$	A. Average conditional estimates for $\beta_1(\mathbf{x})$									
-1.18	0.27	0.27	0.30	0.32	0.33	0.34	0.37	0.39	0.44	0.48
-0.61	0.24	0.25	0.28	0.30	0.31	0.32	0.35	0.38	0.43	0.47
-0.41	0.16	0.19	0.24	0.27	0.29	0.31	0.34	0.37	0.40	0.43
-0.26	0.11	0.16	0.22	0.26	0.29	0.32	0.35	0.38	0.38	0.37
-0.13	0.14	0.18	0.22	0.25	0.28	0.32	0.35	0.37	0.37	0.35
-0.02	0.20	0.20	0.20	0.20	0.22	0.23	0.24	0.25	0.27	0.28
0.09	0.22	0.18	0.14	0.12	0.10	0.08	0.05	0.03	0.04	0.05
0.21	0.16	0.10	0.04	0.00	-0.02	-0.05	-0.10	-0.12	-0.14	-0.19
0.35	0.09	0.03	-0.04	-0.08	-0.12	-0.14	-0.16	-0.18	-0.21	-0.29
0.61	0.05	-0.01	-0.08	-0.13	-0.17	-0.18	-0.20	-0.20	-0.19	-0.24
	B. Average conditional estimates for $\beta_3(\mathbf{x})$									
-1.18	0.39	0.36	0.24	0.11	0.08	0.06	0.05	-0.05	-0.11	-0.10
-0.61	0.42	0.38	0.28	0.19	0.17	0.13	0.07	-0.04	-0.11	-0.13
-0.41	0.50	0.44	0.34	0.27	0.24	0.18	0.09	-0.04	-0.12	-0.17
-0.26	0.56	0.48	0.37	0.30	0.27	0.19	0.08	-0.05	-0.13	-0.16
-0.13	0.52	0.45	0.38	0.34	0.31	0.23	0.10	-0.05	-0.13	-0.15
-0.02	0.47	0.44	0.42	0.41	0.41	0.34	0.21	0.06	-0.05	-0.10
0.09	0.46	0.48	0.49	0.51	0.52	0.48	0.39	0.27	0.17	0.12
0.21	0.51	0.56	0.59	0.61	0.61	0.56	0.50	0.41	0.36	0.39
0.35	0.58	0.61	0.65	0.65	0.64	0.60	0.55	0.45	0.37	0.45
0.61	0.61	0.64	0.66	0.65	0.63	0.61	0.58	0.40	0.25	0.28

Table 7
Out-of-sample variance forecasting performance
Entries report the out-of-sample forecasting R^2 estimates (in percentages) over the 12-month historical variance prediction benchmark from different model specifications (I to IV) and different estimation approaches in terms of pooling across cross-sectional (CS) and time-series (TS) dimensions. Along each dimension, "EW" denotes unconditional equal-weighted pooling and "CW" denotes conditionally weighted pooling. In brackets following each R^2 is the DM t-statistic on the outperformance of the best performer against that estimate.

Estimation approach			Model specification							
No.	CS	TS	I.		II.		III.		IV.	
1	EW	EW	13.0	[16.3]	13.7	[14.8]	13.7	[14.6]	13.7	[14.4]
2	EW	CW	13.5	[17.0]	14.7	[15.6]	14.6	[15.0]	14.5	[14.5]
3	CW	EW	17.0	[9.7]	18.4	[4.8]	18.4	[4.6]	18.4	[4.5]
4	CW	CW	18.0	[8.9]	19.2	[0.0]	19.2	[0.1]	19.1	[1.3]

Table 8

Effects of the degree of cross-sectional and time-series pooling

Entries report the out-of-sample forecasting R^2 estimates (in percentages) from different model specifications (models I to IV) with conditional weighting across both the cross section (CS) and the time series (TS) dimension with different bandwidth multiple (m) choices. A bandwidth of infinity represents the case of no conditional weighting along that dimension.

Bandwidth m choice		Model specification			
CS	TS	I.	II.	III.	IV.
A. Effects of cross-sectional pooling					
0.3	2	15.5	-159.7	-133.7	-149.8
0.5	2	17.5	15.7	15.1	12.0
1	2	18.1	19.0	19.0	18.8
2	2	18.0	19.2	19.2	19.1
3	2	17.8	19.1	19.1	19.0
5	2	17.2	18.9	18.9	18.8
∞	2	13.5	14.7	14.6	14.5
B. Effects of time-series pooling					
2	0.01	4.7	1.7	1.5	0.5
2	0.1	8.3	6.0	5.6	4.2
2	0.5	16.4	16.7	16.6	16.2
2	1	17.8	18.7	18.7	18.5
2	2	18.0	19.2	19.2	19.1
2	3	17.8	19.1	19.1	19.0
2	5	17.4	18.8	18.8	18.7
2	∞	17.0	18.4	18.4	18.4

Table 9

Conditional loading coefficient estimates with different conditioning variables

Entries report the conditional loading coefficient estimates ($\beta_1(x)$ in panel A and $\beta_3(x)$ in panel B) on model II across the deciles of different conditioning variables (x). Each row is for one conditioning variable. Each column is for one conditioning decile. The last two columns report the 10-1 average difference on the conditional coefficient estimates and the absolute Newey-West t value of the average difference (in parentheses). The last row in each panel reports the conditional estimates for each GICS sector, with 10-1 being the difference between the maximum and minimum estimates.

x	1	2	3	4	5	6	7	8	9	10	10-1	
						A. $\beta_1(x)$						
$\mathcal{S}_{1,3}$	0.49	0.47	0.44	0.44	0.45	0.33	0.07	-0.20	-0.37	-0.42	-0.91	(13.6)
$\mathcal{S}_{3,12}$	0.25	0.23	0.21	0.19	0.19	0.20	0.22	0.24	0.28	0.30	0.05	(5.3)
Beta	0.35	0.34	0.32	0.31	0.30	0.29	0.28	0.27	0.26	0.26	-0.09	(8.2)
Size	0.32	0.32	0.31	0.31	0.31	0.31	0.31	0.30	0.29	0.28	-0.04	(3.7)
BM	0.31	0.31	0.31	0.32	0.32	0.31	0.31	0.30	0.30	0.29	-0.02	(2.7)
Momentum	0.24	0.25	0.26	0.27	0.28	0.30	0.32	0.34	0.38	0.40	0.16	(14.2)
Industry	0.28	0.28	0.29	0.28	0.26	0.29	0.35	0.27	0.26	0.28	0.08	(5.9)
						B. $\beta_3(x)$						
$\mathcal{S}_{1,3}$	-0.02	0.02	0.08	0.10	0.11	0.26	0.52	0.75	0.84	0.81	0.82	(13.6)
$\mathcal{S}_{3,12}$	0.47	0.49	0.51	0.61	0.64	0.50	0.22	-0.11	-0.33	-0.43	-0.91	(15.2)
Beta	0.21	0.22	0.22	0.22	0.23	0.23	0.23	0.25	0.28	0.30	0.08	(3.9)
Size	0.23	0.22	0.22	0.21	0.22	0.23	0.24	0.25	0.27	0.28	0.05	(2.3)
BM	0.25	0.24	0.23	0.22	0.22	0.22	0.23	0.24	0.25	0.25	0.00	(0.1)
Momentum	0.22	0.22	0.22	0.23	0.23	0.22	0.23	0.23	0.24	0.26	0.04	(2.3)
Industry	0.28	0.26	0.23	0.24	0.25	0.24	0.23	0.24	0.31	0.31	0.08	(3.3)

Table 10

In-sample and out-of-sample performance with different conditioning variables

Entries in panel A report the time-series averages of the in-sample R^2 estimates in percentages from the cross-sectional regressions with different conditioning variables. Entries in panel B report the out-of-sample forecasting percentage R^2 estimates based on the 10-year rolling window averages of the conditional coefficient estimates. The first row reports the unconditional regression benchmark.

	A. In-sample R^2	B. Out-of-sample R^2
Unconditional	19.74	16.42
$\mathcal{S}_{1,3}$	29.03	21.51
$\mathcal{S}_{3,12}$	27.83	19.32
Bivariate	32.64	22.53
Beta	21.38	16.45
Size	21.30	16.41
BM	20.87	16.41
Momentum	21.26	16.78
Industry	22.01	16.45

Table 11

Asset pricing variation across the information cycle

Entries report the average annualized monthly returns and the absolute Newey-West t values (in parentheses) on the average returns on the stock factor portfolios extracted from the conditional stock return factor model. Panel A reports the statistics from the unconditional factor return model. Panel B reports the conditional regression estimates with the historical variance term structure as the conditioning variable. The last row reports the statistics on the decile spread portfolio that is long the 10th decile portfolio and short the first decile portfolio.

	Beta		Size		BM		Momentum	
	A. Unconditional asset pricing relation							
	-4.84	(3.98)	3.61	(4.40)	4.45	(6.13)	5.68	(8.25)
Decile	B. Conditional asset pricing relation across the $\mathcal{S}_{t,1,12}$ decile							
1	-3.11	(2.63)	2.75	(3.21)	4.76	(6.52)	3.98	(5.96)
2	-3.59	(3.06)	2.37	(2.80)	4.39	(6.06)	4.14	(6.14)
3	-4.03	(3.39)	2.30	(2.77)	4.36	(5.96)	4.51	(6.07)
4	-4.59	(3.76)	2.64	(3.20)	4.46	(6.05)	4.95	(6.28)
5	-4.88	(3.89)	3.19	(3.78)	4.45	(5.75)	5.52	(7.08)
6	-4.92	(3.87)	3.47	(3.97)	4.39	(5.68)	5.77	(7.51)
7	-5.32	(4.12)	3.62	(4.17)	4.42	(5.80)	5.66	(7.03)
8	-5.93	(4.47)	3.91	(4.52)	4.37	(5.69)	6.07	(7.35)
9	-6.31	(4.79)	4.68	(5.38)	4.45	(5.99)	7.31	(9.15)
10	-6.78	(5.35)	5.55	(6.47)	4.44	(5.93)	8.06	(9.78)
10-1	-3.67	(5.69)	2.80	(4.89)	-0.32	(0.58)	4.08	(5.87)

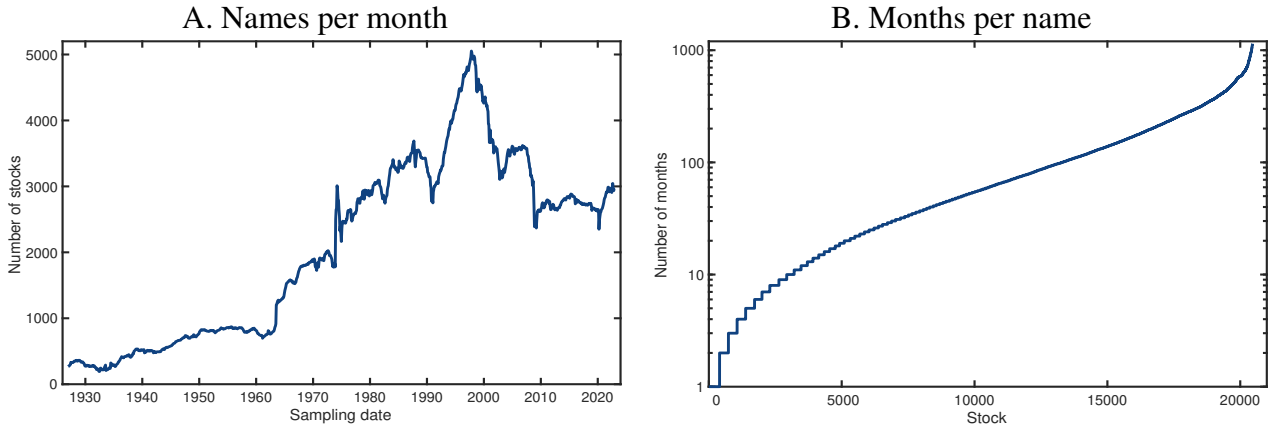


Figure 1

Number of selected stocks per month and numbers of selected months per stock.

Panel A plots the number of selected stocks at each month from December 1926 to November 2022. Panel B plots the number of selected sampling months for each of the 20,462 selected stocks, sorted based on the number of monthly observations and plotted on log scale.

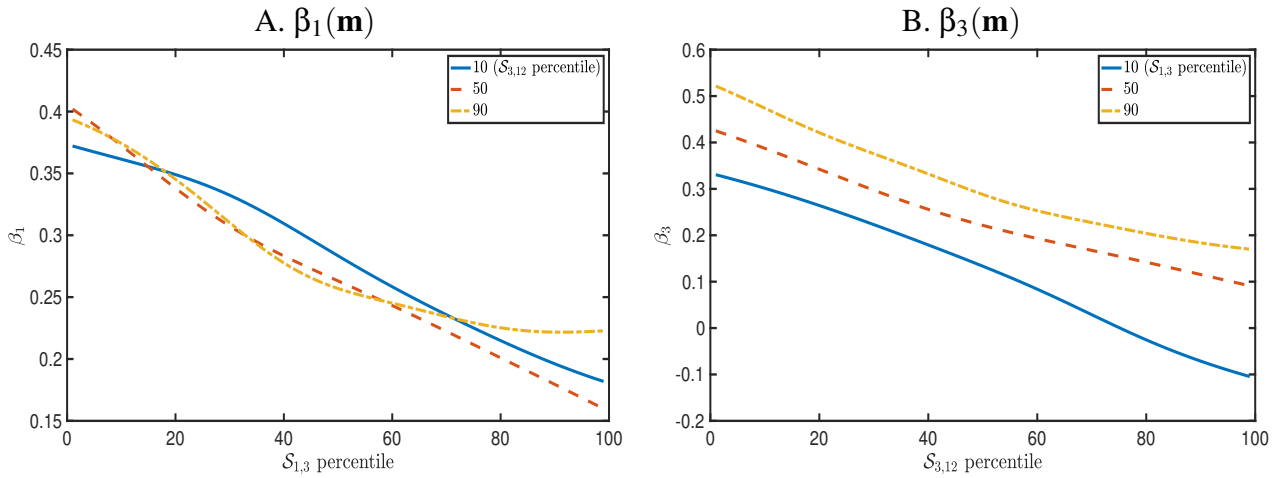


Figure 2

Conditionally weighted averages of loading estimates

The plots show the conditionally weighted time-series averages of loading coefficient estimates $\beta_1(\mathbf{m})$ (panel A) and $\beta_3(\mathbf{m})$ (panel B) from unconditional cross-sectional regressions on model II. Panel A highlights the dependence of the conditional averages of $\beta_1(\mathbf{m})$ on the short end of the market term structure $\mathcal{S}_{t,1,3,m}$ conditional on the long end of the market variance term structure at its 10% (dashed-dotted line), 50% (solid line), and 90% (dashed line) values, respectively. Panel B highlights the dependence of the conditional averages of $\beta_3(\mathbf{m})$ on the long end of the market term structure $\mathcal{S}_{t,3,12,m}$ conditional on the short end of the market variance term structure at its 10% (dashed-dotted line), 50% (solid line), and 90% (dashed line) values, respectively.