The Memory of Stock Return Volatility: Asset Pricing Implications

May 31, 2017

Abstract

This paper examines long memory volatility in the cross-section of stock returns. We show that long memory volatility is widespread in the U.S. and that the degree of memory can be related to firm characteristics such as market capitalization, book-to-market ratio, prior performance and price jumps. Long memory volatility is negatively priced in the cross-section. Buying stocks with shorter memory and selling stocks with longer memory in volatility generates significant excess returns of 1.71% per annum. Consistent with theory, we find that the volatility of stocks with longer memory is more predictable than stocks with shorter memory. This makes the latter more uncertain, which is compensated for with higher average returns.

JEL classification: C22, G12

Keywords: Asset Pricing; Long Memory; Persistence; Volatility

I Introduction

In this paper we investigate the memory of volatility in the cross-section of U.S. stocks. To the best of our knowledge, we are the first to analyze the asset pricing implications of long memory volatility. We show that long memory is prevalent in the volatility of individual stock returns. Long memory can be related to the size, past performance and jump intensity of a firm. Moreover, we provide time-series and cross-sectional evidence for a negative price of long memory volatility in the cross-section of stock returns.

We shed new light on the implication of long memory by combining three strands of literature. First, we extend the research on documenting long memory, which, so far, has only focused on indices or some large firms by investigating the complete cross-section of U.S. stocks. Second, we analyze the time-variation of long memory in volatility. Third, long memory has so far only been analyzed in the time-series dimension, not in the cross-sectional one. We discuss and investigate possible microeconomic fundamentals, which may explain long memory and examine whether memory is a priced factor.

We find that 95% of stocks possesses long memory in volatility with an average memory parameter of 0.22. At the firm level, higher volatility memory estimates are related to larger size, worse prior performance and fewer price jumps. Following the investment strategy of holding stocks with shorter memory in volatility and shorting stocks with longer memory in volatility generates excess returns of 1.71% per annum. This result is supported by cross-sectional regression tests. We find a significant risk premium for the memory parameter where stocks with anti-persistent volatility can earn up to 4.7% per annum more than stocks with long memory in volatility. We show that the volatility of stocks with higher memory parameters is more predictable than stocks with low memory parameters. This indicates that lower uncertainty of stocks with longer memory, i.e.

more persistent volatility, results in the negative premium.[1] Our results are robust to controling for idiosyncratic volatility, size, and other characteristics, as well as to various further tests. At the same time we verify our memory estimates by showing that forecasting volatility for stocks with longer memory works better than for stocks with shorter memory. We also relate our results to existing theoretical models, which show how long memory is generated through heterogeneity in the market.

Long memory processes (also referred to as long-range dependent processes) are present in numerous sciences and fields such as physics, geophysics, hydrology, climatology, biology and, most importantly for the subject of this project, economics and finance. Long memory processes can be described as long-range dependent time series with a hyperbolic decaying autocorrelation function, as opposed to the exponential function of short memory processes such as autoregressive processes. The introduction of long memory processes created a huge wave of new time-series models and methodologies to analyze, estimate, and predict them, since the old methods used for short memory time series were no longer appropriate. The first study to mention is perhaps Hurst (1951), who examines the Nile River in order to understand the persistence of stream flow data. There also exist several papers dealing with long memory in economics and finance. Baillie (1996) provides a detailed survey and review for this purpose. The most common models are the autoregressive fractionally integrated moving average (ARFIMA) model by Granger & Joyeux (1980), Granger (1981) and Hosking (1981) and the fractionally integrated generalized autoregressive conditional heteroskedasticity (FIGARCH) model introduced by Baillie et al. (1996). These are extensions of the short memory

[1] In recent studies, Baltussen et al. (2016) and Hollstein & Prokopczuk (2017) show that volatility-of-volatility is priced in the cross-section of stock returns. Although one might think that volatility-of-volatility is related to the degree of long memory in volatility, we empirically show that (i) it is not, and (ii) it is priced separately.

ARMA and GARCH models, respectively. Long memory properties have been analyzed comprehensively in returns and volatilities and our paper draws from several strands of literature.

The first focuses on the estimation and detection of long memory in the volatility of stock returns. Shortly after the introduction of the FIGARCH model, Bollerslev & Mikkelsen (1996) and Ding & Granger (1996) show that the conditional variance and absolute returns of the S&P 500 index, respectively, possess long memory. Breidt et al. (1998) also find long memory in the variance of equally weighted and value-weighted Center for Research in Security Prices (CRSP) stock market index returns. Lobato & Savin (1998) investigate the long memory properties of the U.S. stock market index and thirty individual stock returns in the U.S., while Sadique & Silvapulle (2001) and Henry (2002) consider the long memory property of various international stock indices, including Germany, Japan, Korea, New Zealand, Malaysia, Singapore, Taiwan and the U.S.

Another strand of the literature analyzes breaks in the long memory parameter, and hence allows memory to vary over time. Leybourne et al. (2007) consider long memory dynamics and introduce a test for a break from stationary long memory to non-stationary long memory. Their test is improved by Sibbertsen & Kruse (2009), since the results may be distorted when the data-generating process exhibits long memory. They apply the test to U.S. inflation data and find a break in the early 1980s. Sibbertsen et al. (2014) test for the persistence of EMU government bond yields for France, Italy and Spain, using the same methodology, and find breaks between 2006 and 2008.

Our paper is mostly related to the asset pricing literature. The research and discovery of anomalies and effects that can explain the cross-section of expected returns is constantly growing since the introduction of the capital asset pricing model (CAPM) (Sharpe, 1964;

Lintner, 1965; Mossin, 1966; Black, 1972). In addition to the market portfolio, Fama & French (1993) show that size and book-to-market ratio are better able to capture the cross-sectional variation in average stock returns. Carhart (1997) adds a momentum factor, and more recently, Fama & French (2015) extend their three-factor model by profitability and investment factors. The list of potential explanatory variables for the cross-sectional variation of stock returns is ongoing. For example, to name only two, Amihud (2002) finds a positive relationship between the illiquidity of stocks and future excess returns while Ang et al. (2006b) show that idiosyncratic volatility is negatively priced in the cross-section. Hou et al. (2014) propose the q-factor model including market, size, investment and profitability factors, and show that the performance of their model is at least as good as the models proposed by Fama & French (1993) and Carhart (1997).

The rest of the paper is organized as follows. Section II describes our data set and estimation procedure for long memory. Section III examines the cross-section of U.S. stocks. Section IV relates long memory to predictability. Section V theoretically discusses the origin of long memory. Section VI presents robustness tests and Section VII concludes.

II Data and Methodology

A Data

The data used for our analyses come from various sources. For our cross-sectional analysis of U.S. stock returns, we obtain equity prices, returns, market capitalization and volume data from the CRSP for the period from January 1926 until December 2015. In our main analysis we investigate four different firm characteristics which have been shown in the existing literature to be priced in the cross-section of stock returns. They include size, value, momentum effects and the liquidity factor. The construction of the variables, which we from now on refer to as Size, Book-to-Market, Momentum and Illiquidity, follows the convention of the literature (see Jegadeesh & Titman, 1993; Amihud, 2002; Fama & French, 2008; Jiang & Yao, 2013, among others) and are based on market capitalizations, returns and trading volumes from CRSP and balance-sheet information from COMPUSTAT.[2]

High-frequency price data are obtained from Thomson Reuters Tick History. When employing high-frequency data, the analysis is restricted to the period from January 1996 until December 2015 and on the S&P 500 constituents only.[3]

B Semi-parametric Estimation of Long Memory in Volatility

Our estimation of the long memory parameter relies on two of the most popular estimators, the GPH estimator and the Local Whittle estimator.

The first is based on the log-periodogram and was developed by Geweke & Porter-Hudak (1983). The GPH estimator employs a linear regression using the first m periodogram ordinates and exploits the shape of the spectral density around the origin. The spectral density of a stationary process X_t is estimated empirically by the periodogram:

$$I_{n,X}(\lambda) = \frac{1}{2\pi N}\left|\sum_{t=1}^{N} X_t e^{-it\lambda}\right|^2, \quad t = 1, ..., N \tag{1}$$

where the periodogram is not affected by centering of the time series for Fourier frequencies $\lambda_j = 2\pi j/N$ ($j = 1, ..., [(N-1)/2]$). The estimator is given by the negative slope estimate β_1 in the regression:

$$log(I(\lambda_j)) = \beta_0 + \beta_1 log[4sin^2(\lambda_j/2)] + \epsilon_j, \quad j = 1, ..., m \tag{2}$$

[2]Even though the size factor is constructed by calculating the logarithm of the market capitalization we refer to this factor as Size rather than log(Size).

[3]This choice is due to the restricted availability of high-frequency data for the complete cross-section, which is crucial for our long memory estimates.

Under mild conditions ($m \to \infty, n \to \infty, \frac{m}{n} \to 0$), Robinson (1995b) derives the asymptotic distribution:

$$\sqrt{m}(\hat{d} - d) \underset{d}{\longrightarrow} N\left(0, \frac{\pi^2}{24}\right) \tag{3}$$

which provides the asymptotic standard errors for the long memory parameter. The estimator is narrowband since the bandwidth parameter m leads to a bias–variance trade-off. While a high m far from the origin leads to bias, a low m too close to the origin leads to a rise in the variance.

The second estimator is the Local Whittle estimator, which is obtained by minimizing the following objective function:

$$\hat{d}_{LW} = \underset{d \in \theta}{\arg\min} \left[log\left(\frac{1}{m}\sum_{j=1}^{m}\frac{I(\lambda_j)}{\lambda_j^{2d}}\right) - \frac{2d}{m}\sum_{j=1}^{m} log\lambda_j \right], \quad \theta \subseteq (-0.5, 0.5) \tag{4}$$

where m is restricted to $m < \frac{n}{2}$. The Local Whittle estimator is an extension of the one originally proposed by Whittle (1951) which relies on an approximate maximum likelihood approach. Under mild assumptions similar to those for the GPH estimator, Robinson (1995a) derives the asymptotic distribution:

$$\sqrt{m}(\hat{d}_{LW} - d_0) \underset{d}{\longrightarrow} N\left(0, \frac{1}{4}\right) \tag{5}$$

For our main analysis we focus on the GPH estimator and the bandwidth $m = N^{0.5}$ following the existing literature (Geweke & Porter-Hudak, 1983; Diebold & Rudebusch, 1989; Hurvich & Deo, 1999; Henry, 2002).[4] Results with alternative bandwidth choices and the Local Whittle estimator are reported in the robustness section, Section VI.

[4]Typically, empirical researchers rely on this bandwidth choice since it is robust against short-range dependencies in the data. In terms of mean squared error (MSE) improvement, Beran et al. (2013) argue that the bandwidth $m = O(N^{0.8})$ is the optimal choice.

We refer to d as the memory parameter and differentiate between three cases: A time series has short memory if $d = 0$. A time series has negative memory or is anti-persistent if $d < 0$. A time series has long memory if $0 < d < 1$ where it is non-stationary if $0.5 < d < 1$.

III Long Memory Volatility in the Cross-Section of Stock Returns

In this section we provide evidence of long memory volatility in the cross-section of U.S. stock returns. First, we show in Section III.A that long memory volatility is prevalent in most stocks but that the degree varies across stocks. Section III.B relates the memory parameter to firm characteristics. Section III.C investigates whether long memory volatility is a priced factor.

A Descriptive Statistics

We apply the GPH estimator to the time series of squared returns for the cross-section of U.S. stocks. Since we are interested in the relationship between memory, firm characteristics and expected returns, we allow for a time-varying memory parameter. More specifically, we estimate the memory parameter at a monthly frequency using a rolling window, which includes the most recent five years of daily return observations.[5] Table 1 provides summary statistics for the memory parameter estimates. In our sample period we have on average 2480 memory parameter estimates at each point of time. The average estimate is 0.22 with a standard deviation of 0.12. The mean t-statistic of 23.34 suggests that the memory parameter is statistically significant on average. Also, we find that

[5]We require at least non-missing return observations on 50% of the days over the examined period for a stock to be included in our analysis.

most of the stocks exhibit long memory in volatility. 95% of the stocks show a memory parameter with $0.0 < d < 0.5$, while 3% of the stocks are anti-persistent and only 2% show non-stationary long memory.

Our results are consistent with the literature and extend the evidence of long memory in stock return volatility to a broader cross-section. Lobato & Savin (1998), for example, find that components of the Dow Jones Index show strong evidence of long memory in squared returns for the period from July 1962 until December 1994. Breidt et al. (1998) find for the equally weighted CRSP portfolio for the period from 1962 until 1989 a memory parameter of $d = 0.22$, which coincides with both the mean and the median from our analysis of the complete cross-section of the U.S. stocks.

B Explaining Long Memory with Firm Characteristics

In this section we relate the memory parameter of a stock's volatility to firm characteristics. We include Size, Book-to-Market, Momentum and Illiquidity. These variables have been shown to be priced in the cross-section of stock returns (Jegadeesh & Titman, 1993; Amihud, 2002; Fama & French, 2008; Jiang & Yao, 2013). We further include two jump measures since recent studies have shown that jumps are an important factor in the cross-section of stock returns. Jiang & Yao (2013) analyze the predictability of cross-sectional stock returns and find that once controling for jumps firm characteristics such as size and liquidity are no longer predictive. Kelly & Jiang (2014) and Cremers et al. (2015) show that the sensitivity of stocks to market tail and jump risk helps to explain the cross-sectional variation in expected returns. We apply the common jump test proposed by Barndorff-Nielsen & Shephard (2006) (BNS).[6] The test relies on the bipower variation,

[6] Pukthuanthong & Roll (2015) show with the help of simulations using different jump size and frequency, that this test is preferable to those proposed by Lee & Mykland (2008), Jiang & Oomen (2008) and Jacod & Todorov (2009).

which decomposes the quadratic variation into its parts due to continuous movements and a jump part. The jump test statistic is given by

$$BNS_t = \frac{(\pi/2)B_t - S_t}{\sqrt{((\pi^2/4) + \pi - 5)(\pi/2)^2 Q_t}} \tag{6}$$

$$Q_t = \frac{1}{K_t - 3} \sum_{k=4}^{K_t} |r_{t,k}||r_{t,k-1}||r_{t,k-2}||r_{t,k-3}| \tag{7}$$

$$S_t = \frac{1}{K_t} \sum_{k=1}^{K_t} r_{t,k}^2 \tag{8}$$

$$B_t = \frac{1}{K_t - 1} \sum_{k=2}^{K_t} |r_{t,k}||r_{t,k-1}| \tag{9}$$

where K_t is the number of observations over the examined period, $r_{t,k}$ is the kth daily observation over the examined period t and BNS_t is normally distributed under the null. First, we compute the BNS jump statistic for each month and stock using daily return data within each calender month following Pukthuanthong & Roll (2015). The first measure of jump intensity is given by the jump test statistic (BNS). Our second measure is a dummy variable indicating whether the current month includes a significant jump at the 5% level, which we denote as BNS-I.

Each month for the period from January 1950 until December 2015, we sort all stocks into quintile portfolios where stocks with the lowest memory parameter are in the first quintile and stocks with the highest memory parameter are in the fifth quintile. We then track the average firm characteristics of these quintile portfolios.[7]

Table 2 shows the results. We report the average memory and firm characteristics for each quintile and for the long memory minus short memory (LMS) portfolio. For the latter we also present t-statistics in square brackets in the last column. Average portfolio size, momentum, and jump measures demonstrate a monotonic pattern that is increas-

[7]We start our analysis in 1950 because book-to-market data is available only from 1950 in COMPUSTAT.

ing/decreasing in the memory parameter. Stocks with higher market capitalization, worse past performance and fewer jumps (higher jump statistics and fewer significant jumps) exhibit longer memory in volatility. These differences are highly statistically significant with absolute t-statistics above 12. There is no monotonic pattern for Book-to-Market and Illiquidity but the hedge portfolio shows positive values for both and the t-statistic is statistically significant for Illiquidity.

We complement the above analysis with cross-sectional regressions. At each point of time, we regress the memory parameter of each firm on the predictor variables in the following regression:

$$d_{i,t} = \alpha_t + \beta_t X_{i,t} + \epsilon_{i,t} \tag{10}$$

where $d_{i,t}$ is the memory estimate of stock i at time t, $X_{i,t}$ is the vector containing the firm characteristics of stock i at time t and ϵ_i is the error term.[8] The slope coefficients are expected to have signs as the LMS portfolio spreads. The coefficients are reported in Table 3 for three regressions. The first row excludes the jump measures, the second includes the BNS jump statistic and the third includes the jump dummy variable. In accordance with our portfolio sorts, stocks with large Size, worse prior performance and fewer jumps (higher jump statistics and fewer significant jumps) exhibit higher memory parameters. The coefficients are all statistically significant at the 1% level. We additionally find that value stocks possess higher memory parameters, while illiquidity is not able to explain the degree of memory in volatility. Intuitively, stocks which tend to exhibit jumps more frequently, are less persistent and predictable and should possess lower memory parame-

[8]We experiment with multiple alternative estimation methods for long memory in order to make sure that the results are robust with respect to the estimation approach. The methods and results are reported in Section VI.C and are qualitatively similar.

ters. We show the close connection of long memory and predictability in Section IV and provide some intuition for how memory is generated for small (large) and loser (winner) stocks in Section V.

C Long Memory Volatility and Expected Stock Returns: Portfolio Sorts

In previous sections we relate the memory of volatility to firm-specific variables, trying to explain the degree of long memory. In the next step, we investigate whether investors demand a compensation for holding assets with higher exposure to this factor by looking at the relationship between the degree of memory in volatility and realized future excess stock returns. Assuming that the degree of memory in volatility is related to the predictability of a stock return's volatility, a highly predictable stock should be less uncertain than an unpredictable stock. We hence expect a negative price for long memory in order to compensate investors for the additional volatility risk of short memory stocks.[9]

As in Section III.B, each month, we sort all stocks into quintile portfolios where stocks with the lowest memory parameter are in the first quintile and stocks with the highest memory parameter are in the fifth quintile. Excess returns of the equally weighted portfolios are tracked over the subsequent month.[10] The analysis is out-of-sample in the sense that there is no overlap between the data used for the memory estimation and the data used to compute the excess returns of the portfolios. The LMS portfolio returns are then regressed on risk factors in order to test whether these returns merely reflect

[9]Section IV confirms the intuitive relationship of memory and predictability of a stock's volatility in a validity check.

[10]Since our memory estimates $d_{i,t}$ rely on rolling window estimates, one might argue that there is barely temporal variation in our estimates. If this is true, this should work against our empirical analysis and we should not find any significant relationship between memory and expected returns, but we do. In the robustness section, Section VI, we repeat the analysis, relying on monthly memory parameters estimated from high-frequency data in that month. The results are qualitatively similar.

passive exposure to standard factors. We include the market portfolio of the CAPM, which controls for systematic risk and the Fama & French (1993) three-factor model (FF3), which additionally includes the size and value effects. Further, we employ the state-of-the art Fama & French (2015) five-factor model (FF5) and the Hou et al. (2014) q-factor model (HXZ).[11] We investigate three different sample periods, which start in 1926, 1963 and 1967, respectively. All periods end in December 2015.[12]

The results are presented in Table 4. We report the mean return of the quintile portfolios and the LMS portfolio (Q5-Q1) in the first row. Below we report the alphas of the three different models. We find that the annualized mean return generally adheres to a decreasing pattern from 13.57% in the first quintile to 11.86% in the fifth quintile. All quintile portfolio returns are statistically significant, just like the difference of −1.71% between the long memory quintile and the short memory quintile (LMS). Controling for risk factors leads to alphas of −2.23%, −2.47%, −2.84% and −2.52% for the CAPM, Fama & French (1993) three-factor model, Fama & French (2015) five-factor and Hou et al. (2014) q-factor model, respectively. The risk adjusted returns are all statistically significant.[13]

Consequently, controling for standard risk factors does not affect our main result that the long memory volatility excess return trade-off is priced with a negative sign.[14]

[11]The factors for the first four models are available from the Kenneth R. French's data library, website:mba.tuck.dartmouth.edu/pages/faculty/ken.french. The factors of the Hou et al. (2014) model were kindly provided by the authors.

[12]The choice of different sample periods is motivated by the availability of the factor models. The Fama & French (2015) factors are available starting in 1963 while the Hou et al. (2014) factors are available starting in 1967.

[13]We focus on equally weighted portfolios. We have redone the analysis with value-weighted portfolios, which leads to a spread return of −2.27% and a FF5 alpha of −2.19%. Both are statistically significant at the 10% level.

[14]As shown in Section III.B, the memory parameter can be explained by firm characteristics such as size, jumps and momentum. Nonetheless, controling for the risk factors delivers statistically significant alphas. As an additional robustness check we investigate whether the isolated effect of long memory, which is orthogonal to firm size and other firm characteristics, is priced in the cross-section as well. Residual long memory is obtained by regressing the memory parameter on the firm characteristics at each point of time following Hong et al. (2000), Nagel (2005) and Hillert et al. (2014). We find a CAPM

D Long Memory Volatility and Expected Stock Returns: Regression Tests

The portfolio sorts present strong evidence that the degree of long memory in volatility is (negatively) related to future excess returns. We now estimate Fama & MacBeth (1973) regressions that simultaneously control for different variables and test if the degree of memory of a stock's volatility contains information about future excess returns beyond that of various other firm characteristics. This exercise, which relies on individual stock returns rather than stock portfolios, presents an alternative method in order to estimate the cross-sectional risk premium associated with long memory volatility. We rely on individual stocks rather than portfolio returns since the formation of portfolios in cross-sectional regressions is shown to influence the results and lead to higher standard errors of the risk premium estimates (Lo & MacKinlay, 1990; Ang et al., 2010; Lewellen et al., 2010). Each month, we regress excess stock returns over the following month on the stock characteristics of the current month:

$$r_{i,t+1} - r_{f,t+1} = \alpha_t + \gamma_t^M d_{i,t} + \gamma_t^C X_{i,t} + \epsilon_{i,t+1} \tag{11}$$

where $r_{i,t}$ is the return of stock i and $r_{f,t}$ is the risk-free rate at time t. $X_{i,t}$ is a vector containing the firm characteristics Size, Book-to-Market, Momentum, Illiquidity and Jumps.[15] γ_t^M and γ_t^C are the risk premia associated with the memory parameter and the remaining firm characteristics, respectively, and $\epsilon_{i,t}$ is the error term. In a second step we perform tests on the time-series averages of the estimated monthly intercept and slope

(FF5) alpha of −1.2% (−1.5%), which is statistically significant at the 10% level or lower. Results are reported in Table 7 in the online appendix.

[15]We use the same firm charateristics as in our portfolio sorts in Section III B. We include further control variables such as the market beta, idiosyncratic volatility and more in the robustness section, Section VI.

coefficients in order to test for significance of the risk premia $\hat{\gamma}_t^M$ and $\hat{\gamma}_t^C$ over the sample period.

Table 5 reports the results of the Fama & MacBeth (1973) regressions presenting the time-series averages of the coefficients, $\hat{\alpha}_t$, $\hat{\gamma}_t^M$ and $\hat{\gamma}_t^C$. Model 1 regresses the excess return of stocks over the following month on the memory parameter only. The market price of long memory is -0.0039, which is statistically significant at the 5% level. Consequently, a stock with anti-persistent volatility can earn average annualized returns of up to 4.7% higher than a stock with long memory volatility.[16] Models 2 to 6 additionally include one of the firm characteristics in the cross-sectional regression. The magnitude and significance of the memory risk premium is slightly reduced when adding Size but barely changes when adding Book-to-Market, Momentum, Illiquidity or Jumps. Nonetheless, the coefficient $\hat{\gamma}^M$ remains statistically significant for all models. The negative (positive) risk premium for Size (Book-to-Market, Momentum and Illiquidity) is consistent with the literature (Fama & French, 1992; Jegadeesh & Titman, 1993; Amihud, 2002). Results are qualitatively similar for the kitchen sink regression (Model 7) where the coefficient of the memory parameter remains statistically significant.

IV Long Memory Volatility and Predictability

A possible explanation for the negative relationship between long memory volatility and expected stock returns is the uncertainty around a stock's volatility. As discussed earlier, long memory represents the hyperbolic decay of the autocorrelation function which on the other hand allows for (high and long-run) volatility predictability. One can argue

[16] The lowest possible memory parameter for a anti-persistent stock is given by the lower bound of the interval $(-0.5; 0)$ while the highest possible stationary memory parameter is given by the upper bound of the interval $(0; 0.5)$. The highest possible annualized spread returns can thus be approximated by $1 * (-0.0039) * 12 = -0.0468$.

that in times of financial distress large negative shocks are more persistent for stocks with long memory, which makes these stocks less favorable than short memory stocks. But even though negative shocks are more persistent, the volatility predictability is still higher for long memory stocks, which makes them less uncertain regarding their level of risk.

In Sections III.C and III.D, we provide evidence that stocks with long memory volatility earn on average lower returns than stocks with short memory using both portfolio sorts and cross-sectional regressions. In this section, we supply empirical evidence that long memory is associated with predictability and hence confirm our channel of negative expected returns through volatility uncertainty. Further, this exercise is a validity check of our long memory estimates. If our memory estimates are not biased by data quality or spurious long memory, a higher memory parameter should be directly linked to higher forecasting performance.[17]

For each stock, we conduct monthly predictability regressions of realized volatility both in-sample and out-of-sample. The time series of monthly realized volatility is obtained by summing squared daily returns for each month (Bollerslev et al., 2014). Following the spirit of Corsi (2009), we use (heterogenous) autoregressive models of realized volatility (HAR-RV).[18] The regressions include lagged observations of the realized volatil-

[17]We acknowledge the issue of spurious long memory where higher memory parameters can be caused by structural breaks. Even though we work with rolling window estimates, which should be only marginally affected by breaks, we control for this in three different ways. First, both our portfolio sorts and cross-sectional regressions include the BNS jump statistic and the alpha or long memory risk premium remain statistically significant. Hence, our results are not driven by the BNS variable. Second, the validity check in this section relates the memory parameter to predictability. If our parameters are biased by structural breaks or jumps, we should not find any clear relationship, however we do. Third, we repeat our portfolio sorts but rely on returns purged from jumps following Pukthuanthong & Roll (2015). Buying stocks with long memory volatility and selling stocks with short memory volatility, where long memory is estimated from raw returns, leads to a statistically significant spread return of -1.73% and a Fama & French (2015) five-factor alpha of -2.89%, which is statistically significant as well. Both are of similar magnitudes to those in our main analysis.

[18]We also experimented with simple autoregressive (AR) models including the lags 1, 6, 12, 24 and 60, leading to qualitatively similar results.

ity and we allow for five different specifications by including the volatility from the previous month (HAR(1)), six months (HAR(2)), one year (HAR(3)), two years (HAR(4)) and 5 years (HAR(5)):

$$HAR(1) : RV_{t+1}^{M} = \alpha + \beta RV_t^{M} + \epsilon_{t+1} \tag{12}$$

$$HAR(2) : RV_{t+1}^{M} = \alpha + \beta RV_t^{M} + \beta RV_t^{6M} + \epsilon_{t+1} \tag{13}$$

$$HAR(3) : RV_{t+1}^{M} = \alpha + \beta RV_t^{M} + \beta RV_t^{6M} + \beta RV_t^{1Y} + \epsilon_{t+1} \tag{14}$$

$$HAR(4) : RV_{t+1}^{M} = \alpha + \beta RV_t^{M} + \beta RV_t^{6M} + \beta RV_t^{1Y} + \beta RV_t^{2Y} + \epsilon_{t+1} \tag{15}$$

$$HAR(5) : RV_{t+1}^{M} = \alpha + \beta RV_t^{M} + \beta RV_t^{6M} + \beta RV_t^{1Y} + \beta RV_t^{2Y} + \beta RV_t^{5Y} + \epsilon_{t+1} \tag{16}$$

The multiperiod volatilities are normalized sums of the one-month realized volatilities. The six-months' realized volatility is exemplarily given by:

$$RV_t^{6M} = \frac{1}{6}(RV_t^{M} + RV_{t-1}^{M} + ... + RV_{t-5}^{M}) \tag{17}$$

Despite the simplicity of these models, they are shown to be able to mimic long memory behavior and exhibit good forecasting performance. We form quintile portfolios by sorting the cross-section of stock returns by the memory parameter. We then compute the average adjusted R^2, F-statistic and out-of-sample R^2_{OOS} for each quintile portfolio.[19] The calculation of the out-of-sample R^2_{OOS} follows Campbell & Thompson (2008), and measures the differences in mean squared prediction errors (MSPE) for the predictive model, Equations (12)-(16) and the historical mean.

The results are reported in Table 6. Panel A shows the adjusted R^2 of the in-

[19]We report t-statistics of the slope coefficient for HAR(1) and F-statistics for the joint significance of the slope coefficients for the remaining models. For the out-of-sample analysis, the R^2_{OOS} for some stocks show extremely bad performance, with values below -100% due to large spikes. We winsorize the data at the 1% and 99% level to minimize the effect of these outliers. Cleaning the time series of the outliers delivers qualitatively similar results.

sample predictability regressions. There is a strictly monotonic pattern of explanatory power, which is increasing in the memory parameter. This is supported by the increasing t-statistics and F-statistics in Panel B. Stocks with higher memory parameters show stronger explanatory power and the predictor variables are more statistically significant than stocks with lower memory parameters. Lastly, the R^2_{OOS} also show that the out-of-sample forecasting performance of long memory stocks is stronger than short memory stocks and exhibits a generally monotonic pattern. A graphical illustration of the results is presented in Figure 1. One can see that the bars are monotonically increasing for all five models and all three colors (adj R^2, F-statistic and R^2_{OOS}).

We thus show that the memory of stocks is a proxy for predictability, which explains the negative spread returns of the LMS portfolio. At the same time, this exercise validates our estimation approach to memory. Our results are true for both in-sample and out-of-sample, while we allow for various model specifications including short memory processes and long memory mimicking processes.

V Implication for Existing Models

In this section we discuss the connection of our empirical results with theoretical models of how long memory in volatility is generated for individual stocks using the proposed "Agent-based" model of LeBaron (2006) and the "Interacting Agent View" of Alfarano & Lux (2007). These models rely on heterogeneity across market agents. Müller et al. (1993), Peters (1994) and Corsi (2009) also consider markets with heterogenous traders. Motivated by the memory-generating models, we discuss how large and loser stocks in these models differ from small and winner stocks.[20]

[20]For these characteristics, we find statistical significance concerning memory parameter spreads for both portfolio sorts and cross-sectional regressions.

A Interacting Agent View

Alfarano & Lux (2007) divide traders in a market into two groups – fundamentalists and chartists – whose interactions are based on the mechanism introduced by Kirman (1993). The noise traders (chartists) are driven by herd instincts and buy (sell) if they are optimistic (pessimistic). The long memory in volatility is then generated by the interaction of agents with heterogenous beliefs and strategies. The numbers of fundamentalists and chartists are fixed, but transition from optimists to pessimists and vice versa is allowed by a two-state model. They derive an equilibrium distribution with two equilibria where a transition between them has a finite probability. The average time for the transition is denoted as the mean first passage time T_0. From the ratio of mean first passage time T_0 and available data observations T, conclusions on the memory of the process can be drawn. For higher T_0 relative to T, the memory parameter of squared returns decreases starting with a Hurst exponent close to 1 and converging to 0 for $T >> T_0$. The mean first passage time is negatively related to the number of agents N in the market. We divide the cross-section of stock returns into several segments by firm characteristics. The relation of T and T_0 for each submarket allows for conclusions on the memory of the submarket. We focus on the effect of these two variables, assuming that all other variables are the same for the two markets in comparison.[21]

First, our main analysis shows that stocks with higher market capitalization exhibit longer memory in volatility. Gompers & Metrick (2001) find that the demand for large and liquid stocks has grown due to the increasing share of the U.S. equity market. Additionally, investment decisions in small stocks are harder to justify to sponsors by professional managers, as argued by Lakonishok et al. (1992). Further, Merton (1987) argues that

[21] The impact of other variables is neglible, since the memory parameter is high for low T relative to T_0 and always converges to zero for $T \to \infty$.

small stocks exhibit incomplete information. This makes smaller stocks less favorable as well. All these findings suggest that the number of investors in large stocks dominates those of small stocks. The larger number of agents for large stocks leads to a higher mean first passage time and hence intuitively to longer memory in volatility, as we empirically document.

Second, we find that stocks with longer memory in volatility tend to be loser stocks. This result can be explained by the disposition effect, as labeled by Shefrin & Statman (1985). The effect states that investors tend to hold on their losing stocks too long and sell their winner stocks too soon in financial markets. This effect can be explained in the context of the prospect theory of Kahneman & Tversky (1979) and the mental accounting framework of Thaler (1980). The results suggest that the number of agents investing in winner stocks tends to decrease while the number for the loser stocks tends to remain constant or even increase. This leads to longer memory for loser stocks, as shown in our main analysis.

B Agent-based Models

LeBaron (2006) divides the market into groups according to their investment horizon and hence considers a heterogenous agent framework. The agents rely on past information such as lagged returns, dividend–price ratios and trend indicators to evaluate rules for investment decisions. This evaluation varies across agents. Some agents rely only on more recent data, e.g. only the past six months (short memory investor), while others use thirty years' worth of data (long memory investor). The trading rules may evolve over time and a Walrasian equilibrium is reached by clearing the market. The author shows that in a market consisting of homogeneous investors (long memory investors only), the price converges to the equilibrium price through the learning mechanism, which results in a short memory process for squared returns. If the market consists of all types of agents (all memory), on the other hand, the price takes large swings from the equilibrium and large crashes and shows long memory behavior for volatility. The persistence is driven by the short-term investors. As argued by Corsi (2009), short-term investors are influenced by the long-term variance, which again has an impact on the short-term variance while long-term investors are not influenced by changes in short-term volatility. The model can be transferred to parts of the complete market as proxied by the cross-section of U.S. stock returns. We compare the fraction of short- and long-term investors in various markets and conclude on the degree of memory in these markets.

Perez-Quiros & Timmermann (2000) argue that small firms with little collateral show the highest asymmetry in their risk across recession and expansion states. Their expected returns are thus more sensitive to credit market conditions. Chan & Chen (1991) present similar arguments for small firms being more sensitive to news about the state of the business cycle. This implies that investors in small firms are generally mid- to long-term oriented, while investors of large and better collateralized firms may be both short- and long-term oriented. This is supported by the argument from the "Interacting Agent View model". Large stocks are more favorable and hence attract all different kinds of agents. The higher degree of heterogeneity of large firm investors lead to the higher memory parameter compared to small firms.[22]

Daniel & Moskowitz (2016) argue that the momentum strategy generates abnormally high returns on average, but at the same time experiences abnormally high losses. This is because the loser stocks embed features of a short call option on the market portfolio. Especially during times of volatile bear markets, past-loser stocks lose a large fraction of

[22]Even though small firm investors are rather short-term oriented, this does not mean there are no long-term investors. The same is true for large firms. Hence we consider the relative proportion of long-term and short-term investors and talk about the degree of heterogeneity.

their market value and contain high financial leverage. The equity of theses firms is similar to out-of-the money call options on the underlying firm values, which are correlated with the market. This implies that loser stocks are much more sensitive to the state of the market (turbulent vs. calm), which may change quickly. Consequently, the fraction of short-term investors in the market of loser stocks should be larger than in the market of winner stocks, which leads to higher memory estimates for loser stocks. This is what we find empirically.

VI Extensions and Robustness Tests

In this section we run further analyses of long memory volatility in the cross-section and various robustness tests including alternative estimators and portfolio sorts, and extend our cross-sectional analysis with further control variables. Detailed results are reported in the online appendix.

A Long Memory Volatility and Industries

In Section III we consider different firm characteristics and how they are able to explain the memory parameter of volatility in the cross-section of U.S. returns. We find that higher memory parameters can be related to large, loser stocks and stocks with fewer jumps. In this section, we investigate whether firms in certain industries possess higher or lower memory parameters. More specifically, we use the twelve industry portfolio identifiers obtained from Kenneth R. French's data library. The industries are Consumer Non-Durables, Consumer Durables, Manufacturing, Energy, Chemicals, Business Equipment, Telecommunication, Utilities, Shops, Healthcare, Money & Finance and Others. We apply the GPH estimator and a bandwidth parameter of $m = N^{0.5}$ as in our main analysis. Table 8 of the online appendix reports the results. The mean and median are very close to the value for the complete cross-section (0.22). Since the degree of memory is similar for all industries, industry codes, unlike firm characteristics, are not able to explain the cross-sectional variation of the memory parameter.

B Fama–French Portfolios

In Section III we sort stocks by their memory parameter and investigate the average firm characteristics of quintile portfolios. In this section, we validate our results by comparing the memory of Fama–French decile portfolios, which are sorted by size, book-to-market or momentum. There are two major differences with this approach. First, instead of sorting by the memory parameter, stocks are sorted by their firm characteristics. Second, we consider decile instead of quintile portfolios.[23] The portfolio returns are obtained from Kenneth R. French's data library. We apply the GPH estimator with the bandwidth parameter of $m = N^{0.5}$ as in our main analysis and report the memory parameter for each decile portfolio and the high-minus-low ($D10 - D1$) in Table 9 of the online appendix. Consistent with our main results, portfolios with larger size, higher book-to-market and worse prior performance exhibit higher memory parameters.[24] The book-to-market (momentum) portfolios demonstrate a monotonically increasing (decreasing) pattern in memory.

C Estimation of the Memory Parameter

For our main analysis we follow the existing literature and choose the ad hoc bandwidth parameter of $m = N^{0.5}$. We repeat the estimation using a bandwidth parameter of

[23] The results for the Fama–Frech quintile portfolios are qualitatively similar.

[24] The magnitude of the memory parameters are somewhat higher than in our main analysis. This is because we here use the complete time series of daily returns over more than 60 years, compared to the 5 years in our main analysis.

$m = N^{0.6}$, $m = N^{0.7}$ and $m = N^{0.8}$ and alternative estimators in this section and report the results in Tables 10, 11 and 12 of the online appendix.[25]

We report the portfolio sorts for the cross-section of U.S. returns using the GPH estimator and alternative parameters in Table 10, Panels A, B and C. We find that sorting by the memory parameter and holding stocks with long memory and selling stocks with short memory still generates negative excess returns. Using the alternative bandwidth parameters $m = N^{0.6}$, $m = N^{0.7}$ and $m = N^{0.8}$ leads to returns of -1.80%, -2.71% and -2.32% per annum, respectively. Adjusting for the additional risk factors of the Fama & French (2015) model leads to significant alphas of similar magnitudes as in our main analysis.

Further, we apply the GPH estimator to the absolute returns rather than the squared returns as in our main analysis (Bollerslev & Wright, 2000). The results are reported in Panel D and are consistent with our main findings. Stocks with short memory earn on average 2.94% per annum more than stocks with long memory. This spread return is statistically significant at the 1% level and remains significant when controling for the Fama & French (2015) risk factors.

A commonly used alternative approach to estimate long memory is the Local Whittle (LW) estimator. We repeat the estimation with the LW estimator and the same bandwidth parameter as in our main results, $m = N^{0.5}$ (Bandi & Perron, 2006). Results are provided in Table 11. For the portfolio sorts we find a negative spread return of 2.09% for the LMS portfolio which is statistically significant at the 5% level (Panel A). The Fama & French (2015) five-factor alpha with a value of -3.21% is statistically significant as well. In addition, we apply the LW estimator with bandwidth parameters of $m = N^{0.6}$,

[25] These alternative bandwidth parameters are the most common choices in the literature, see Hurvich & Ray (2003), Hurvich et al. (2005), Bandi & Perron (2006), Berger et al. (2009), Hou & Perron (2014), among others, and include the MSE-optimal one for the GPH estimator.

$m = N^{0.7}$ and $m = N^{0.8}$ to the squared returns and a bandwidth parameter of $m = N^{0.5}$ to the absolute returns. Panels B to E report the results. The spread returns are all negative, varying from −1.82% to −3.03%, and the Fama & French (2015) five-factor alphas vary from −2.54% to −3.93%, while all returns and risk-adjusted returns are statistically significant.

Table 12 reports the coefficient estimates from the cross-sectional regressions in Equation (11) using the alternative long memory estimator and bandwidths. We rely on simple regressions where individual stock returns are regressed on the long memory parameter in Panel A and multiple regressions where we additionally include Size, Book-to-Market, Momentum, Illiquidity and the BNS jump test statistic as explanatory variables. The results are consistent with our main analysis. For the simple regressions we find that long memory is negatively priced in the cross-section with a risk premium estimate varying from −0.0104 to −0.0039, depending on the estimator and bandwidth, which is statistically significant at the 5% level or lower. Including the control variables slightly changes the magnitude of the long memory premium but they remain statistically significant. In addition, we find a negative (positive) price for the size (book-to-market ratio and momentum) of a stock which is consistent with both our main analysis and the literature.

D Holding Period Returns

In our main analysis, portfolios are rebalanced monthly and held for one month. We now track whether the negative risk premium associated with long memory volatility persists for longer holding periods. Each month, we sort all stocks into quintile portfolios where stocks with the lowest memory parameter are in the first quintile and stocks with the highest memory parameter are in the fifth quintile. Excess returns of the portfolios are tracked over the subsequent one, two, three, four and five years. To account for the overlapping returns, we adjust the standard errors following Newey & West (1987), using lags according to the return horizon expressed in months.

The results are reported in Table 13 of the online appendix. Average returns and Fama & French (2015) risk adjusted returns for the one-, two-, three-, four- and five-year holding period are reported in Panels A, B, C, D and E. The annualized mean returns are of similar magnitude as for the one-month holding period. The LMS spreads are −1.88%, −1.93%, −1.88%, −1.90% and −1.91%, respectively, and are all statistically significant at the 5% level or lower. The risk adjusted returns only change slightly, and vary between −1.66% to −2.29% and are generally statistically significant.

E High-Frequency Data

We repeat our analysis, but rely on high-frequency instead of daily returns. We obtain 5-min returns for the S&P 500 constituents for the period from 1996 until 2015 from Thomson Reuters Tick History. Our choice of the sample period and stocks is restricted by their availability. The data is cleaned following Barndorff-Nielsen et al. (2009). Zhang et al. (2005) argue that high-frequency data should always result in a more accurate estimate when used correctly due to the basic statistical principle that more data are always better. Bollerslev & Wright (2000) show that the high-frequency data allow for a superior and nearly unbiased estimation of the long memory parameter using 5-min return observations. We apply the GPH estimator and a bandwidth parameter of $m = N^{0.5}$ to a month of 5-min returns, which counts up to 1738 ($= 22 * 79$) data points per estimation window. This window is comparable to 8 years of daily observations.

The results are reported in Table 14 of the online appendix. We find a negative return of −8.83% for the LMS portfolio, which is statistically significant at the 1% level. Controling for additional risk factors generally slightly mitigates the risk premium but the alphas remain significant. This section thus confirms our main results and shows that the negative risk premium is not dependent on the source and frequency of data and the sample period. We implicitly investigate four subsamples and thus show that our main results are robust against various sample periods. Our choice of subsamples is motivated by the availability of the data. The longest period from 1926 until 2015 is chosen according to the availability of the CRSP stock data. We control for Fama & French (2015) (Hou et al., 2014) risk factors, which are available from 1963 until 2015 (1967 until 2015). Lastly, we also investigate the most recent 20 years from 1996 until 2015, which is chosen due to the availability of high-frequency data from Thomson Reuters Tick History.

F Additional Control Variables

In Section III.D we conduct regression tests including Size, Book-to-Market, Momentum, Illiquidity and Jumps. We now also control for further effects and anomalies which have been shown to be good predictors of expected returns. More specifically, we include the market beta (BETA), reversal (REV), cokurtosis with the market (CKT), coskewness with the market (CSK), idiosyncratic volatility (IVOL), realized kurtosis (KURT), realized skewness (SKEW) and demand for lottery (MAX). Further, we include a stock's volatility-of-volatility (Vol-of-Vol). In our empirical analysis we relate the long memory of volatility to the predictability of volatility and uncertainty. We relate higher volatility predictability to lower uncertainty regarding a stock's level of risk. In the literature, uncertainty has been measured by the volatility-of-volatility for both individual stocks and the aggregate market (Baltussen et al., 2016; Hollstein & Prokopczuk, 2017).[26] We calcu-

[26]Both studies investigate the asset pricing implication of the volatility-of-volatility and find a negative price, just as we find for long memory.

late the volatility-of-volatility as the 5-year rolling window volatility of monthly realized volatility.[27] We find an average cross-sectional correlation of 0.11 between the degree of long memory volatility of a stock and its volatility-of-volatility. While both are intuitively related to uncertainty, the measures are barely correlated and we hence do not expect that our findings can be explained by the volatility-of-volatility of a stock. The market beta is estimated from daily return regressions of excess stock returns on an intercept and the market excess return over the examined period. Following Ang et al. (2006b), idiosyncratic volatility equals the standard deviation of the residuals from the same regression as for the market beta, but additionally includes the size and book-to-market factors of the Fama & French (1993) model. The short-term reversal at the end of a month is defined as the return of that month following Jegadeesh (1990). The coskewness and cokurtosis of a stock at the end of a month is estimated from the daily returns in that month following Ang et al. (2006a). The kurtosis and skewness of a stock at the end of a month is given by the sample kurtosis and skewness estimated from the daily returns in that month. Lastly, the demand for lottery is given by the maximum total daily return observation of a month (Bali et al., 2011).

Table 15 of the online appendix presents the results of the cross-sectional regressions. Models 7 to 15 show the time-series averages of the additional coefficients in multiple regressions. Most importantly, the risk premium of the long memory volatility remains negative and statistically significant for all additional control variables, varying from -0.0043 to -0.0036. The signs of statistically significant risk premia for variables besides long memory are generally consistent with the literature. Frazzini & Pedersen (2014) find that portfolios with higher betas have lower alphas and Sharpe ratios than portfo-

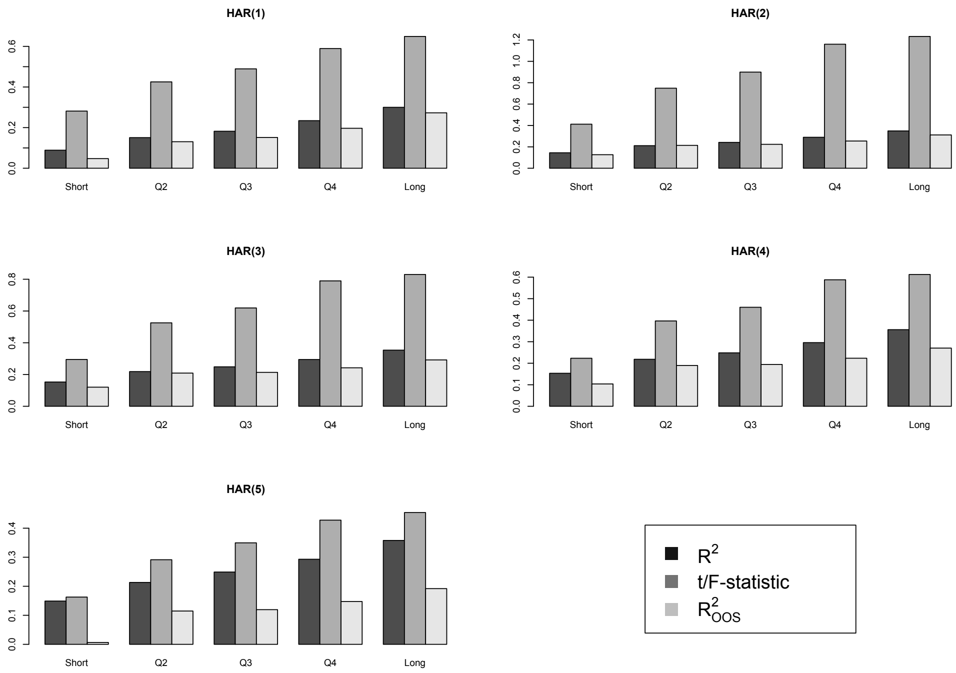
[27] It is not possible to compute the measure of Baltussen et al. (2016) for our sample since they rely on options data of individual stocks which are available starting in 1996 from OptionMetrics. Our approach for calculating the volatility-of-volatility closely follows the approach for our long memory estimates.

lios of low-beta assets. Amaya et al. (2015) show that buying stocks with low realized skewness and selling stocks with high realized skewness generates statistically significant and positive excess returns at a weekly frequency while there is no clear relationship for realized kurtosis. The negative and statistically significant premium for idiosyncratic volatility is consistent with the results of Ang et al. (2006b). Bali et al. (2011) argue that investors are willing to pay more for stocks that exhibit extreme positive returns. As a consequence, these stocks exhibit lower future returns, which is consistent with the negative premium we find. Model 16 includes the memory parameter and all additional control variables in this section while Model 17 presents the kitchen sink regression. The coefficient of the memory parameter remains statistically significant at the 5% level or lower.

VII Conclusion

In this paper we shed new light on the asset pricing implication of long memory in stock return volatility. Using portfolio sorts and cross-sectional regressions, we analyze how the degree of long memory of a firm's return volatility can be explained by its size, book-to-market, prior performance or jumps. Based on existing theoretical models, we discuss how long memory is generated in high market capitalization (winner) stocks compared to low market capitalization (loser) stocks. We estimate a cross-sectional price of long memory of -4.7% per annum. This estimate is robust to controling for size, value, momentum, liquidity effects and more. We relate the compensation for holding short memory stocks to higher risk, which is given by the low predictability of short memory stocks. Our results are robust against different variations of the estimation approach and the examined models.

Figure 1: Predictability of Quintile Portfolios



This figure reports adjusted R^2, F-statistics and R^2_{OOS} for quintile portfolios of the cross-section of U.S. stock returns. For comparison reasons the test statistics are all divided by 100.

Table 1: Summary Statistics

This table presents summary statistics for the memory estimates of individual stocks' volatility. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. Obs. stands for the average number of long memory estimates per month. SD stands for the standard deviation. The second column reports selected quantiles of the averages. t-statistic reports the mean t-statistic. Sign. at 5% reports the proportion of significant long memory estimates, while the remainder of the last column reports the proportion of the memory parameter being in a certain interval.

Descriptive		Quantiles		Memory	
Obs.	2479.73	5%	0.04	t-statistic	23.34
Mean	0.22	25%	0.15	Sign. at 5%	0.96
SD	0.12	Median	0.22	$-0.5 < d < 0.0$	0.03
Skewness	0.40	75%	0.29	$0.0 < d < 0.5$	0.95
Kurtosis	1.48	95%	0.43	$0.5 < d < 1$	0.02

Table 2: Portfolio Sorts and Characteristics

This table presents firm characteristics of portfolios sorted by the memory of volatility. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. From 1950 until 2015 we sort stocks each month and form and hold the portfolio for one month. We report the average long memory parameter, size, momentum and illiquidity, BNS statistic and BNS indicator function of quintile portfolios. The Q5-Q1 column reports the averages for the long memory minus short memory portfolio (LMS) with the according t-statistics in square brackets.

	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (LMS)	
Memory	0.0044	0.1295	0.2118	0.2975	0.4471	0.4427	[202.7567]
Size	11.6610	11.8630	12.0161	12.1707	12.3560	0.6950	[23.3435]
Book-to-Market	0.8934	0.9168	0.8993	0.8758	0.8996	0.0062	[0.5910]
Momentum	0.1681	0.1558	0.1522	0.1483	0.1284	−0.0397	[−14.2697]
Illiquidity	0.0044	0.0040	0.0038	0.0040	0.0055	0.0010	[3.9205]
BNS	−0.1994	−0.0620	−0.0255	−0.0110	0.0036	0.2030	[12.5035]
BNS-I	0.0177	0.0126	0.0106	0.0087	0.0074	−0.0103	[−24.2588]

Table 3: Cross-sectional Regression

This table presents the results from cross-sectional regressions for the period from 1950 until 2015. Each month, we regress the memory parameter of the cross-section on size, book-to-market, momentum, illiquidity and BNS. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. We report the average β coefficients and the according standard errors in parentheses below. The first row excludes any jump measures. The second row includes the BNS jump statistic while the thrid row includes the BNS jump indicator. Stars indicate significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Intercept	Size	Book-to-Market	Momentum	Illiquidity	BNS	BNS-I
β	0.0292***	0.0160***	0.0019***	−0.0186***	0.3126		
	(0.0061)	(0.0006)	(0.0006)	(0.0013)	(0.2696)		
β	0.0286***	0.0161***	0.0017***	−0.0184***	0.3720	0.0052***	
	(0.0061)	(0.0006)	(0.0006)	(0.0013)	(0.2738)	(0.0005)	
β	0.0301***	0.0160***	0.0018***	−0.0185***	0.3701		−0.0491***
	(0.0061)	(0.0006)	(0.0006)	(0.0013)	(0.2754)		(0.0025)

Table 4: Sorted Portfolio Returns

This table reports average returns and risk-adjusted returns of equally weighted quintile portfolios for the period from 1926 until 2015. Each month, stocks are sorted by the degree of long memory in volatility and we track the portfolio returns over the subsequent month. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. The one-month-ahead portfolio returns are regressed on risk factors in the Capital Asset Pricing Model (CAPM), the Fama & French (1993) 3-factor model (FF3), the Fama & French (2015) 5-factor model (period starts in 1963) (FF5) and the Hou et al. (2014) q-model (period starts in 1967) (HXZ). The corresponding alphas are reported. We report Newey–West standard errors using lags equal to the return horizon in parentheses. Stars indicate significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (LMS)
Mean return	0.1357***	0.1288***	0.1344***	0.1263***	0.1186***	−0.0171**
	(0.0334)	(0.0326)	(0.0343)	(0.0346)	(0.0356)	(0.0086)
CAPM	0.0385***	0.0328***	0.0337***	0.0238**	0.0162	−0.0223***
	(0.0125)	(0.0115)	(0.0110)	(0.0103)	(0.0108)	(0.0083)
FF3	0.0136**	0.0103**	0.0084*	−0.0016	−0.0111*	−0.0247***
	(0.0062)	(0.0051)	(0.0048)	(0.0048)	(0.0062)	(0.0077)
FF5	0.0238**	0.0146*	0.0137*	0.0045	−0.0046	−0.0284***
	(0.0108)	(0.0087)	(0.0076)	(0.0075)	(0.0095)	(0.0099)
HXZ	0.0450***	0.0340***	0.0335***	0.0270**	0.0198	−0.0252*
	(0.0160)	(0.0129)	(0.0114)	(0.0113)	(0.0133)	(0.0129)

Table 5: Fama–MacBeth Regressions

This table reports results from Fama–MacBeth regressions for the period from 1950 until 2015. Each month, excess stock returns are regressed on lagged firm characteristics including the memory parameters, market capitalization (Size), book-to-market values, prior returns (Momentum), illiquidity and jump statistics (BNS). The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. Stars indicate significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Intercept	0.0091***	0.0144***	0.0075***	0.0076***	0.0087***	0.0091***	0.0106*
	(0.0025)	(0.0051)	(0.0025)	(0.0025)	(0.0025)	(0.0025)	(0.0056)
Long Memory	−0.0039**	−0.0021*	−0.0038**	−0.0038**	−0.0044***	−0.0043***	−0.0024**
	(0.0016)	(0.0012)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0011)
Size		−0.0006*					−0.0005
		(0.0003)					(0.0003)
Book-to-Market			0.0019***				0.0024***
			(0.0005)				(0.0006)
Momentum				0.0067***			0.0095***
				(0.0016)			(0.0013)
Illiquidity					0.2010**		0.0991
					(0.1010)		(0.1768)
BNS						0.0024***	0.0020***
						(0.0004)	(0.0003)

Table 6: Long Memory and Predictability

This table reports results of predictive regressions. We run heterogenous autoregressive regressions of the monthly realized variance for each stock including the previous one, six, twelve, twenty-four and sixty observations. We form quintile portfolios where stocks with the lowest memory parameter are in the first quintile and stocks with the highest memory parameter in the fifth quintile portfolio. We report average adjusted R^2 in Panel A, average t-statistics and F-statistics in panel B and out-of-sample R^2 in Panel C.

	Q1	Q2	Q3	Q4	Q5
Panel A: Adjusted R^2					
HAR(1)	0.0888	0.1507	0.1822	0.2343	0.3000
HAR(2)	0.1447	0.2111	0.2418	0.2897	0.3491
HAR(3)	0.1529	0.2185	0.2486	0.2946	0.3536
HAR(4)	0.1535	0.2184	0.2484	0.2958	0.3561
HAR(5)	0.1491	0.2132	0.2490	0.2931	0.3579
Panel B: T-statistic/F-statistic					
HAR(1)	5.6276	8.5058	9.7878	11.7858	12.9780
HAR(2)	41.2025	74.8700	89.9142	116.0092	123.2804
HAR(3)	29.4787	52.5572	61.9348	78.9834	82.9948
HAR(4)	22.3186	39.6614	46.0103	58.7847	61.2399
HAR(5)	16.2773	29.1439	34.9617	42.7776	45.3960
Panel C: R^2_{OOS}					
HAR(1)	0.0474	0.1306	0.1515	0.1967	0.2729
HAR(2)	0.1266	0.2139	0.2237	0.2546	0.3117
HAR(3)	0.1203	0.2090	0.2136	0.2424	0.2921
HAR(4)	0.1039	0.1896	0.1944	0.2233	0.2704
HAR(5)	0.0064	0.1147	0.1194	0.1475	0.1919

References

Alfarano, S., & Lux, T. (2007). A noise trader model as a generator of apparent financial power laws and long memory. *Macroeconomic Dynamics*, *11*(S1), 80–101.

Amaya, D., Christoffersen, P., Jacobs, K., & Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, *118*(1), 135–167.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56.

Ang, A., Chen, J., & Xing, Y. (2006a). Downside risk. *Review of Financial Studies*, *19*(4), 1191–1239.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006b). The cross-section of volatility and expected returns. *Journal of Finance*, *61*(1), 259–299.

Ang, A., Liu, J., & Schwarz, K. (2010). Using individual stocks or portfolios in tests of factor models. Working Paper, Columbia University.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, *73*(1), 5–59.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*(1), 3–30.

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, *99*(2), 427–446.

Baltussen, G., Van Bekkum, S., & Van der Grient, B. (2016). Unknown unknowns: Vol-of-vol and the cross section of stock returns. *Journal of Financial and Quantitative Analysis, forthcoming*.

Bandi, F. M., & Perron, B. (2006). Long memory and the relation between implied and realized volatility. *Journal of Financial Econometrics*, *4*(4), 636–670.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *Econometrics Journal*, *12*(3), 1–33.

Barndorff-Nielsen, O. E., & Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, *4*(1), 1–30.

Beran, J., Feng, Y., Ghosh, S., & Kulik, R. (2013). Long-memory processes: Probabilistic properties and statistical methods. *Springer, Heidelberg, 2013*.

Berger, D., Chaboud, A., & Hjalmarsson, E. (2009). What drives volatility persistence in the foreign exchange market? *Journal of Financial Economics*, *94*(2), 192–213.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business*, *45*(3), 444–455.

Bollerslev, T., Marrone, J., Xu, L., & Zhou, H. (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence. *Journal of Financial and Quantitative Analysis*, *49*(3), 633–661.

Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, *73*(1), 151–184.

Bollerslev, T., & Wright, J. H. (2000). Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data. *Journal of Econometrics*, *98*(1), 81–106.

Breidt, F. J., Crato, N., & De Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, *83*(1), 325–348.

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

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*(1), 57–82.

Chan, K., & Chen, N.-F. (1991). Structural and return characteristics of small and large firms. *Journal of Finance*, *46*(4), 1467–1484.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*(2), 174–196.

Cremers, M., Halling, M., & Weinbaum, D. (2015). Aggregate jump and volatility risk in the cross-section of stock returns. *Journal of Finance*, *70*(2), 577–614.

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, *122*(2), 221–247.

Diebold, F. X., & Rudebusch, G. D. (1989). Long memory and persistence in aggregate output. *Journal of Monetary Economics*, *24*(2), 189–209.

Ding, Z., & Granger, C. W. (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, *73*(1), 185–215.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, *47*(2), 427–465.

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

Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *Journal of Finance*, *63*(4), 1653–1678.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25.

Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, *4*(4), 221–238.

Gompers, P. A., & Metrick, A. (2001). Institutional investors and equity prices. *Quarterly Journal of Economics*, *116*(1), 229–259.

Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, *16*(1), 121–130.

Granger, C. W., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, *1*(1), 15–29.

Henry, O. T. (2002). Long memory in stock returns: Some international evidence. *Applied Financial Economics*, *12*(10), 725–729.

Hillert, A., Jacobs, H., & Müller, S. (2014). Media makes momentum. *Review of Financial Studies*, *27*(12), 3467–3501.

Hollstein, F., & Prokopczuk, M. (2017). How aggregate volatility-of-volatility affects stock returns. *Working Paper*.

Hong, H., Lim, T., & Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance*, *55*(1), 265–295.

Hosking, J. R. (1981). Fractional differencing. *Biometrika*, *68*(1), 165–176.

Hou, J., & Perron, P. (2014). Modified local Whittle estimator for long memory processes in the presence of low frequency (and other) contaminations. *Journal of Econometrics*, *182*(2), 309–328.

Hou, K., Xue, C., & Zhang, L. (2014). Digesting anomalies: An investment approach. *Review of Financial Studies*, *28*, 650–705.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *American Society of Civil Engineers*, *116*, 770–799.

Hurvich, C. M., & Deo, R. S. (1999). Plug-in selection of the number of frequencies in regression estimates of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, *20*(3), 331–341.

Hurvich, C. M., Moulines, E., & Soulier, P. (2005). Estimating long memory in volatility. *Econometrica*, *73*(4), 1283–1328.

Hurvich, C. M., & Ray, B. K. (2003). The local Whittle estimator of long-memory stochastic volatility. *Journal of Financial Econometrics*, *1*(3), 445–470.

Jacod, J., & Todorov, V. (2009). Testing for common arrivals of jumps for discretely observed multidimensional processes. *The Annals of Statistics*, *37*, 1792–1838.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance*, *45*(3), 881–898.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, *48*(1), 65–91.

Jiang, G. J., & Oomen, R. C. (2008). Testing for jumps when asset prices are observed with noise: A swap variance approach. *Journal of Econometrics*, *144*(2), 352–370.

Jiang, G. J., & Yao, T. (2013). Stock price jumps and cross-sectional return predictability. *Journal of Financial and Quantitative Analysis*, *48*(5), 1519–1544.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kelly, B., & Jiang, H. (2014). Editor's choice tail risk and asset prices. *Review of Financial Studies*, *27*(10), 2841–2871.

Kirman, A. (1993). Ants, rationality, and recruitment. *Quarterly Journal of Economics*, *108*(1), 137–156.

Lakonishok, J., Shleifer, A., Vishny, R. W., Hart, O., & Perry, G. L. (1992). The structure and performance of the money management industry. *Brookings Papers on Economic Activity. Microeconomics*, *1992*, 339–391.

LeBaron, B. (2006). Agent-based financial markets: Matching stylized facts with style. *Post Walrasian Macroeconomics: Beyond the DSGE Model*, *Edited by: Colander, DC*, 221–236. Cambridge: Cambridge University Press.

Lee, S. S., & Mykland, P. A. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Studies*, *21*(6), 2535–2563.

Lewellen, J., Nagel, S., & Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics*, *96*(2), 175–194.

Leybourne, S., Taylor, R., & Kim, T.-H. (2007). Cusum of squares-based tests for a change in persistence. *Journal of Time Series Analysis*, *28*(3), 408–433.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, *47*(1), 13–37.

Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, *3*(2), 175–205.

Lobato, I. N., & Savin, N. E. (1998). Real and spurious long-memory properties of stock-market data. *Journal of Business & Economic Statistics*, *16*(3), 261–268.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, *42*(3), 483–510.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, *34*(4), 768–783.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Pictet, O. V., Olsen, R. B., & Ward, J. R. (1993). Fractals and intrinsic time: A challenge to econometricians. *39th International AEA Conference on Real Time Econometrics, 14-15 October 1993, Luxembourg*.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics*, *78*(2), 277–309.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703–08.

Perez-Quiros, G., & Timmermann, A. (2000). Firm size and cyclical variations in stock returns. *Journal of Finance*, *55*(3), 1229–1262.

Peters, E. E. (1994). *Fractal market analysis: Applying chaos theory to investment and economics*, vol. 24. John Wiley, Hoboken, NJ.

Pukthuanthong, K., & Roll, R. (2015). Internationally correlated jumps. *Review of Asset Pricing Studies*, *5*(1), 92–111.

Robinson, P. M. (1995a). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, *23*(5), 1630–1661.

Robinson, P. M. (1995b). Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, *23*(3), 1048–1072.

Sadique, S., & Silvapulle, P. (2001). Long-term memory in stock market returns: International evidence. *International Journal of Finance & Economics*, *6*(1), 59–67.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, *19*(3), 425–442.

Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance*, *40*(3), 777–790.

Sibbertsen, P., & Kruse, R. (2009). Testing for a break in persistence under long-range dependencies. *Journal of Time Series Analysis*, *30*(3), 263–285.

Sibbertsen, P., Wegener, C., & Basse, T. (2014). Testing for a break in the persistence in yield spreads of EMU government bonds. *Journal of Banking & Finance*, *41*, 109–118.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39–60.

Whittle, P. (1951). *Hypothesis testing in time series analysis*. Almquist & Wiksells, Uppsala.

Zhang, L., Mykland, P. A., & Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, *100*(472), 1394–1411.

The Memory of Stock Return Volatility: Asset Pricing Implications

Online Appendix

Table 7: Sorted Portfolio Returns: Residual Long Memory

This table reports average returns and risk-adjusted returns of equally weighted quintile portfolios. Each month, stocks are sorted by their residual long memory and we track the portfolio returns over the subsequent month. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. Residual memory is calculated by regressing the memory parameter on size, book-to-market, momentum and illiquidity (Model 1). Model 2 additionally includes the BNS jump test statistic. The one-month-ahead portfolio returns are regressed on risk factors in the Capital Asset Pricing Model (CAPM) and the Fama & French (2015) five-factor model (FF5). The corresponding alphas are reported. We report Newey–West standard errors using lags equal to the return horizon in parentheses. Stars indicate significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (LMS)
Panel A: Model 1						
CAPM	0.0261***	0.0227**	0.0286***	0.0188*	0.0146	−0.0115*
	(0.0101)	(0.0100)	(0.0098)	(0.0101)	(0.0097)	(0.0060)
FF5	0.0049	−0.0007	0.0034	−0.0090*	−0.0100*	−0.0149**
	(0.0047)	(0.0044)	(0.0041)	(0.0048)	(0.0058)	(0.0069)
Panel B: Model 2						
CAPM	0.0261***	0.0236**	0.0293***	0.0176*	0.0141	−0.0120**
	(0.0100)	(0.0100)	(0.0099)	(0.0100)	(0.0097)	(0.0060)
FF5	0.0050	0.0006	0.0042	−0.0102**	−0.0099*	−0.0149**
	(0.0047)	(0.0043)	(0.0041)	(0.0048)	(0.0058)	(0.0068)

Table 8: Long Memory and Industries

This table reports descriptive statistics for the memory parameter of industry portfolios. for the period from 1926 until 2015. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. SD stands for the standard deviation. Min and Max stand for the minimum and maximum observation over the sample period.

	Non-Durables	Durables	Manufacturing	Energy	Chemicals	Business Equipment
Mean	0.21	0.22	0.22	0.21	0.24	0.19
Median	0.21	0.22	0.21	0.21	0.23	0.20
SD	0.06	0.05	0.06	0.08	0.10	0.08
Min	0.02	−0.02	0.11	−0.03	−0.04	−0.11
Max	0.37	0.39	0.44	0.55	0.80	0.56
Skewness	0.32	−0.06	1.64	0.34	1.08	−0.29
Kurtosis	3.48	4.22	6.33	4.00	6.13	4.22
	Telecommunication	Utilities	Shops	Healthcare	Money Finance	Other
Mean	0.20	0.21	0.23	0.23	0.21	0.21
Median	0.21	0.21	0.22	0.22	0.20	0.21
SD	0.09	0.08	0.05	0.08	0.07	0.07
Min	−0.30	−0.15	0.10	−0.01	−0.02	−0.05
Max	0.47	0.53	0.39	0.58	0.45	0.43
Skewness	−0.78	−0.29	0.82	1.08	0.03	−0.47
Kurtosis	5.77	5.39	3.53	5.39	3.92	5.10

Table 9: Long Memory and Fama–French Portfolios

This table reports the memory parameter for decile portfolios sorted Size, Book-to-Market and Momentum for the period from 1950 until 2015. The last column reports the average of the High-Minus-Low ($D10 - D1$) portfolio. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D10-D1
Size	0.3425	0.5483	0.5162	0.4799	0.4955	0.4489	0.4349	0.4397	0.4159	0.3860	0.0436
Book-to-Market	0.3382	0.4249	0.4334	0.4544	0.4808	0.5062	0.5090	0.5326	0.4905	0.6149	0.2767
Momentum	0.6184	0.6202	0.6138	0.5527	0.5215	0.4896	0.4237	0.3635	0.3034	0.1952	−0.4232

Table 10: Sorted Portfolio Returns: Alternative GPH Estimators

This table reports average returns and risk-adjusted returns of equally weighted quintile portfolios for the period from 1926 until 2015. Each month, stocks are sorted by their memory parameter estimate and we track the portfolio returns over the subsequent month. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.6}$, $m = N^{0.7}$ or $m = N^{0.8}$ in Panels A-C. The GPH estimator is applied to absolute returns and $m = N^{0.5}$ in Panel D. The one-month-ahead portfolio returns are regressed on risk factors in the Fama & French (2015) five-factor model (FF5). The average return and the corresponding alphas are reported. We report Newey–West standard errors using lags equal to the return horizon in parentheses. Stars indicate significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (LMS)
Panel A: GPH $m = N^{0.6}$						
Mean return	0.1347***	0.1353***	0.1316***	0.1248***	0.1167***	−0.0180**
	(0.0345)	(0.0347)	(0.0338)	(0.0345)	(0.0331)	(0.0089)
FF5	0.0219**	0.0174*	0.0081	0.0052	−0.0009	−0.0228**
	(0.0106)	(0.0090)	(0.0074)	(0.0076)	(0.0094)	(0.0095)
Panel B: GPH $m = N^{0.7}$						
Mean return	0.1426***	0.1313***	0.1286***	0.1256***	0.1155***	−0.0271***
	(0.0357)	(0.0345)	(0.0343)	(0.0331)	(0.0330)	(0.0096)
FF5	0.0291***	0.0131	0.0074	0.0070	−0.0043	−0.0334***
	(0.0105)	(0.0093)	(0.0078)	(0.0076)	(0.0088)	(0.0097)
Panel C: GPH $m = N^{0.8}$						
Mean return	0.1415***	0.1379***	0.1248***	0.1208***	0.1183***	−0.0232**
	(0.0361)	(0.0361)	(0.0335)	(0.0335)	(0.0313)	(0.0099)
FF5	0.0293***	0.0170**	0.0090	−0.0010	−0.0022	−0.0314***
	(0.0103)	(0.0082)	(0.0086)	(0.0085)	(0.0084)	(0.0095)
Panel D: GPH Absolute Returns $m = N^{0.5}$						
Mean return	0.1417***	0.1321***	0.1306***	0.1264***	0.1123***	−0.0294***
	(0.0335)	(0.0331)	(0.0341)	(0.0342)	(0.0360)	(0.0103)
FF5	0.0202**	0.0145	0.0106	0.0074	−0.0026	−0.0228**
	(0.0102)	(0.0091)	(0.0074)	(0.0074)	(0.0105)	(0.0105)

Table 11: Sorted Portfolio Returns: Alternative LW Estimators

This table reports average returns and risk-adjusted returns of equally weighted quintile portfolios for the period from 1926 until 2015. Each month, stocks are sorted by their memory parameter estimate and we track the portfolio returns over the subsequent month. The memory parameter is estimated with the LW estimator and a bandwidth parameter of $m = N^{0.5}$, $m = N^{0.6}$, $m = N^{0.7}$ or $m = N^{0.8}$ in Panels A-D. The LW estimator is applied to absolute returns and $m = N^{0.5}$ in Panel E. The one-month-ahead portfolio returns are regressed on risk factors in the Fama & French (2015) five-factor model (FF5). The average return and the corresponding alphas are reported. We report Newey–West standard errors using lags equal to the return horizon in parentheses. Stars indicate significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (LMS)
Panel A: LW $m = N^{0.5}$						
Mean return	0.1391***	0.1299***	0.1309***	0.1256***	0.1181***	−0.0209**
	(0.0333)	(0.0332)	(0.0333)	(0.0351)	(0.0358)	(0.0100)
FF5	0.0309***	0.0133	0.0092	0.0009	−0.0012	−0.0321***
	(0.0115)	(0.0086)	(0.0077)	(0.0076)	(0.0097)	(0.0109)
Panel B: LW $m = N^{0.6}$						
Mean return	0.1363***	0.1355***	0.1309***	0.1227***	0.1182***	−0.0182*
	(0.0342)	(0.0344)	(0.0341)	(0.0345)	(0.0334)	(0.0099)
FF5	0.0254**	0.0176**	0.0079	0.0017	0.0000	−0.0254**
	(0.0110)	(0.0088)	(0.0079)	(0.0075)	(0.0094)	(0.0103)
Panel C: LW $m = N^{0.7}$						
Mean return	0.1435***	0.1324***	0.1307***	0.1238***	0.1137***	−0.0298***
	(0.0352)	(0.0349)	(0.0338)	(0.0343)	(0.0326)	(0.0101)
FF5	0.0324***	0.0131	0.0093	0.0054	−0.0069	−0.0393***
	(0.0106)	(0.0092)	(0.0081)	(0.0079)	(0.0090)	(0.0105)
Panel D: LW $m = N^{0.8}$						
Mean return	0.1427***	0.1370***	0.1275***	0.1230***	0.1135***	−0.0292***
	(0.0366)	(0.0351)	(0.0344)	(0.0334)	(0.0315)	(0.0112)
FF5	0.0298***	0.0191**	0.0080	0.0014	−0.0053	−0.0351***
	(0.0108)	(0.0088)	(0.0082)	(0.0078)	(0.0093)	(0.0106)
Panel E: LW Absolute Returns $m = N^{0.5}$						
Mean return	0.1445***	0.1327***	0.1336***	0.1175***	0.1141***	−0.0303**
	(0.0337)	(0.0324)	(0.0344)	(0.0337)	(0.0369)	(0.0121)
FF5	0.0264**	0.0147	0.0099	0.0021	−0.0029	−0.0293***
	(0.0103)	(0.0091)	(0.0075)	(0.0076)	(0.0108)	(0.0112)

Table 12: Cross-sectional Regressions: Alternative Estimators

This table reports results from Fama–MacBeth regressions for the period from 1950 until 2015. Each month, excess stock returns are regressed on the lagged memory parameters in Panel A. Panel B further includes additional lagged firm characteristics, which are market capitalization (Size), book-to-market values, prior returns (Momentum), illiquidity and jump statistics (BNS). The memory parameter is estimated by applying the GPH or the LW estimator and a bandwidth parameter of $m = N^{0.5}$, $m = N^{0.6}$, $m = N^{0.7}$ or $m = N^{0.8}$ to squared or absolute returns. Stars indicate the significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	GPH				LW				
	$N^{0.6}$	$N^{0.7}$	$N^{0.8}$	Abs. $N^{0.5}$	$N^{0.5}$	$N^{0.6}$	$N^{0.7}$	$N^{0.8}$	Abs. $N^{0.5}$
Panel A: Simple Regressions									
Intercept	0.0089***	0.0091***	0.0091***	0.0094***	0.0092***	0.0091***	0.0094***	0.0095***	0.0097***
	(0.0021)	(0.0022)	(0.0022)	(0.0021)	(0.0021)	(0.0022)	(0.0022)	(0.0022)	(0.0021)
Long Memory	−0.0039**	−0.0062**	−0.0075**	−0.0045***	−0.0047**	−0.0053**	−0.0084**	−0.0104**	−0.0060***
	(0.0019)	(0.0025)	(0.0031)	(0.0017)	(0.0019)	(0.0026)	(0.0034)	(0.0043)	(0.0022)
Panel B: Multiple Regressions									
Intercept	0.0106**	0.0109**	0.0110**	0.0111**	0.0106**	0.0106**	0.0109**	0.0111**	0.0110**
	(0.0045)	(0.0045)	(0.0046)	(0.0045)	(0.0045)	(0.0045)	(0.0046)	(0.0046)	(0.0045)
Long Memory	−0.0026*	−0.0044**	−0.0047*	−0.0027**	−0.0030**	−0.0033*	−0.0063**	−0.0071**	−0.0036**
	(0.0014)	(0.0019)	(0.0024)	(0.0011)	(0.0013)	(0.0018)	(0.0025)	(0.0033)	(0.0015)
Size	−0.0005*	−0.0005*	−0.0005**	−0.0005**	−0.0005*	−0.0005*	−0.0005*	−0.0005*	−0.0005*
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Book-to-Market	0.0024***	0.0024***	0.0023***	0.0024***	0.0024***	0.0024***	0.0024***	0.0023***	0.0024***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Momentum	0.0094***	0.0095***	0.0094***	0.0095***	0.0095***	0.0095***	0.0094***	0.0094***	0.0095***
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Illiquidity	0.0987	0.0952	0.0926	0.0927	0.0993	0.0991	0.0971	0.0924	0.0933
	(0.1730)	(0.1720)	(0.1715)	(0.1725)	(0.1726)	(0.1730)	(0.1718)	(0.1709)	(0.1724)
BNS	0.0020***	0.0020***	0.0020***	0.0020***	0.0020***	0.0020***	0.0020***	0.0020***	0.0020***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)

Table 13: Sorted Portfolio Returns: Alternative Holding Periods

This table reports average returns and risk-adjusted returns of equally weighted quintile portfolios for the period from 1926 until 2015. Each month, stocks are sorted by their memory parameter estimate and we track the portfolio returns over the subsequent one, two, three, four and five years in Panel A, B, C, D and E, respectively. The memory parameter is estimated with the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. The one-month-ahead portfolio returns are regressed on risk factors in the Capital Asset Pricing Model (CAPM) and the Fama & French (2015) five-factor model (period starts in 1963) (FF5). The mean returns and the corresponding alphas are reported. We report Newey–West standard errors using lags equal to the return horizon in parentheses. Stars indicate significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (LMS)
Panel A: One Year Holding Period						
Mean return	0.1404***	0.1371***	0.1384***	0.1329***	0.1216***	−0.0188**
	(0.0503)	(0.0504)	(0.0472)	(0.0438)	(0.0448)	(0.0095)
FF5	0.1566***	0.1592***	0.1616***	0.1478***	0.1400***	−0.0166
	(0.0502)	(0.0488)	(0.0477)	(0.0470)	(0.0473)	(0.0150)
Panel B: Two Years Holding Period						
Mean return	0.1453***	0.1431***	0.1412***	0.1371***	0.1260***	−0.0193**
	(0.0423)	(0.0403)	(0.0386)	(0.0359)	(0.0371)	(0.0097)
FF5	−0.0059	−0.0037	−0.0076	−0.0200**	−0.0288***	−0.0229**
	(0.0092)	(0.0079)	(0.0075)	(0.0087)	(0.0111)	(0.0100)
Panel C: Three Years Holding Period						
Mean return	0.1445***	0.1441***	0.1421***	0.1378***	0.1256***	−0.0188**
	(0.0390)	(0.0391)	(0.0357)	(0.0325)	(0.0332)	(0.0092)
FF5	−0.0072	−0.0025	−0.0093	−0.0202*	−0.0292**	−0.0219**
	(0.0089)	(0.0092)	(0.0078)	(0.0115)	(0.0119)	(0.0091)
Panel D: Four Years Holding Period						
Mean return	0.1510***	0.1509***	0.1478***	0.1438***	0.1319***	−0.0190**
	(0.0424)	(0.0434)	(0.0394)	(0.0352)	(0.0352)	(0.0094)
FF5	−0.0008	0.0017	−0.0031	−0.0152	−0.0212	−0.0204**
	(0.0188)	(0.0234)	(0.0292)	(0.0302)	(0.0224)	(0.0081)
Panel E: Five Years Holding Period						
Mean return	0.1534***	0.1537***	0.1493***	0.1464***	0.1343***	−0.0191**
	(0.0341)	(0.0379)	(0.0356)	(0.0321)	(0.0303)	(0.0093)
FF5	−0.0191	−0.0204	−0.0263*	−0.0352***	−0.0393***	−0.0203**
	(0.0141)	(0.0180)	(0.0142)	(0.0124)	(0.0122)	(0.0091)

Table 14: Sorted Portfolio Returns: High Frequency Data

This table reports average returns and risk-adjusted returns of quintile portfolios for the period from 1996 until 2015. Each month, stocks are sorted by their long memory parameter estimate and we track the portfolio returns over the subsequent month. The one-month-ahead portfolio returns are regressed on risk factors in the Capital Asset Pricing Model (CAPM), the Fama & French (1993) 3-factor model (FF3), the Fama & French (2015) 5-factor model (FF5) and the Hou et al. (2014) q-model (HXZ). The corresponding alphas are reported. We report Newey–West standard errors using lags equal to the return horizon in parentheses. The memory parameter is estimated using a month of 5-min returns and the GPH estimator and a bandwidth parameter of $m = N^{0.5}$. Stars indicate significance: * significant at $p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Q1	Q2	Q3	Q4	Q5	Q5-Q1 (LMS)
Mean return	0.1638***	0.1246***	0.1091**	0.1025**	0.0754*	−0.0883***
	(0.0440)	(0.0472)	(0.0459)	(0.0437)	(0.0444)	(0.0176)
CAPM	0.1508***	0.1138***	0.0980**	0.0881**	0.0600	−0.0908***
	(0.0403)	(0.0431)	(0.0441)	(0.0404)	(0.0417)	(0.0182)
FF3	0.1369***	0.1040**	0.0798*	0.0768*	0.0633	−0.0736***
	(0.0423)	(0.0431)	(0.0452)	(0.0428)	(0.0432)	(0.0193)
FF5	0.0963	0.0643	0.0395	0.0558	0.0367	−0.0597**
	(0.0628)	(0.0625)	(0.0601)	(0.0569)	(0.0555)	(0.0258)
HXZ	0.0558	0.0194	−0.0088	0.0259	−0.0169	−0.0727**
	(0.0850)	(0.0823)	(0.0810)	(0.0770)	(0.0727)	(0.0302)

Table 15: Fama–MacBeth Regressions: Additional Control Variables

This table reports results from Fama–MacBeth regressions for the period from 1950 until 2015. Each month, excess stock returns are regressed on lagged firm characteristics including, memory parameters, market capitalization (Size), book-to-market values, prior returns (Momentum), illiquidity and jump statistics (BNS). We further control for Beta, Cokurtosis (CKT), Coskewness (CSK), idiosyncratic volatility (IVOL), kurtosis (KURT), skewness (SKEW), demand for lottery (MAX) and volatility of volatility (Vol-of-Vol). We report Newey–West standard errors using lags equal to the return horizon in parentheses. Stars indicate significance: * significant at $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15	Model 16	Model 17
Intercept	0.0091***	0.0144***	0.0075***	0.0076***	0.0087***	0.0091***	0.0096***	0.0095***	0.0087***	0.0089***	0.0116***	0.0114***	0.0096***	0.0122***	0.0089***	0.0106***	0.0267***
	(0.0025)	(0.0051)	(0.0025)	(0.0025)	(0.0025)	(0.0025)	(0.0027)	(0.0024)	(0.0026)	(0.0026)	(0.0019)	(0.0025)	(0.0025)	(0.0021)	(0.0017)	(0.0019)	(0.0040)
Long Memory	−0.0039**	−0.0021*	−0.0038**	−0.0038**	−0.0044***	−0.0043***	−0.0040**	−0.0037**	−0.0036**	−0.0037**	−0.0042***	−0.0043***	−0.0039**	−0.0041***	−0.0038***	−0.0041***	−0.0021**
	(0.0016)	(0.0012)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0017)	(0.0016)	(0.0016)	(0.0016)	(0.0014)	(0.0017)	(0.0016)	(0.0015)	(0.0014)	(0.0013)	(0.0010)
Size		−0.0006*															−0.0013***
		(0.0003)															(0.0002)
Book-to-Market			0.0019***														0.0012**
			(0.0005)														(0.0005)
Momentum				0.0067***													0.0096***
				(0.0016)													(0.0013)
Illiquidity					0.2010**												0.4686**
					(0.1010)												(0.1836)
BNS						0.0024***											−0.0001
						(0.0004)											(0.0003)
Beta							−0.0510***									−0.0006	−0.0006
							(0.0042)									(0.0005)	(0.0006)
REV								−0.0005								−0.0577***	−0.0395***
								(0.0005)								(0.0043)	(0.0037)
CKT									0.0000							0.0007	0.0011***
									(0.0005)							(0.0005)	(0.0004)
CSK										−0.0005						0.0003	−0.0010
										(0.0008)						(0.0007)	(0.0007)
IVOL											−0.1483***					−0.0725	−0.2892***
											(0.0447)					(0.0598)	(0.0535)
KURT												−0.0006***				−0.0003**	−0.0002**
												(0.0001)				(0.0001)	(0.0001)
SKEW													−0.0020***			0.0014***	0.0009***
													(0.0003)			(0.0003)	(0.0003)
MAX														−0.0610***		−0.0159	0.0156
														(0.0112)		(0.0164)	(0.0165)
Vol-of-Vol															−0.0074	0.0207	−0.0206
															(0.0309)	(0.0244)	(0.0244)