

A New Look at Expected Stock Returns and Volatility

Russell P. Robins¹ and Geoffrey Peter Smith*²

¹AB Freeman School of Business, Tulane University

²WP Carey School of Business, Arizona State University

Abstract

We replicate French, Schwert, and Stambaugh (1987) (FSS) with up-to-date data and new tools from the modern toolbox of econometric methods. As we proceed, we highlight the main technical details and econometric methods from the original study and, when necessary, update them. While our main goal is to replicate FSS as carefully as possible, we also aim to help new researchers quickly gain an in-depth understanding of the major features of the original study, and to demonstrate why FSS is fundamental to the asset pricing literature. We finish by text mining the titles and abstracts of over one thousand citing studies for information on why other studies cite FSS and which parts of FSS receive the most attention. After careful replication, we confirm that the main results in FSS hold and continue to hold through 2019.

1 Introduction

Lying at the heart of the asset pricing literature, French, Schwert, and Stambaugh (1987) (hereafter FSS) pose a question central to the discipline; what is the relation between the expected market risk premium, defined as the expected return on a stock market portfolio minus the risk-free interest rate, and risk, as measured by the volatility of the stock market? Thirty-three years later, we replicate this study with up-to-date data and

**Corresponding author:* Department of Finance, WP Carey School of Business, Arizona State University, Tempe, AZ 85287 USA; Phone: (480) 965-8623; Fax: (480) 965-8539; E-mail: gps@asu.edu.

new tools from the modern toolbox of econometric methods. As we proceed, we highlight the main technical details and econometric methods from the original study and, when necessary, update them. While our main goal is to replicate FSS as carefully as possible, we also aim to help new researchers quickly gain an in-depth understanding of the major features of the original study, and to demonstrate why FSS is fundamental to the asset pricing literature.

At its core, FSS is an empirical study on the intertemporal relation between risk and expected stock returns. More specifically, FSS investigate relations of the form

$$E(R_{mt} - R_{ft} | \hat{\sigma}_{mt}) = \alpha + \beta \hat{\sigma}_{mt}^p, \quad p = 1, 2, \quad (1)$$

where R_{mt} is the return on a stock market portfolio, R_{ft} is the risk-free interest rate, $\hat{\sigma}_{mt}$ is an ex ante measure of the portfolio's standard deviation, and $\hat{\sigma}_{mt}^2$ is an ex ante measure of the variance. If $\beta = 0$ in (1), the expected risk premium is unrelated to the ex ante volatility. If $\alpha = 0$ and $\beta > 0$, the expected risk premium is proportional to the standard deviation ($p = 1$) or variance ($p = 2$) of stock market returns.

Knowledge of FSS is now essential for understanding the market-wide risk-return relation. Yet, at the time it was published, there were already a number of closely-related studies on the topic of risk and return.¹ Before FSS, Pindyck (1984) argued that the relation between expected returns and volatility is strong, while Poterba and Summers (1986) argued that the time-series properties of volatility made this unlikely. Before that, Merton (1980) estimated a risk-return relation very similar to (1) with contemporaneous measures of volatility.

What, then, is the critical insight that makes FSS different from these earlier studies? We suggest it is the recognition that realized market volatility is in part predicted and

¹See, for example, Black, Jensen, and Scholes (1972), Fama and MacBeth (1973), Roll and Ross (1980), Breeden, Gibbons, and Litzenberger (1989), and Chen, Roll, and Ross (1986).

in part unpredicted by the market. Predicted volatility is the expected level of volatility based on past observations of realized volatility. Unpredicted volatility is the difference between the level of volatility predicted and the amount of volatility realized. French et al. (1987) show how combining predicted and unpredicted volatility into a single measure of realized volatility obscures the positive relation between the expected risk premium and predicted volatility. Having the insight to decouple predicted volatility from realized volatility enables FSS to discover the *volatility feedback effect*, where a volatility surprise causes the market to reevaluate its expectation of future risk premiums and make an immediate adjustment in stock prices.²

After careful replication, we confirm that the main results in FSS hold and continue to hold through 2019. We confirm that there is a positive relation between the expected risk premium on common stocks and the level of predicted volatility, though the variability of stock returns is so large that it is difficult to discriminate among alternative specifications of this relation. We also confirm that there is a negative and significant relation between the unpredicted level of volatility and excess holding period returns. In sum, we confirm that there is a market-wide volatility feedback effect through 2019. We also confirm that the magnitude of this effect is too large to be attributed to the leverage effect discussed in Black (1976) and Christie (1982).

2 Why is FSS Fundamental to Asset Pricing?

According to the Web of Science database, FSS is cited in 1,189 studies published through June 2020. In Appendix A, we discuss the results from text mining the titles and abstracts of these studies for information on why other studies cite FSS and which parts of FSS receive the most attention. In this section, we discuss why FSS is fundamental to the asset

²Pindyck (1984), French et al. (1987), and Campbell and Hentschel (1992) are often cited together as the foundational studies on the volatility feedback effect.

pricing literature.

The focus of FSS is market volatility and market volatility is a common measure of risk. French et al. (1987) show how higher (lower) than expected volatility drives up (down) expected risk premiums, leading to an immediate decrease (increase) in stock market value. French et al. (1987) establish this relation and show how it represents a positive relation between conditional volatility and expected returns through the discount rate channel. If stock values drop in response to a volatility spike, then discount rates must have risen, which are identical to expected returns. Most (if not all) of the major multiperiod asset pricing models predict such a relation.

For example, the risk-return relation established in FSS underlies the intertemporal capital asset pricing model (ICAPM) of Merton (1973), which recognizes investors' desire to hedge against future changes in the investment opportunity set. This desire induces an additional systematic risk premium beyond the usual CAPM beta. Starting from the ICAPM, other major asset pricing models establish links between stock returns and key macroeconomic variables thought to represent uncertainty over future investment opportunities and expected returns. Prominent among these is the growth rate in consumption per capita in Campbell and Cochrane (1999) and Bansal and Yaron (2004). Other key macroeconomic variables include default spreads and term spreads in Fama and French (1989), short-term-interest rates in Shanken (1990), lagged production growth, interest rates, and the market dividend-price ratio in Chen (1991), and changes in the federal funds rate in Bernanke and Kuttner (2005). For international stock markets, Ferson and Harvey (1993) add foreign currency returns, global inflation, world interest rates, and world industrial production. In sum, the evidence on the market-wide risk-return relation established in FSS substantiates the predictions of most of the major multiperiod asset pricing models.

With this in mind, we now turn our attention onto the replication itself. As we pro-

ceed, we highlight the main technical details and econometric methods from the original study and, when necessary, update them. For those who would like to see how we make our calculations, we give major segments of the R code used for the replication in Appendix B.³

3 Time Series Properties of the Data

French et al. (1987) recognize that realized market volatility is in part predicted and in part unpredicted by the market. Predicted volatility is the expected level of volatility based on past observations of realized volatility. Unpredicted volatility is the difference between the level of volatility predicted and the amount of volatility realized. It follows that a major focus of FSS is making correct estimates for realized and predicted market volatility.

3.1 Standard Deviations of Stock Market Returns

French et al. (1987) calculate realized monthly market volatility from the daily returns on the S&P composite portfolio from 1928–84. We extend their sample period to 2019 for this calculation and for the remainder of the replication.⁴

Equation (2) in FSS is the function for calculating realized monthly market volatility (variance) from daily stock returns

$$\sigma_{mt}^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2 \sum_{i=1}^{N_t-1} r_{it} r_{i+1,t}, \quad (2)$$

where there are N_t daily returns, r_{it} , in month t . French et al. (1987) assume that the mean

³All of our data analysis is conducted using R software by the R Core Team (2020).

⁴Daily returns on the S&P composite portfolio from January 3, 1928 to July 2, 1962 are from Schwert (1990) and from the Center for Research in Security Prices (CRSP) thereafter.

daily return is zero and that the returns are autocorrelated for one lag due to nonsynchronous trading.⁵ Equation (2) follows from the *variance sum law*, where the sum of two correlated random variables, σ_{x+y}^2 , is equal to $\sigma_x^2 + \sigma_y^2 + 2\sigma_{xy}$, where σ_{xy} is the covariance.⁶ The calculation of monthly variance in (2) is critical because σ_{mt}^2 and the monthly standard deviation, σ_{mt} , (more generally, σ_{mt}^p) are the only estimates for realized monthly market volatility in the remainder of the study.

Here we call attention to two concerns with (2). First, Schwert (1989b) points out that the variance estimates from (2) are not guaranteed to be positive. Negative estimates are nonsensible because variance is the expected squared deviation around the mean and therefore always nonnegative. Though none of the monthly variances are negative, σ_{mt}^2 would have been negative in September 1906, which is outside the original sample period.⁷ The second concern is that monthly volatility is calculated as a function of i , the number of trading days in the month. Some months have more trading days than others. This means that months having more trading days might automatically have higher volatility than those with fewer trading days because (2) has more terms in some months than in others. A simple correction is to calculate monthly volatility per number of trading days in the month.

We calculate (2) and then replicate FSS Figure 1a by plotting the monthly percent realized standard deviation estimates in Figure 1. We also add plots of the autocorrelation function (ACF), partial autocorrelation function (PACF), and a histogram with an overlay of the normal distribution.

[Figure 1 here]

The top of Figure 1 is a plot of the time-series of monthly percent realized standard

⁵See, for example, Fisher (1966) and Scholes and Williams (1977).

⁶To see how (2) works, assume $i = \{1, 2\}$, so the variance on day $i = 1$ is $r_{1t}^2 + 2r_{1t}r_{2t}$ and the variance on day $i = 2$ is r_{2t}^2 . The sum of these is $r_{1t}^2 + r_{2t}^2 + 2r_{1t}r_{2t}$, which follows the variance sum law exactly.

⁷We discover the negative variance in September 1906 by calculating (2) using all of the available Schwert (1990) data, which begins on February 17, 1885.

deviations from 1928–2019. This part of Figure 1 highlights how much realized volatility changes from month to month. There are some periods of high and others of low volatility, meaning that stock returns are heteroskedastic. Prior evidence of heteroskedastic stock returns is also found in Officer (1973). To isolate stable periods, FSS divide the full sample of monthly realized volatility into two subperiods, 1928–52 and 1953–84. We add 1985–2019.

We replicate FSS Table 1 and report descriptive statistics for the realized standard deviations in Table 1, Panel A.

[Table 1 here]

We confirm that the mean and standard deviation of σ_{mt} are higher in 1928–52 than in 1953–84. Our new subperiod, 1985–2019, is in between. The autocorrelations of σ_{mt} are large and decay slowly in each subperiod. This feature can also be seen at the bottom of Figure 1, where we plot the ACF and PACF for 1928–2019. Table 1, Panel A, shows that σ_{mt} has significant positive skewness. This feature can also be seen at the bottom of Figure 1, where we plot the histogram for σ_{mt} with an overlay of the normal distribution. Under the hypothesis of a stationary normal distribution, $\sqrt{6T(T-1)/(T-2)(T+1)(T+3)}$ is the exact asymptotic standard error for the sample skewness.⁸ French et al. (1987) calculate the approximate asymptotic standard error for the sample skewness as $\sqrt{6/T}$.

With realized volatility in hand, FSS estimate predicted volatility, $\hat{\sigma}_{mt}^p$. Predicted volatility is unobservable, so it must be estimated from realized volatility. French et al. (1987) calculate predicted volatility in two ways. The first follows the approach of Box and Jenkins (1970b). Box-Jenkins forecasting involves estimating an autoregressive integrated moving average (ARIMA) model to find the best fit of a time series to past values of itself. Box-Jenkins forecasting generally involves five steps: (i) data preparation, (ii) model

⁸The formula for calculating the standard error of the sample skewness is from Weinberg, Harel, and Abramowitz (2021).

selection, (iii) parameter estimation, (iv) model checking, and (v) prediction.

Figure 1 shows that some months are outliers in the sense that the σ_{mt} in that month is much more extreme than in others. To deemphasize these outliers, FSS calculate the natural logarithm of monthly realized standard deviation as $\ln \sigma_{mt}$. Table 1, Panel B, reports descriptive statistics for $\ln \sigma_{mt}$. Notice that the skewness is gone and that the autocorrelations are close to zero beyond lag three.

We add here that another way to deemphasize outliers is with the Box-Cox transformation of Box and Cox (1964). Box-Cox transforms nonnegative nonnormal data into something more normal through

$$f_{\lambda}(x) = \begin{cases} \frac{x^{\lambda}-1}{\lambda} & \text{if } \lambda \neq 0; \\ \ln(x) & \text{if } \lambda = 0, \end{cases} \quad (3)$$

where values from -1 to $+2$ are considered for λ and the final λ is the one that results in the best approximation of the normal distribution. The Box-Cox transformation equals the log transformation when $\lambda = 0$.

We calculate λ as -0.42 from 1928–84, -0.79 from 1928–52, -0.13 from 1953–84, and -0.57 from 1985–2019, though there is little practical difference between the transformed volatility estimates. The correlation between the log volatility and the Box-Cox volatility ranges from 0.9349 from 1928–52 to 0.9991 from 1953–84.

It is also true that, for regression models in general, the dependent variable does not need to be normal anyway. All that matters are the properties of the regression estimates, and the central limit theorem guarantees that these estimates will be normal (or essentially equivalent; follow a t -distribution) as the sample size becomes larger. However, the properties of time-series regression coefficient estimates are not normal when the variable in the regression is nonstationary, meaning that the mean, variance, and autocorrelations

are not constant through time. The central limit theorem breaks down when the variable is nonstationary and the coefficient estimates will follow a nonstandard distribution. It is preferable to work with stationary data so that the coefficient estimates follow a distribution we understand and with which we are more familiar.

French et al. (1987) assess the stationarity of $\ln \sigma_{mt}$ by calculating the Q -statistic of Box and Pierce (1970a) as a test for the absence of autocorrelation up to 12 lags. We calculate the Ljung and Box (1978) Q -statistic instead because it is a modified version of the Box et al. (1970a) Q -statistic with increased power in finite samples. The null hypothesis of both tests is no autocorrelation. We confirm that the Q -statistics are significant in every subperiod, and the null of no autocorrelation is rejected. To achieve stationarity, FSS take first differences in $\ln \sigma_{mt}$.

We add here that another way to assess stationarity is with the augmented Dickey-Fuller (ADF) test of Dickey and Fuller (1979) and Said and Dickey (1984). The null hypothesis of the ADF test is the presence of a unit root, while the alternative is stationarity. We calculate the ADF test for the first difference in $\ln \sigma_{mt}$ for 12 lags. We reject the null hypothesis and find that the first difference in $\ln \sigma_{mt}$ is stationary in every subperiod.⁹

Figure 2 is a re-plot of Figure 1 with the first difference in $\ln \sigma_{mt}$ in place of σ_{mt} . Table 1 and Figure 2 show the log transformation and first-differencing work as intended. The autocorrelations are lower and the data are much more normal looking.

[Figure 2 here]

With the data preparation step complete, step two in Box-Jenkins is model selection. French et al. (1987) select an ARIMA regression model of the form

$$(1 - L) \ln \sigma_{mt} = \theta_0 + (1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3) u_t, \quad (4)$$

⁹Using subjective judgment as to the need for a drift or time trend can enhance the power of unit root tests. See, for example, Kennedy (2003, p. 403). We include a drift term in our ADF tests.

which is an ARIMA(0, 1, 3) model with zero autoregressive (AR) terms, first-differencing of the dependent variable (I), three moving average (MA) terms, and a drift term. This is known as a short-memory model because $\ln \sigma_{mt}$ is completely uncorrelated with lags of itself beyond lag three.

We add here that our ACF and PACF plots are useful for identifying the number of AR or MA terms for an ARIMA($p, d, 0$) or an ARIMA($0, d, q$) model. The data may follow an ARIMA($p, d, 0$) model if the ACF is exponentially decaying or sinusoidal and there is a significant spike at lag p in the PACF, but none beyond lag p . The data may follow an ARIMA($0, d, q$) model if the PACF is exponentially decaying or sinusoidal and there is a significant spike at lag q in the ACF, but none beyond lag q . Looking at the ACF and PACF plots in Figure 2, it seems that the ARIMA(0, 1, 3) model is a good choice. The PACF is decaying and there are no significant spikes in the ACF beyond lag three but for lag 12.

Step three in Box-Jenkins is parameter estimation. Table 1, Panel C, summarizes the results from regression (4) in each subperiod.¹⁰ We confirm that the estimates for the constant term, θ_0 , are small compared to their standard errors, suggesting no deterministic drift in the standard deviation of the stock market return. The MA estimate at lag one is large in all periods, whereas the estimate at lag two is largest from 1928–52, and the estimate at lag three is largest from 1953–84. The F -statistic of 0.99 testing the hypothesis that the parameters are the same in 1928–52 and 1953–84 is not significant. We do not adjust this calculation for unequal variances.

With the parameter estimates in hand, step four in Box-Jenkins is model checking. This involves confirming that the model errors, u_t , are uncorrelated. French et al. (1987) calculate the Q -statistic of Box et al. (1970a) for 12 lags of u_t to check for the absence of autocorrelation.

¹⁰We estimate the ARIMA models using the forecast package of Hyndman, Athanasopoulos, Bergmeir, Caceres, Chhay, O'Hara-Wild, Petropoulos, Razbash, Wang, and Yasmeeen (2020).

We add here that the Q -statistic follows the chi-squared distribution with h degrees of freedom, where h is the number of lags being tested. When the Ljung et al. (1978) test is applied to the residuals of an ARIMA model, though, the degrees of freedom are h minus the number of parameters in the ARIMA model. If we use the critical value for $h = 12$, which is 21.03, we confirm that the residuals from (4) are not significantly autocorrelated. If we use the critical value for $h = 8$, which is 15.51, then the Ljung et al. (1978) test rejects the null hypothesis that the residuals are not autocorrelated in the 1928–52 subperiod.

Step five in Box-Jenkins is prediction. French et al. (1987) calculate predicted volatility, $\ln \hat{\sigma}_{mt}$, as the fitted values from regression (4). The fitted values estimated by an ARIMA model are predictions calculated using only the prior realized values of the dependent variable. It follows that the predicted volatility from regression (4) is not subject to look-ahead bias because, conditional on the parameters, which are treated as if they are known to investors, it is dependent only on the past realized values of monthly volatility. There is one major caveat, though. In order to be useful, the fitted values, $\ln \hat{\sigma}_{mt}$, must be back-transformed in terms of the original data, $\hat{\sigma}_{mt}$ (without logs). It is well-known that if $\ln \hat{\sigma}_{mt}$ is normally distributed with mean μ and variance u^2 , then $\hat{\sigma}_{mt}$ is log-normally distributed with mean $\exp(\mu + u^2/2)$. Equations (4a) and (4b) in FSS give the exact back-transformations as

$$\hat{\sigma}_{mt} = \exp[\ln \hat{\sigma}_{mt} + 0.5V(u_t)] \quad (5)$$

$$\hat{\sigma}_{mt}^2 = \exp[2 \ln \hat{\sigma}_{mt} + 2V(u_t)], \quad (6)$$

where $\ln \hat{\sigma}_{mt}$ are the fitted values and $V(u_t)$ is the variance of the residuals from regression (4). We calculate predicted monthly standard deviation, $\hat{\sigma}_{mt}$, and predicted monthly variance, $\hat{\sigma}_{mt}^2$, following (5) and (6).

We replicate FSS Figure 1b and plot in Figure 3 the predicted monthly standard devi-

ations from (5). To make the comparison easier, we plot the realized monthly standard deviations from (1) above the predicted monthly standard deviations from (5).

[Figure 3 here]

Predicted monthly standard deviation, $\hat{\sigma}_{mt}$, tracks realized monthly standard deviation, σ_{mt} , closely, although the predicted series is smoother. We calculate R^2 as the squared correlation between $\hat{\sigma}_{mt}$ and σ_{mt} . The R^2 values are 0.54 from 1928–84, 0.49 from 1928–52, 0.41 from 1953–84, and 0.28 from 1985–2019.

The back-transformed fitted values from (5) and (6), $\hat{\sigma}_{mt}^p$, are the first estimates of predicted volatility in FSS. Unpredicted volatility, σ_{mt}^{pu} , then equals, $\sigma_{mt}^p - \hat{\sigma}_{mt}^p$, the difference between realized and predicted volatility.¹¹

3.2 ARCH Models

One of the major assumptions of volatility feedback is that conditional volatility is autocorrelated and predictable. To help establish predictability, FSS fit the autoregressive conditional heteroskedasticity (ARCH) model of Engle (1982)

$$r_t = \alpha + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_t^2), \quad (7)$$

$$\sigma_t^2 = a + b\varepsilon_{t-1}^2, \quad (8)$$

to represent a series with changing volatility. Unlike ARIMA, ARCH models the return, r_t , and variance, σ_t^2 , processes jointly in two equations. The mean equation (7) describes the return as a function of other variables plus an error term. The variance equation (8) describes the evolution of the conditional variance of the error from the mean equation as a function of past lagged errors.

¹¹Amihud (2002, 2019) use the same predicted/unpredicted approach to analyze the relation between expected/unexpected stock liquidity and future stock returns.

French et al. (1987) model the conditional excess return process as an MA(1) model

$$R_{mt} - R_{ft} = \alpha + \varepsilon_t - \theta\varepsilon_{t-1}, \quad (9)$$

to account for nonsynchronous trading. The term $R_{mt} - R_{ft}$ is the percentage change in the S&P composite portfolio minus the risk-free interest rate on Treasury bills from CRSP.

French et al. (1987) model the conditional variance process using the GARCH (generalized ARCH) model of Bollerslev (1986), where

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2 \quad (10)$$

is the GARCH variance equation. The GARCH variance equation describes the evolution of the conditional variance of the error from the mean equation as a function of lagged errors plus past values of the conditional variance. Equation (10) is a GARCH(2, 1) model with two lags of the error from the mean equation and one lagged value of the conditional variance.

We replicate FSS Table 2 and report the estimates for regression (9) and (10) with daily risk premiums in Table 2.¹²

[Table 2 here]

The GARCH estimates indicate that the variance of the daily risk premiums is highly autocorrelated. The sum $(b + c_1 + c_2)$ must be less than 1.0 for the volatility process to be stationary. This sum equals 0.997, 0.997, 0.993, and 0.984, for the 1928–84, 1928–52, 1953–84, and 1985–2019 subperiods, respectively. French et al. (1987) calculate a chi-square test to check the constancy of the GARCH parameters.

We add here that we now know from Robins and Smith (2020) that the chi-square test

¹²We estimate the GARCH models using the `rugarch` package of Ghalanos (2020).

for constancy is not valid unless the breakpoint dividing the subperiods is exogenous. Robins et al. (2020) also show how the chi-square test is very sensitive to the breakpoint, where the results can change from significant to insignificant simply by moving the breakpoint backward or forward by just one period. In light of this, we check the constancy of the GARCH parameters with the Nyblom (1989) test, which is a correct test for the constancy of GARCH parameters. We confirm that the GARCH parameters are not constant from 1928–84.

3.3 *Stock Market Risk Premiums*

We follow FSS and use the value-weighted portfolio of all New York Stock Exchange (NYSE) stocks from CRSP to measure monthly stock market returns. We calculate monthly excess holding period returns by subtracting the monthly risk-free interest rate on Treasury bills from CRSP. The mean excess holding period return is then an estimate of the mean expected risk premium.

We replicate FSS Table 3 and report the mean, standard deviation, and skewness of the monthly excess holding period returns in Table 3.

[Table 3 here]

We estimate the mean three ways: (i) ordinary least squares (OLS); (ii) weighted least squares (WLS), where the weight for each observation is $1/\sigma_{mt}$, the reciprocal of the realized monthly standard deviation estimated from the daily S&P returns; and (iii) WLS, where the weight is $1/\hat{\sigma}_{mt}$, the reciprocal of the predicted standard deviation from the ARIMA model in Table 1.

The difference between OLS and WLS is that WLS minimizes the weighted sum of squares instead of the residual sum of squares. An estimate by OLS is only efficient in terms of having the smallest mean square error if the data have a constant variance. The

purpose of WLS is to improve the efficiency of the estimate by accounting for nonconstant variance (heteroskedasticity). Weighted least squares deemphasizes values with a high standard deviation to produce more efficient estimates.

The means weighted by realized standard deviation give larger estimates of the expected risk premium and larger t -statistics than either of the other estimates. This foreshadows a result in the following section. In periods of unexpectedly high volatility (meaning when $\sigma_{mt} > \hat{\sigma}_{mt}$), realized stock returns are lower than average. These lower returns receive less weight when $1/\sigma_{mt}$ is the weight used to estimate the expected risk premium.

We add here that Moreira and Muir (2017, 2019) produce a novel result by exploiting this feature of stock returns. They find that the standard mean-variance trade-off weakens in periods of high volatility. This suggests that investors should time volatility by taking more risk when volatility is low and taking less risk when volatility is high. They find that managed portfolios taking less risk when volatility is high produce large alphas, substantially increase factor Sharpe ratios, and produce large utility gains for mean-variance investors.

4 Estimating Relations Between Risk-Premiums and Volatility

French et al. (1987) argue that a positive relation between the expected risk premium and ex ante predicted volatility induces a negative relation between the excess holding period return and unexpected volatility. Combining predicted and unpredicted volatility into one measure of realized volatility obscures the ex ante relation.

We illustrate their argument in Figure 4 with a series of three scatterplots, where the excess holding period returns are on the y-axis and the estimates of realized, predicted,

and unpredicted volatility are on the x-axes.

[Figure 4 here]

Figure 4 shows how combining predicted and unpredicted volatility into one realized volatility measure obscures the positive relation between predicted volatility and the expected risk premium. Decoupling predicted volatility from realized volatility exposes a positive relation between the predicted level of volatility and the expected risk premium and a strong negative relation between the excess holding period return and the unpredicted level of volatility.

4.1 *Regressions of Excess Holding Period Returns on ARIMA Forecasts of Volatility*

French et al. (1987) estimate the relation between expected risk premiums and volatility by regressing excess holding period returns on the predictable components of the stock market standard deviation or variance

$$R_{mt} - R_{ft} = \alpha + \beta \hat{\sigma}_{mt}^p + \varepsilon_t, \quad (11)$$

where each observation is weighted by predicted volatility, $\hat{\sigma}_{mt}^p$, from the ARIMA model to account for heteroskedasticity. If $\beta = 0$ in (11), the expected risk premium is unrelated to predicted volatility. If $\alpha = 0$ and $\beta > 0$, the expected risk premium is proportional to the predicted standard deviation ($p = 1$) or variance ($p = 2$) of stock returns.

We replicate FSS Table 4 and report the estimates from regression (11) in Table 4.

[Table 4 here]

We confirm that the estimates from (11) provide little evidence of a relation between expected risk premiums and predicted volatility. The 1928–84 estimate of β is 0.066 with

a standard error of 0.139 for the standard deviation specification, and 0.379 with a standard error of 0.909 for the variance specification. All of the estimates for β are within one standard error of zero.

French et al. (1987) then estimate regressions measuring the relation between excess holding period returns and contemporaneous unpredicted changes in market volatility

$$R_{mt} - R_{ft} = \alpha + \beta \hat{\sigma}_{mt}^p + \gamma \sigma_{mt}^{pu} + \varepsilon_t, \quad (12)$$

where σ_{mt}^{pu} is unpredicted volatility calculated as the difference between realized volatility and predicted volatility from the ARIMA model (4). If $\gamma < 0$, then there is a volatility feedback effect.

French et al. (1987) explain the volatility feedback effect as follows.

“Suppose this month’s standard deviation is larger than predicted. Then the model in Table 1, Panel C, implies that predicted standard deviations will be revised upward for all future time periods. If the risk premium is positively related to the predicted standard deviation, the discount rate for future cash flows will increase. If the cash flows are unaffected, the higher discount rate reduces both their present value and the current stock price. Thus, a positive relation between the predicted stock market volatility and the expected risk premium induces a negative relation between the unpredicted component of volatility and excess holding period returns.”

Table 4 also reports the estimates from regression (12). We confirm that there is a reliably negative relation between excess holding period returns and the level of unpredicted volatility. The estimates for γ for the standard deviation specification range from -0.807 to -0.879 with t -statistics from -3.959 to -8.979 . The estimates for γ for the variance specification range from -3.528 to -8.631 with t -statistics from -3.865 to -5.919 . This

confirms the main result in FSS. We confirm that there is a market-wide volatility feedback effect where unexpected volatility causes the market to reevaluate its expectation of future risk premiums and make an immediate adjustment in stock prices.

Regression (12) again provides little direct evidence of a relation between the expected risk premium and predicted volatility. Six of the eight estimates for β are negative and none are statistically significant. On the other hand, all of the estimates for α are positive in (12) and four of the eight are significant. This implies that the expected risk premium is not proportional to the predicted standard deviation nor to the predicted variance of the stock market return. This means that the ratio of excess return to volatility does not follow a constant proportionality and the expected slope of the capital market line conditional on $\hat{\sigma}_{mt}$, $E_{t-1} [(R_{mt} - R_{ft}) / \hat{\sigma}_{mt}]$, is not constant.

Table 4 summarizes the main results on volatility feedback. To make it easy to compare the current estimates to the original estimates in FSS, we add Table 5, which places our replication results from regression (12) next to the original results from regression (7) in FSS Table 4.

[Table 5 here]

Notice that our estimates for γ for the standard deviation and variance specifications are essentially the same as the original estimates in FSS in terms of size and statistical significance. This confirms that there is a market-wide volatility feedback effect and that it extends through 2019.

4.2 *GARCH-in-Mean Models*

French et al. (1987) calculate their second estimate for predicted volatility as the fitted values from the variance equation of the GARCH-in-mean model of Engle, Lilien, and Robins (1987). The GARCH-in-mean extends the GARCH by allowing the conditional

mean return to be a function of volatility. French et al. (1987) estimate two forms of the GARCH-in-mean model, one for the standard deviation and one for the variance

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t + \varepsilon_t - \theta\varepsilon_{t-1} \quad (13)$$

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t^2 + \varepsilon_t - \theta\varepsilon_{t-1}, \quad (14)$$

which is a MA(1) in the mean equation (to account for nonsynchronous trading) plus volatility. The variance equation is the same GARCH(2, 1) variance equation in (10). GARCH-in-mean is an attractive alternative to ARIMA because the return and variance processes are estimated jointly.

With the daily excess holding period returns on the S&P composite portfolio used to estimate the model, β has the same interpretation in (14) as it does in the monthly ARIMA regression (11) with ($p = 2$) because both return and variance are approximately proportional to the length of the measurement interval. On the other hand, the estimate for β in (13) should be about $\sqrt{22}$ times smaller than the monthly estimate in (11) with ($p = 1$) because standard deviation is proportional to the square root of the approximately 22-days-per-month measurement interval. The intercept α should be about 22 times smaller than the monthly estimate in (11) because it is an average daily risk premium in (13) and (14).

We replicate FSS Table 5 and report the GARCH-in-mean estimates for the daily returns in Table 6.

[Table 6 here]

We confirm that there is a positive relation between expected risk premiums and predicted volatility. The estimated coefficient of predicted volatility, β , for 1928–84 is 0.078, with a standard error of 0.040, in the standard deviation specification (13) and 2.484, with a standard error of 0.753, in the variance specification (14).

We add here that we also calculate White (1982) robust standard errors for the GARCH-in-mean parameters in addition to the uncorrected errors in FSS. While the estimates are still positive, with robust instead of uncorrected standard errors, β changes from significant to insignificant in the 1953–84 and 1985–2019 subperiods for both the standard deviation and variance specifications. The Nyblom (1989) test for constancy of the GARCH-in-mean parameters from 1928–84 confirms that the parameters are not constant in the standard deviation and variance specifications.

4.3 Comparisons of ARIMA and GARCH Models

The ARIMA models in Table 4, which use monthly excess holding period returns, and the GARCH-in-mean models in Table 6, which use daily data, yield different results vis-a-vis the risk-return relation, so FSS explore the relation between these two models further.

We replicate FSS Table 6a and report estimates of the GARCH-in-mean models (13) and (14) using monthly excess holding period returns in Table 7.

[Table 7 here]

The estimates of β in Table 7 are large compared to those in Table 4, although we do not find any that are more than two standard errors above zero with robust standard errors.

Using the same specification as (11) and (12) in Table 4, FSS estimate

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t + \varepsilon_t \quad (15)$$

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t^p + \varepsilon_t, \quad (16)$$

where, instead of the fitted values from the ARIMA model, σ_t^p represents the fitted values from the GARCH-in-mean variance equations estimated with monthly data.

We replicate FSS Table 6b and report the estimates from regressions (15) and (16) in

Table 8.

[Table 8 here]

We confirm that the estimates for β are small in relation to their standard errors for 1928–84. The estimates for β are negative for 1928–52 and positive and significant for 1953–84. The estimates for β are positive but not significant for 1985–2019. While these regressions use the GARCH-in-mean estimates of predicted volatilities, they still provide little evidence of a relation between expected risk premiums and predicted volatility.

As a final comparison of the regression and GARCH-in-mean models, we follow FSS and create a series of monthly predicted standard deviations from the daily GARCH-in-mean model in Table 6 by using the fitted GARCH process (10) to forecast σ_t^2 for each trading day in the month. We compute implied monthly standard deviation by summing the squared fitted values within the month and taking the square root of the sum. We estimate the expected monthly risk premium as the sum of the daily fitted values from the mean equation of the daily GARCH-in-mean model.

The GARCH-in-mean prediction of the monthly standard deviation is similar to the ARIMA prediction, $\hat{\sigma}_{mt}$ (the correlation is 0.874 for 1928–84 and 0.818 for 1928–2019 and the means are virtually identical). The GARCH-in-mean and ARIMA predictions have approximately the same correlation with the realized monthly standard deviation σ_{mt} (0.897 and 0.734 for 1928–84 and 0.885 and 0.686 for 1928–2019) and the sample variances for the GARCH predictions are over one third larger. The two models have similar abilities to predict monthly volatility.¹³

The behavior of the expected risk premiums implied by the regression and GARCH-in-mean models is quite different, though. French et al. (1987) plot the monthly expected risk premium from regression (11) with ($p = 1$) from Table 4 in Figure 2a and the monthly

¹³The monthly GARCH-in-mean predictor of the standard deviation has a correlation with σ_{mt} of 0.692 for 1928–84 and 0.652 for 1928–2019.

expected risk premium from the daily GARCH-in-mean model (13) in Figure 2b. To make the comparison easier, we combine their two figures into one Figure 5.

[Figure 5 here]

The correlation between the two measures is 0.75 from 1928–84 and 0.82 from 1928–2019, although the predicted risk premiums from the daily GARCH model have a much higher mean and variance than the predictions from regression (11). (The scale in Figure 5, Panel A, is from 0 to 1% and in Panel B is from 0 to 10% per month.)

The higher variability of predicted risk premiums in Panel B is caused by two factors: (i) the greater variability of the predicted standard deviation from the GARCH-in-mean model and (ii) the larger coefficient of the predicted standard deviation, β , in the GARCH-in-mean model. The estimate of β from regression (11) when ($p = 1$) is 0.066 from 1928–84 in Table 4. The comparable estimate of β is 0.078 for the standard deviation specification of the daily GARCH-in-mean model in Table 6. As discussed, the daily estimate of β in Table 6 must be multiplied by $\sqrt{N} \approx 4.5$ to make it comparable to the monthly estimate in Table 4, so $(4.5)(0.078)/(0.066) = 5.3$. Although the ARIMA and GARCH-in-mean models have similar ability to predict volatility, the GARCH-in-mean model implies greater variability in the expected risk premiums. The mean of the expected risk premiums is much higher for the GARCH-in-mean model than for the ARIMA model (1.32% versus 0.60% per month for 1928–84 and 1.42% versus 0.63% per month for 1928–2019). The GARCH-in-mean predictions seem too high because they are more than two times the mean realized premium of 0.60% reported in Table 3 for 1928–84. French et al. (1987) conclude that it is likely neither model is entirely adequate for predicting expected risk premiums.

5 Analysis of the Results

5.1 *Interpreting the Estimated Coefficients*

Merton (1980) notes that in a model of capital market equilibrium where a “representative investor” has constant relative risk aversion, there are conditions under which the expected market risk premium will be approximately proportional to the predicted variance of the market return

$$E_{t-1}(R_{mt} - R_{ft}) = C\hat{\sigma}_{mt}^2. \quad (17)$$

The parameter C in (17) is the representative investor’s coefficient of relative risk aversion (CRRA). Ignoring intercepts, the CRRA equals β in regression (11) for ($p = 2$) and in the GARCH-in-mean model (14).

The estimate of relative risk aversion, β , from regression (11) in Table 4 from 1928–84 is 0.379, although the large standard error of 0.909 does not distinguish this coefficient from zero. The corresponding GARCH-in-mean estimate of β in Table 6 is 2.484, which is over three times its robust standard error of 0.753. French et al. (1987) conclude that both estimates are economically reasonable because they are well within the range of estimates produced by other studies using different approaches.¹⁴

5.2 *The Effect of Leverage*

A negative intertemporal relation between stock returns and stock return volatility motivates an alternative hypotheses. Many of the firms whose common stocks are in the S&P composite portfolio have debt. Black (1976) and Christie (1982) suggest that leverage can induce a negative ex post relation between returns and volatility for common stocks, even if the volatility and the expected return for the total firm are constant. The *leverage effect*

¹⁴See, for example, Friend and Blume (1975), Hansen and Singleton (1982), and Brown and Gibbons (1985).

hypothesis posits that a fall (rise) in stock value increases (decreases) financial leverage, making future stock returns more (less) volatile.

The basic premise of the leverage effect hypothesis is as follows. Let V be the market value of the firm and $V = D + E$, where D is the market value of the debt and E is the market value of the equity. With risk-free debt, $dV = dE$ and stockholders bear all of the variation in market value due to changes in the expected cash flows. The stock return as represented by the percent change in the market value of the equity is then

$$\frac{dE}{E} = \frac{dV}{V} \frac{V}{E} = \frac{dV}{V} \frac{(D + E)}{E} = \frac{dV}{V} \left(1 + \frac{D}{E}\right). \quad (18)$$

With nonnegative D and E , the standard deviation of the stock return is then

$$\sigma_E = \sigma_V \left(1 + \frac{D}{E}\right). \quad (19)$$

The change in the standard deviation of the stock return divided by the stock return (the elasticity of stock return volatility with respect to stock return) is then

$$\frac{\partial \sigma_E / \sigma_E}{\partial E / E} = \frac{\partial \sigma_E}{\partial E} \times \frac{E}{\sigma_E} = \frac{-D}{D + E}, \quad (20)$$

where the lower bound is -1.0 if the equity has no value.

French et al. (1987) test if the relation between realized risk premiums and unexpected volatility is caused only by leverage by estimating

$$\ln(\sigma_{mt} / \sigma_{mt-1}) = \alpha_0 + \alpha_1 \ln(1 + R_{mt}) + \varepsilon_t, \quad (21)$$

where σ_{mt} is the realized monthly standard deviation of the S&P composite portfolio, calculated from the daily returns, and $\ln(1 + R_{mt})$ is the continuously compounded return,

calculated as the sum of the daily continuously compounded returns on that portfolio.

We estimate (21) and confirm that α_1 is significant -1.58 from 1928–84, significant -1.55 from 1928–52, significant -1.68 from 1953–84, and significant -2.36 from 1985–2019. We confirm that leverage is probably not the sole explanation for the negative relation between stock returns and volatility.

5.3 Extensions

French et al. (1987) suggest a number of extensions where the focus is on making better estimates of predicted volatility. They suggest trying the predicted variability of the real interest rate, the predicted covariance between the stock market return and consumption, and the predicted variability of decile portfolios formed on the basis of firm size as alternative measures of volatility. They also suggest trying to improve the forecasts by including other predictive variables in the models. Examples include the nominal interest rate, from Fama and Schwert (1977), and the yield spread between long-term low-grade corporate bonds and short-term Treasury bills, the level of the S&P composite index in relation to its average, and the average share price of the firms in the smallest quintile of NYSE firms, from Keim and Stambaugh (1986).

We add here that there are now automatic processes that will select the best-fitting ARIMA model by trying various combinations of $ARIMA(p, d, q)$. For example, the Hyndman-Khandakar algorithm of Hyndman and Khandakar (2008) combines unit root tests and minimization of the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) to obtain the best-fitting ARIMA model. The AIC and BIC criterion reward goodness of fit, but also include a penalty to discourage overfitting.

ARIMA models are also capable of modeling data having a seasonal pattern where, for example, high values tend to occur in some months and low values tend to occur in

other months. A seasonal ARIMA is written as

$$\text{ARIMA} \quad \underbrace{(p, d, q)}_{\substack{\text{nonseasonal part} \\ \text{of the model}}} \quad \underbrace{(P, D, Q)_m}_{\substack{\text{seasonal part} \\ \text{of the model}}} \quad (22)$$

where m is the number of observations per year. For monthly data, $m = 12$. With monthly data, a value of $P = 1$ indicates that the model uses the values of σ_{mt-12} to predict the values of σ_{mt} . With $P = 2$, the model uses the values of σ_{mt-12} and σ_{mt-24} to predict the values of σ_{mt} . A value of $Q = 1$ indicates that the model uses the values of u_{t-12} to predict σ_{mt} . The seasonal term, D , is the seasonal difference, for example, when $D = 1$ the seasonal difference is $\sigma_{mt} - \sigma_{mt-12}$.

We extend FSS by checking if automatic model selection and allowing for a seasonal pattern produces better forecasts of predicted market volatility. We use the Hyndman-Khandakar algorithm to find the best-fitting $\text{ARIMA}(p, d, q)(P, D, Q)_{12}$ model for each subperiod. Model selection is based on minimizing the AICc, a bias-corrected version of the AIC.

The ARIMA models selected by minimum AICc are, $\text{ARIMA}(1, 1, 1)(0, 0, 1)_{12}$ with Box-Cox $\lambda = -0.42$ for 1928–84, $\text{ARIMA}(2, 1, 3)$ with $\lambda = -0.79$ for 1928–52, $\text{ARIMA}(1, 1, 2)$ with $\lambda = -0.13$ for 1953–84, and $\text{ARIMA}(2, 1, 1)(1, 0, 1)_{12}$ with $\lambda = -0.57$ for 1985–2019. In terms of model accuracy, the root mean square error (RMSE) for the $\text{ARIMA}(0, 1, 3)$ models are 0.354 for 1928–84, 0.395 for 1928–52, 0.316 for 1953–84, and 0.400 for 1985–2019. The corresponding RMSE for the automatic ARIMA models are 0.023, 0.032, 0.013, and 0.022. Automatic model selection does not produce very different estimates for predicted volatility, though. The correlations between the predicted standard deviations are 0.996 for 1928–84, 0.973 for 1928–52, 0.985 for 1953–84, and 0.983 for 1985–2019. Substituting the fitted values from the automatic ARIMA models in regressions (11) and (12) does not change the main results in a meaningful way.

6 Conclusions

French et al. (1987) pose a question lying at the heart of the asset pricing literature; what is the relation between the expected market risk premium, defined as the expected return on a stock market portfolio minus the risk-free interest rate, and risk, as measured by the volatility of the stock market? French et al. (1987) then show how higher (lower) than expected volatility drives up (down) expected risk premiums, leading to an immediate decrease (increase) in stock value. French et al. (1987) not only establish this relation, but they also show how it represents a positive relation between conditional volatility and expected returns through the discount rate channel (equivalent to expected returns). These findings are significant because most (if not all) of the major multiperiod asset pricing models predict such a relation.

After careful replication, we confirm that the main results in FSS hold and continue to hold through 2019. In regressions of market excess returns on the predicted and unpredicted levels of market volatility, we confirm that there is a positive relation between the expected risk premium on common stocks and the predictable level of volatility, though it is difficult to discriminate among alternate specifications of this relation. We also confirm that there is a negative and significant relation between the unpredicted component of stock market volatility and excess holding period returns. In sum, we confirm that there is a market-wide volatility feedback effect through 2019. We also confirm that the magnitude of this effect is too large to be attributed to the leverage effect discussed in Black (1976) and Christie (1982).

References

- Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002). "Range-based estimation of stochastic volatility models". *Journal of Finance* 57 (3), pp. 1047–1091.
- Allen, F. and R. Karjalainen (1999). "Using genetic algorithms to find technical trading rules". *Journal of Financial Economics* 51, pp. 245–271.
- Amihud, Y. (2002). "Illiquidity and stock returns: cross-section and time-series effects". *Journal of Financial Markets* 5, pp. 31–56.
- Amihud, Y. (2019). "Illiquidity and stock returns: a revisit". *Critical Finance Review* 8, pp. 203–221.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001a). "The distribution of realized exchange rate volatility". *Journal of the American Statistical Association* 96 (453), pp. 42–55.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001b). "The distribution of realized stock return volatility". *Journal of Financial Economics* 61, pp. 43–76.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). "Modeling and forecasting realized volatility". *Econometrica* 71 (2), pp. 579–625.
- Andrei, D., M. Hasler, and A. Jeanneret (2019). "Asset pricing with persistence risk". *Review of Financial Studies* 32 (7), pp. 2809–2849.
- Ang, A. and J. Chen (2002). "Asymmetric correlations of equity portfolios". *Journal of Financial Economics* 63, pp. 443–494.
- Antweiler, W. and M. Z. Frank (2004). "Is all that talk just noise? The information content of internet stock message boards". *Journal of Finance* 59 (3), pp. 1259–1294.
- Baillie, R. T. and R. P. DeGennaro (1990). "Stock returns and volatility". *Journal of Financial and Quantitative Analysis* 25 (2), pp. 203–214.

- Bakshi, G. and N. Kapadia (2003). "Delta-hedged gains and the negative market volatility risk premium". *Review of Financial Studies* 16 (2), pp. 527–566.
- Bandi, F. M. and J. R. Russell (2006). "Separating microstructure noise from volatility". *Journal of Financial Economics* 79, pp. 655–692.
- Bansal, R. and A. Yaron (2004). "Risks for the long run: a potential resolution of asset pricing puzzles". *Journal of Finance* 59 (4), pp. 1481–1509.
- Bekaert, G., C. R. Harvey, and C. Lundblad (2007). "Liquidity and expected returns: lessons from emerging markets". *Review of Financial Studies* 20 (6), pp. 1783–1831.
- Bernanke, B. S. and K. N. Kuttner (2005). "What explains the stock market's reaction to Federal Reserve policy?" *Journal of Finance* 60 (3), pp. 1221–1257.
- Bessembinder, H. and P. J. Seguin (1993). "Price volatility, trading volume, and market depth: evidence from futures markets". *Journal of Financial and Quantitative Analysis* 28 (1), pp. 21–39.
- Bibinger, M., C. Neely, and L. Winkelmann (2019). "Estimation of the discontinuous leverage effect: evidence from the NASDAQ order book". *Journal of Econometrics* 209, pp. 158–184.
- Black, F. (1976). "Studies of stock price volatility changes". *Proceedings of the business and economic statistics section. American Statistical Association*, pp. 177–181.
- Black, F., M. C. Jensen, and M. Scholes (1972). "The capital asset pricing model: some empirical tests". *Studies in the theory of capital markets*. Ed. by M. C. Jensen. Praeger, New York.
- Bollerslev, T. (1986). "Generalized autoregressive conditional heteroskedasticity". *Journal of Econometrics* 31, pp. 307–327.
- Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). "ARCH modeling in finance: a review of the theory and empirical evidence". *Journal of Econometrics* 52, pp. 5–59.

- Box, G. E. P. and D. R. Cox (1964). "An analysis of transformations". *Journal of the Royal Statistical Society. Series B (Methodological)* 26 (2), pp. 211–252.
- Box, G. E. P. and D. A. Pierce (1970a). "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models". *Journal of the American Statistical Association* 65 (332), pp. 1509–1526.
- Box, G. E. P. and G. M. Jenkins (1970b). *Time series analysis: forecasting and control*. San Francisco: Holden-Day.
- Breeden, D. T., M. R. Gibbons, and R. H. Litzenberger (1989). "Empirical tests of the consumption-oriented CAPM". *Journal of Finance* 44 (2), pp. 231–262.
- Brock, W., J. Lakonishok, and B. LeBaron (1992). "Simple technical trading rules and the stochastic properties of stock returns". *Journal of Finance* 47 (5), pp. 1731–1764.
- Brown, D. P. and M. R. Gibbons (1985). "A simple econometric approach for utility-based asset pricing models". *Journal of Finance* 40 (2), pp. 359–381.
- Bushee, B. J. and C. F. Noe (2000). "Corporate disclosure practices, institutional investors, and stock return volatility". *Journal of Accounting Research* 38, pp. 171–202.
- Cai, J. (1994). "A Markov model of switching-regime ARCH". *Journal of Business & Economic Statistics* 12 (3), pp. 309–316.
- Campbell, J. Y. and L. Hentschel (1992). "No news is good news: an asymmetric model of changing volatility in stock returns". *Journal of Financial Economics* 31, pp. 281–318.
- Campbell, J. Y. and J. H. Cochrane (1999). "By force of habit: a consumption-based explanation of aggregate stock market behavior". *Journal of Political Economy* 107 (2), pp. 205–251.
- Campbell, J. Y., M. Lettau, B. G. Malkiel, and Y. Xu (2001). "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk". *Journal of Finance* 56 (1), pp. 1–43.

- Chan, K., K. C. Chan, and G. A. Karolyi (1991). "Intraday volatility in the stock index and stock index futures markets". *Review of Financial Studies* 4 (4), pp. 657–684.
- Chen, J., H. Hong, and J. C. Stein (2001). "Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices". *Journal of Financial Economics* 61, pp. 345–381.
- Chen, N.-F. (1991). "Financial investment opportunities and the macroeconomy". *Journal of Finance* 46 (2), pp. 529–554.
- Chen, N.-F., R. Roll, and S. A. Ross (1986). "Economic forces and the stock market". *Journal of Business* 59 (3), pp. 383–403.
- Christensen, B. and N. Prabhala (1998). "The relation between implied and realized volatility". *Journal of Financial Economics* 50, pp. 125–150.
- Christie, A. A. (1982). "The stochastic behavior of common stock variances". *Journal of Financial Economics* 10, pp. 407–432.
- Cutler, D. M., J. M. Poterba, and L. H. Summers (1989). "What moves stock prices?" *Journal of Portfolio Management* 15 (3), pp. 4–12.
- Dickey, D. A. and W. A. Fuller (1979). "Distribution of the estimators for autoregressive time series with a unit root". *Journal of the American Statistical Association* 74 (366), pp. 427–431.
- Drienko, J., T. Smith, and A. von Reibnitz (2019). "A review of the return-illiquidity relationship". *Critical Finance Review* 8, pp. 127–171.
- Easley, D. and M. O'Hara (1992). "Time and the process of security price adjustment". *Journal of Finance* 47 (2), pp. 577–605.
- Edmans, A., D. García, and O. Norli (2007). "Sports sentiment and stock returns". *Journal of Finance* 62 (4), pp. 1967–1998.
- Engle, R. (2002). "New frontiers for ARCH models". *Journal of Applied Econometrics* 17, pp. 425–446.

- Engle, R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". *Econometrica* 50 (4), pp. 987–1007.
- Engle, R. F., D. M. Lilien, and R. P. Robins (1987). "Estimating time varying risk premia in the term structure: the ARCH-M model". *Econometrica* 55 (2), pp. 391–407.
- Engle, R. F., V. K. Ng, and M. Rothschild (1990). "Asset pricing with a factor-ARCH covariance structure: empirical estimates for Treasury bills". *Journal of Econometrics* 45, pp. 213–237.
- Engle, R. F. and G. González-Rivera (1991). "Semiparametric ARCH models". *Journal of Business & Economic Statistics* 9 (4), pp. 345–359.
- Fama, E. F. (1976). *Foundations of finance: Portfolio decisions and securities prices*. Basic Books, Inc.
- Fama, E. F. (1990). "Stock returns, expected returns, and real activity". *Journal of Finance* 45 (4), pp. 1089–1108.
- Fama, E. F. and J. D. MacBeth (1973). "Risk, return, and equilibrium: empirical tests". *Journal of Political Economy* 81 (3), pp. 607–636.
- Fama, E. F. and G. W. Schwert (1977). "Asset returns and inflation". *Journal of Financial Economics* 5, pp. 115–146.
- Fama, E. F. and K. R. French (1988). "Permanent and temporary components of stock prices". *Journal of Political Economy* 96 (2), pp. 246–273.
- Fama, E. F. and K. R. French (1989). "Business conditions and expected returns on stocks and bonds". *Journal of Financial Economics* 25, pp. 23–49.
- Ferson, W. E. and C. R. Harvey (1993). "The risk and predictability of international equity returns". *Review of Financial Studies* 6 (3), pp. 527–566.
- Fisher, L. (1966). "Some new stock-market indexes". *Journal of Business* 39 (1), pp. 191–225.

- Fleming, J., B. Ostdiek, and R. E. Whaley (1996). "Trading costs and the relative rates of price discovery in stock, futures, and option markets". *Journal of Futures Markets* 16 (4), pp. 353–387.
- French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). "Expected stock returns and volatility". *Journal of Financial Economics* 19, pp. 3–29.
- Friend, I. and M. E. Blume (1975). "The demand for risky assets". *American Economic Review* 65 (5), pp. 900–922.
- Fu, F. (2009). "Idiosyncratic risk and the cross-section of expected stock returns". *Journal of Financial Economics* 91, pp. 24–37.
- Ghalanos, A. (2020). *rugarch: Univariate GARCH models*. R package version 1.4-4.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2005). "There is a risk-return trade-off after all". *Journal of Financial Economics* 76, pp. 509–548.
- Goyal, A. and P. Santa-Clara (2003). "Idiosyncratic risk matters!" *Journal of Finance* 58 (3), pp. 975–1007.
- Hamao, Y., R. W. Masulis, and V. Ng (1990). "Correlations in price changes and volatility across international stock markets". *Review of Financial Studies* 3 (2), pp. 281–307.
- Hansen, L. P. and K. J. Singleton (1982). "Generalized instrumental variables estimation of nonlinear rational expectations models". *Econometrica* 50 (5), pp. 1269–1286.
- Harvey, C. R. (1989). "Time-varying conditional covariances in tests of asset pricing models". *Journal of Financial Economics* 24, pp. 289–317.
- Harvey, C. R. and A. Siddique (1999). "Autoregressive conditional skewness". *Journal of Financial and Quantitative Analysis* 34 (4), pp. 465–487.
- Haugen, R. A. and N. L. Baker (1996). "Commonality in the determinants of expected stock returns". *Journal of Financial Economics* 41, pp. 401–439.
- Hentschel, L. (1995). "All in the family: nesting symmetric and asymmetric GARCH models". *Journal of Financial Economics* 39, pp. 71–104.

- Hodrick, R. J. (1992). "Dividend yields and expected stock returns: alternative procedures for inference and measurement". *Review of Financial Studies* 5 (3), pp. 357–386.
- Hong, H. and J. C. Stein (2003). "Differences of opinion, short-sales constraints, and market crashes". *Review of Financial Studies* 16 (2), pp. 487–525.
- Hsieh, D. A. (1991). "Chaos and nonlinear dynamics: application to financial markets". *Journal of Finance* 46 (5), pp. 1839–1877.
- Hutton, A. P., A. J. Marcus, and H. Tehranian (2009). "Opaque financial reports, R^2 , and crash risk". *Journal of Financial Economics* 94, pp. 67–86.
- Hyndman, R., G. Athanasopoulos, C. Bergmeir, G. Caceres, L. Chhay, M. O'Hara-Wild, F. Petropoulos, S. Razbash, E. Wang, and F. Yasmeeen (2020). *forecast: Forecasting functions for time series and linear models*. R package version 8.12. URL: <http://pkg.robjhyndman.com/forecast>.
- Hyndman, R. J. and Y. Khandakar (2008). "Automatic time series forecasting: the forecast package for R". *Journal of Statistical Software* 27 (3), pp. 1–22.
- Jang, J. and J. Kang (2019). "Probability of price crashes, rational speculative bubbles, and the cross-section of stock returns". *Journal of Financial Economics* 132, pp. 222–247.
- Jiang, F., J. Lee, X. Martin, and G. Zhou (2019). "Manager sentiment and stock returns". *Journal of Financial Economics* 132, pp. 126–149.
- Jiang, G. J. and Y. S. Tian (2005). "The model-free implied volatility and its information content". *Review of Financial Studies* 18 (4), pp. 1305–1342.
- Jondeau, E., Q. Zhang, and X. Zhu (2019). "Average skewness matters". *Journal of Financial Economics* 134, pp. 29–47.
- Kandel, S. and R. F. Stambaugh (1996). "On the predictability of stock returns: an asset-allocation perspective". *Journal of Finance* 51 (2), pp. 385–424.
- Keim, D. B. and R. F. Stambaugh (1986). "Predicting returns in the stock and bond markets". *Journal of Financial Economics* 17, pp. 357–390.

- Kennedy, P. (2003). *A guide to econometrics*. 5th ed. The MIT Press.
- Kilic, M. and I. Shaliastovich (2019). "Good and bad variance premia and expected returns". *Management Science* 65 (6), pp. 2522–2544.
- Kothari, S. P., X. Li, and J. E. Short (2009). "The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a study using content analysis". *Accounting Review* 84 (5), pp. 1639–1670.
- Kottimukkalur, B. (2019). "Attention to market information and underreaction to earnings on market moving days". *Journal of Financial and Quantitative Analysis* 54 (6), pp. 2493–2516.
- Lamoureux, C. G. and W. D. Lastrapes (1990). "Persistence in variance, structural change, and the GARCH model". *Journal of Business & Economic Statistics* 8 (2), pp. 225–234.
- Lee, W. Y., C. X. Jiang, and D. C. Indro (2002). "Stock market volatility, excess returns, and the role of investor sentiment". *Journal of Banking & Finance* 26, pp. 2277–2299.
- Lewellen, J. and S. Nagel (2006). "The conditional CAPM does not explain asset-pricing anomalies". *Journal of Financial Economics* 82, pp. 289–314.
- Lewis, D. D., Y. Yang, T. G. Rose, and F. Li (2004). "RCV1: a new benchmark collection for text categorization research". *Journal of Machine Learning Research* 5, pp. 361–397.
- Liu, J. (2007). "Portfolio selection in stochastic environments". *Review of Financial Studies* 20 (1), pp. 1–39.
- Ljung, G. M. and G. E. P. Box (1978). "On a measure of lack of fit in time series models". *Biometrika* 65 (2), pp. 297–303.
- Lo, A. W. and A. C. MacKinlay (1988). "Stock market prices do not follow random walks: evidence from a simple specification test". *Review of Financial Studies* 1 (1), pp. 41–66.
- Ludvigson, S. C. and S. Ng (2007). "The empirical risk-return relation: a factor analysis approach". *Journal of Financial Economics* 83, pp. 171–222.

- McQueen, G. and V. V. Roley (1993). "Stock prices, news, and business conditions". *Review of Financial Studies* 6 (3), pp. 683–707.
- Merton, R. C. (1973). "An intertemporal capital asset pricing model". *Econometrica* 41 (5), pp. 867–887.
- Merton, R. C. (1980). "On estimating the expected return on the market: an exploratory investigation". *Journal of Financial Economics* 8, pp. 323–361.
- Moreira, A. and T. Muir (2017). "Volatility-managed portfolios". *Journal of Finance* 72 (4), pp. 1611–1644.
- Moreira, A. and T. Muir (2019). "Should long-term investors time volatility?" *Journal of Financial Economics* 131, pp. 507–527.
- Nelson, D. B. (1990). "ARCH models as diffusion approximations". *Journal of Econometrics* 45, pp. 7–38.
- Nelson, D. B. (1991). "Conditional heteroskedasticity in asset returns: a new approach". *Econometrica* 59 (2), pp. 347–370.
- Nelson, D. B. and C. Q. Cao (1992). "Inequality constraints in the univariate GARCH model". *Journal of Business & Economic Statistics* 10 (2), pp. 229–235.
- Nyblom, J. (1989). "Testing for the constancy of parameters over time". *Journal of the American Statistical Association* 84 (405), pp. 223–230.
- Officer, R. R. (1973). "The variability of the market factor of the New York Stock Exchange". *Journal of Business* 46 (3), pp. 434–453.
- Pagan, A. R. and G. W. Schwert (1990). "Alternative models for conditional stock volatility". *Journal of Econometrics* 45, pp. 267–290.
- Pindyck, R. S. (1984). "Risk, inflation, and the stock market". *American Economic Review* 74 (3), pp. 335–351.
- Poterba, J. M. and L. H. Summers (1986). "The persistence of volatility and stock market fluctuations". *American Economic Review* 76 (5), pp. 1142–1151.

- Pyun, S. (2019). "Variance risk in aggregate stock returns and time-varying return predictability". *Journal of Financial Economics* 132, pp. 150–174.
- R Core Team (2020). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.r-project.org/>.
- Robins, R. P. and G. P. Smith (2020). "Selection bias and pseudo discoveries on the constancy of stock return anomalies". *Review of Quantitative Finance and Accounting*, forthcoming.
- Roll, R. and S. A. Ross (1980). "An empirical investigation of the arbitrage pricing theory". *Journal of Finance* 35 (5), pp. 1073–1103.
- Said, S. E. and D. A. Dickey (1984). "Testing for unit roots in autoregressive-moving average models of unknown order". *Biometrika* 71 (3), pp. 599–607.
- Saunders Jr., E. M. (1993). "Stock prices and Wall Street weather". *American Economic Review* 83 (5), pp. 1337–1345.
- Scholes, M. and J. Williams (1977). "Estimating betas from nonsynchronous data". *Journal of Financial Economics* 5, pp. 309–327.
- Schwert, G. W. (1989a). "Tests for unit roots: a Monte Carlo investigation". *Journal of Business & Economic Statistics* 7 (2), pp. 147–159.
- Schwert, G. W. (1989b). "Why does stock market volatility change over time?" *Journal of Finance* 44 (5), pp. 1115–1153.
- Schwert, G. W. (1990). "Indexes of U.S. stock prices from 1802 to 1987". *Journal of Business* 63 (3), pp. 399–426.
- Shanken, J. (1990). "Intertemporal asset pricing: an empirical investigation". *Journal of Econometrics* 45, pp. 99–120.
- Tauchen, G. and R. Hussey (1991). "Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models". *Econometrica* 59 (2), pp. 371–396.

- Veronesi, P. (1999). "Stock market overreaction to bad news in good times: a rational expectations equilibrium model". *Review of Financial Studies* 12 (5), pp. 975–1007.
- Weinberg, S. L., D. Harel, and S. K. Abramowitz (2021). *Statistics using R: an integrative approach*. New York: Cambridge University Press.
- White, H. (1980). "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity". *Econometrica* 48 (4), pp. 817–838.
- White, H. (1982). "Maximum likelihood estimation of misspecified models". *Econometrica* 50 (1), pp. 1–25.
- Yu, J. and Y. Yuan (2011). "Investor sentiment and the mean-variance relation". *Journal of Financial Economics* 100, pp. 367–381.

Table 1: Time-Series Properties of Estimates of the Standard Deviation of the Return to the Standard & Poor's Composite Portfolio.^a

Description: This Table summarizes descriptive statistics for monthly realized standard deviation in (A) and the first difference in its natural logarithm in (B). The results for regression (4) are given in (C).

Interpretation: This Table summarizes monthly realized volatility and gives the parameters for calculating monthly predicted volatility.

Period	Mean	Std. dev.	Skewness ^d	Autocorrelation at lags												Std. error	Q(12)	
				1	2	3	4	5	6	7	8	9	10	11	12			
(A) Monthly standard deviation of S & P composite returns estimated from daily data																		
1928-84	0.0477	0.0328	2.87 ^b	0.71	0.57	0.54	0.53	0.48	0.48	0.49	0.53	0.53	0.48	0.50	0.48	0.45	0.00	2,310.0 ^b
1928-52	0.0610	0.0422	2.15 ^b	0.67	0.51	0.49	0.48	0.42	0.44	0.50	0.50	0.50	0.43	0.45	0.44	0.39	0.00	851.0 ^b
1953-84	0.0374	0.0170	1.75 ^b	0.62	0.49	0.37	0.34	0.31	0.27	0.24	0.24	0.22	0.24	0.25	0.20	0.17	0.00	522.5 ^b
1985-2019	0.0418	0.0255	3.63 ^b	0.50	0.39	0.29	0.23	0.25	0.24	0.24	0.21	0.20	0.21	0.19	0.13	0.12	0.00	364.1 ^b
(B) Log differences of monthly standard deviation of S & P composite returns estimated from daily data																		
2/28-12/84	-0.0000	0.4073	0.15	-0.34	-0.08	-0.09	0.08	-0.05	-0.04	0.00	0.00	0.05	-0.04	0.02	-0.01	0.08	0.02	102.1 ^b
2/28-12/52	-0.0012	0.4638	0.22	-0.32	-0.13	-0.09	0.14	-0.10	-0.07	0.02	0.12	0.12	-0.06	0.00	-0.01	0.13	0.03	60.5 ^b
2/53-12/84	0.0008	0.3582	0.03	-0.35	-0.01	-0.10	0.00	0.01	0.00	-0.02	-0.04	-0.04	-0.01	0.04	-0.02	0.00	0.02	53.2 ^b
2/85-12/19	-0.0015	0.4736	0.17	-0.45	0.04	-0.04	-0.06	-0.02	0.03	0.01	0.00	0.00	-0.01	0.09	-0.10	0.05	0.02	95.7 ^b
(C) ARIMA models for the logarithm of the monthly standard deviation of S & P composite returns estimated from daily data ^c																		
$(1-L)\ln\sigma_{mt} = \theta_0 + (1-\theta_1L - \theta_2L^2 - \theta_3L^3)u_t$ (4)																		
Period	θ_0	θ_1	θ_2	θ_3	$S(u_t)$	R^2	Q(12)	Skewness ^d	SR(u_t)	F-test for stability ^e								
1928-84	0.0001 (0.0030)	0.534 (0.039)	0.156 (0.040)	0.090 (0.037)	0.354	0.538	10.4	0.16	11.16 ^b	0.99								
1928-52	0.0014 (0.0050)	0.568 (0.060)	0.189 (0.058)	0.030 (0.055)	0.395	0.489	19.1 ^b	0.06	10.04 ^b									
1953-84	-0.0006 (0.0038)	0.513 (0.051)	0.093 (0.058)	0.163 (0.052)	0.316	0.406	3.2	0.29 ^b	6.64 ^b									
1985-2019	0.0002 (0.0043)	0.656 (0.049)	-0.003 (0.063)	0.128 (0.050)	0.396	0.283	7.5	0.29 ^b	8.04 ^b									

^a The monthly standard deviation estimator, σ_{mt} , is calculated from the daily rates of return to the Standard & Poor's composite portfolio for each day in the month

$$\sigma_{mt}^2 = \sum_{i=1}^{N_t} r_{it}^2 + 2 \sum_{i=1}^{N_t-1} r_{it}r_{i+1,t}$$

where r_{it} is the return on day i within month t , and there are N_t days in the month. $Q(12)$ is the Ljung and Box (1978) statistic for 12 lags of the autocorrelation function, and $SR(u_t)$ is the studentized range, the sample range divided by the standard deviation. See Fama (1976), chapter 1, for a discussion and fractiles of SR under the hypothesis of a stationary normal distribution.

^b Greater than the 0.95 fractile of the sampling distribution under the hypothesis of a stationary, serially uncorrelated normal distribution.

^c Standard errors are in parentheses.

^d The asymptotic standard error for the sample skewness (SES) is $\sqrt{\frac{6t(T-1)}{(T-2)(T+1)(T+3)}}$ under the hypothesis of a stationary normal distribution. The standard error equals 0.093, 0.141, 0.125, and 0.119 for $T = 684, 300, 384$, and 420.

^e The F -test for stability of the time series models is based on the residual sums of squares (RSS) from the subperiods and for the overall sample period. The RSS is 85.44 for 1928-84, 46.72 for 1928-52, and 38.22 for 1953-84.

Table 2: Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Models for Daily Excess Holding Period Returns to the Standard & Poor's Composite Portfolio.^a

$$R_{mt} - R_{ft} = \alpha + \varepsilon_t - \theta\varepsilon_{t-1} \quad (9)$$

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2 \quad (10)$$

Description: This Table summarizes the results from the GARCH regression (9) and (10).

Interpretation: Conditional volatility is autocorrelated and predictable.

GARCH model equations	$\alpha \times 10^3$	$a \times 10^3$	b	c_1	c_2	θ	Nyblom test for stability
<i>(A) January 1928 to December 1984, T = 15,370</i>							
GARCH (9), (10)	0.405 (0.065) [0.097]	0.001 (0.000) [0.003]	0.908 (0.006) [0.069]	0.089 (0.008) [0.013]	0.000 (0.005) [0.073]	0.157 (0.009) [0.012]	1,209.7
<i>(B) January 1928 to December 1952, T = 7,327</i>							
GARCH (9), (10)	0.679 (0.111) [0.139]	0.001 (0.001) [0.010]	0.894 (0.014) [0.141]	0.103 (0.013) [0.021]	0.000 (0.019) [0.137]	0.089 (0.013) [0.015]	
<i>(C) January 1953 to December 1984, T = 8,043</i>							
GARCH (9), (10)	0.298 (0.081) [0.123]	0.001 (0.000) [0.002]	0.907 (0.007) [0.068]	0.086 (0.011) [0.018]	0.000 (0.008) [0.073]	0.213 (0.012) [0.015]	
<i>(D) January 1985 to December 2019, T = 8,822</i>							
GARCH (9), (10)	0.544 (0.082) [0.089]	0.002 (0.001) [0.009]	0.879 (0.014) [0.128]	0.094 (0.011) [0.033]	0.011 (0.018) [0.121]	-0.014 (0.012) [0.011]	

^a $R_{mt} - R_{ft}$ is the daily excess holding period return to the Standard & Poor's composite portfolio (the percentage price change minus the risk-free interest rate). Nonlinear optimization techniques are used to calculate maximum likelihood estimates. Asymptotic standard errors are in parentheses under the coefficient estimates. The numbers in brackets are robust standard errors calculated using the method of White (1982).

Table 3: Means, Standard Deviations, and Skewness of the Monthly CRSP Value-Weighted Market Excess Holding Period Returns (*t*-statistics in parentheses).^a

Description: This Table summarizes descriptive statistics for monthly risk premiums. Means are calculated by ordinary least squares and weighted least squares (WLS).

Interpretation: Monthly risk premiums are heteroskedastic. In periods of unexpectedly high volatility, realized stock returns are lower than average.

Period	Mean	WLS mean ^b	WLS mean ^c	Std. dev.	Skewness
1928–84	0.0060 (0.002)	0.0091 (0.002)	0.0055 (0.002)	0.0586	0.49 ^d
1928–52	0.0075 (0.004)	0.0120 (0.003)	0.0075 (0.004)	0.0753	0.49 ^d
1953–84	0.0048 (0.002)	0.0074 (0.002)	0.0045 (0.002)	0.0411	−0.05
1985–2019	0.0068 (0.002)	0.0101 (0.002)	0.0069 (0.002)	0.0413	−0.92 ^d

^a The one-month risk-free rate is subtracted from the CRSP value-weighted stock market return to create an excess holding period return.

^b Sample mean estimated by weighted least squares, where the standard deviation of the Standard & Poor’s composite portfolio estimated from the days within the month, σ_{mt} , is used to weight the observations.

^c Sample mean estimated by weighted least squares, where the predicted standard deviation of the Standard & Poor’s composite portfolio estimated from the ARIMA model in Table 1, Panel C, is used to weight the observations.

^d Greater than the 0.95 fractile of the sampling distribution under the hypothesis of a stationary, serially uncorrelated normal distribution.

Table 4: Weighted Least Squares Regressions of Monthly CRSP Value-Weighted Excess Holding Period Returns against the Predictable and Unpredictable Components of the Standard Deviations or Variances of Stock Market Returns.^a

$$R_{mt} - R_{ft} = \alpha + \beta \hat{\sigma}_{mt}^p + \varepsilon_t \quad (11)$$

$$R_{mt} - R_{ft} = \alpha + \beta \hat{\sigma}_{mt}^p + \gamma \sigma_{mt}^{pu} + \varepsilon_t \quad (12)$$

Description: This Table summarizes the results from regressions (11) and (12).

Interpretation: There is a strong negative relation between unpredicted volatility and excess returns.

Volatility measure	Eq. (11)		Eq. (12)						
	α	β	α	β	γ	$S(\varepsilon)$	R^2	$Q(12)$	$SR(\varepsilon)$
<i>(A) January 1928 to December 1984, T = 684</i>									
σ_{mt}	0.0028 (0.0045) [0.0054]	0.066 (0.102) [0.139]	0.0050 (0.0042) [0.0055]	0.016 (0.097) [0.143]	-0.830 (0.093) [0.127]	0.0568	0.106	14.4	9.18
σ_{mt}^2	0.0047 (0.0022) [0.0022]	0.379 (0.938) [0.909]	0.0054 (0.0021) [0.0021]	0.187 (0.892) [0.938]	-3.973 (0.463) [0.795]	0.0579	0.098	16.5	6.49
<i>(B) January 1928 to December 1952, T = 300</i>									
σ_{mt}	0.0076 (0.0086) [0.0099]	-0.001 (0.154) [0.202]	0.0107 (0.0081) [0.0100]	-0.058 (0.145) [0.210]	-0.807 (0.125) [0.158]	0.0733	0.122	14.8	8.11
σ_{mt}^2	0.0089 (0.0043) [0.0043]	-0.250 (1.103) [1.086]	0.0108 (0.0039) [0.0040]	-0.649 (1.018) [1.128]	-3.528 (0.479) [0.596]	0.0748	0.154	9.7	6.57
<i>(C) January 1953 to December 1984, T = 384</i>									
σ_{mt}	0.0012 (0.0064) [0.0063]	0.096 (0.177) [0.183]	0.0049 (0.0062) [0.0064]	-0.012 (0.171) [0.194]	-0.879 (0.154) [0.222]	0.0399	0.079	14.3	6.29
σ_{mt}^2	0.0035 (0.0032) [0.0031]	0.759 (2.213) [2.020]	0.0050 (0.0031) [0.0030]	-0.350 (2.140) [2.159]	-8.631 (1.548) [2.233]	0.0408	0.076	11.7	6.22
<i>(D) January 1985 to December 2019, T = 420</i>									
σ_{mt}	0.0081 (0.0054) [0.0052]	-0.033 (0.136) [0.144]	0.0123 (0.0049) [0.0049]	-0.141 (0.123) [0.134]	-0.862 (0.086) [0.096]	0.0369	0.195	18.9	6.34
σ_{mt}^2	0.0079 (0.0027) [0.0027]	-0.460 (1.424) [1.371]	0.0086 (0.0024) [0.0025]	-0.697 (1.300) [1.246]	-4.440 (0.478) [0.772]	0.0370	0.172	15.8	6.36

^a $\hat{\sigma}_{mt}$ is the prediction and σ_{mt}^u is the prediction error for the estimate of the monthly stock market standard deviation from the ARIMA model in Table 1, Panel C. $\hat{\sigma}_{mt}^2$ and σ_{mt}^{2u} are the prediction and prediction error for the variance of stock returns. The estimated time series model for σ_{mt} is reported in Table 1, Panel C. Standard errors are in parentheses below the coefficient estimates. The numbers in brackets are standard errors based on the White (1980) consistent heteroskedasticity correction. $S(\varepsilon)$ is the standard deviation of the residuals, R^2 is the coefficient of determination, $Q(12)$, is the Ljung and Box (1978) statistic for 12 lags of the residual autocorrelation function, and $SR(\varepsilon)$ is the studentized range of the residuals. These regressions are estimated using weighted least squares, where the predicted standard deviation of the S&P composite portfolio $\hat{\sigma}_{mt}$ is used to standardize each observation. R^2 , $Q(12)$ and $SR(\varepsilon)$ are based on the weighted residuals, but the standard deviation of the residuals is based on the unweighted residuals (in the same units as the original data).

Table 5: Comparison of Weighted Least Squares Regressions of Monthly CRSP Value-Weighted Excess Holding Period Returns against the Predictable and Unpredictable Components of the Standard Deviations or Variances of Stock Market Returns.^a

$$R_{mt} - R_{ft} = \alpha + \beta \hat{\sigma}_{mt}^p + \gamma \sigma_{mt}^{pu} + \varepsilon_t \quad (12)$$

Description: This Table compares the current estimates from regression (12) to the original estimates from regression (7) in FSS.

Interpretation: The main results in FSS hold and continue to hold through 2019. The relation between unpredicted volatility and excess returns is negative and significant. There is a market-wide volatility feedback effect.

Volatility measure	Current estimates			Original estimates		
	α	β	γ	α	β	γ
<i>(A) January 1928 to December 1984, T = 684</i>						
σ_{mt}	0.0050 (0.0042) [0.0055]	0.016 (0.097) [0.143]	-0.830 (0.093) [0.127]	0.0077 (0.0039) [0.0039]	-0.050 (0.107) [0.105]	-1.010 (0.092) [0.111]
σ_{mt}^2	0.0054 (0.0021) [0.0021]	0.187 (0.892) [0.938]	-3.973 (0.463) [0.795]	0.0057 (0.0020) [0.0021]	0.088 (0.889) [0.930]	-4.438 (0.496) [0.886]
<i>(B) January 1928 to December 1952, T = 300</i>						
σ_{mt}	0.0107 (0.0081) [0.0100]	-0.058 (0.145) [0.210]	-0.807 (0.125) [0.158]	0.0199 (0.0076) [0.0081]	-0.230 (0.163) [0.175]	-1.007 (0.115) [0.131]
σ_{mt}^2	0.0108 (0.0039) [0.0040]	-0.649 (1.018) [1.128]	-3.528 (0.479) [0.596]	0.0114 (0.0038) [0.0039]	-0.671 (1.042) [1.140]	-3.985 (0.515) [0.707]
<i>(C) January 1953 to December 1984, T = 384</i>						
σ_{mt}	0.0049 (0.0062) [0.0064]	-0.012 (0.171) [0.194]	-0.879 (0.154) [0.222]	0.0068 (0.0056) [0.0051]	-0.071 (0.172) [0.161]	-1.045 (0.152) [0.205]
σ_{mt}^2	0.0050 (0.0031) [0.0030]	-0.350 (2.140) [2.159]	-8.631 (1.548) [2.233]	0.0046 (0.0031) [0.0031]	-0.349 (2.118) [2.186]	-9.075 (1.571) [2.382]
<i>(D) January 1985 to December 2019, T = 420</i>						
σ_{mt}	0.0123 (0.0049) [0.0049]	-0.141 (0.123) [0.134]	-0.862 (0.086) [0.096]			
σ_{mt}^2	0.0086 (0.0024) [0.0025]	-0.697 (1.300) [1.246]	-4.440 (0.478) [0.772]			

^a $\hat{\sigma}_{mt}$ is the prediction and σ_{mt}^u is the prediction error for the estimate of the monthly stock market standard deviation from the ARIMA model in Table 1, Panel C. $\hat{\sigma}_{mt}^2$ and σ_{mt}^{2u} are the prediction and prediction error for the variance of stock returns. The estimated time series model for σ_{mt} is reported in Table 1, Panel C. Standard errors are in parentheses below the coefficient estimates. The numbers in brackets are standard errors based on the White (1980) consistent heteroskedasticity correction.

Table 6: Generalized Autoregressive Conditional Heteroskedasticity-in-Mean (GARCH-in-Mean) Models for Daily Excess Holding Period Returns to the Standard & Poor's Composite Portfolio.^a

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t + \varepsilon_t - \theta\varepsilon_{t-1} \quad (13)$$

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t^2 + \varepsilon_t - \theta\varepsilon_{t-1} \quad (14)$$

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2 \quad (10)$$

Description: This Table summarizes the results from GARCH-in-mean regressions with daily data.

Interpretation: The relation between predicted volatility and excess returns is positive.

GARCH-in-mean equations	$\alpha \times 10^3$	β	$a \times 10^3$	b	c_1	c_2	θ	Nyblom test for stability
<i>(A) January 1928 to December 1984, T = 15,370</i>								
Std. dev. (13), (10)	-0.111 (0.168) [0.343]	0.078 (0.023) [0.040]	0.001 (0.000) [0.003]	0.908 (0.006) [0.067]	0.090 (0.008) [0.013]	0.000 (0.005) [0.072]	0.157 (0.009) [0.012]	1,224.6
Variance (14), (10)	0.276 (0.080) [0.100]	2.484 (0.921) [0.753]	0.001 (0.000) [0.001]	0.908 (0.003) [0.021]	0.090 (0.009) [0.080]	0.000 (0.003) [0.057]	0.156 (0.009) [0.015]	1,209.4
<i>(B) January 1928 to December 1952, T = 7,327</i>								
Std. dev. (13), (10)	0.287 (0.278) [0.362]	0.048 (0.031) [0.035]	0.001 (0.001) [0.009]	0.893 (0.014) [0.133]	0.103 (0.013) [0.021]	0.000 (0.019) [0.130]	0.090 (0.013) [0.015]	
Variance (14), (10)	0.564 (0.136) [0.151]	1.435 (0.990) [0.814]	0.001 (0.001) [0.009]	0.893 (0.014) [0.131]	0.103 (0.013) [0.021]	0.000 (0.019) [0.127]	0.090 (0.013) [0.015]	
<i>(C) January 1953 to December 1984, T = 8,043</i>								
Std. dev. (13), (10)	-0.297 (0.257) [0.588]	0.101 (0.042) [0.082]	0.001 (0.000) [0.002]	0.907 (0.007) [0.063]	0.086 (0.011) [0.018]	0.000 (0.008) [0.070]	0.213 (0.012) [0.015]	
Variance (14), (10)	0.057 (0.128) [0.300]	6.286 (2.584) [5.328]	0.001 (0.000) [0.002]	0.906 (0.007) [0.066]	0.086 (0.011) [0.019]	0.000 (0.009) [0.068]	0.212 (0.012) [0.016]	
<i>(D) January 1985 to December 2019, T = 8,822</i>								
Std. dev. (13), (10)	-0.257 (0.274) [0.703]	0.104 (0.034) [0.083]	0.002 (0.001) [0.007]	0.878 (0.013) [0.106]	0.093 (0.011) [0.032]	0.014 (0.017) [0.098]	-0.014 (0.012) [0.011]	
Variance (14), (10)	0.309 (0.116) [0.177]	3.580 (1.243) [1.867]	0.002 (0.001) [0.008]	0.878 (0.014) [0.122]	0.094 (0.011) [0.033]	0.012 (0.017) [0.114]	-0.014 (0.012) [0.011]	

^a $R_{mt} - R_{ft}$ is the daily excess holding period return to the Standard & Poor's composite portfolio (the percentage price change minus the risk-free interest rate). Nonlinear optimization techniques are used to calculate maximum likelihood estimates. Asymptotic standard errors are in parentheses under the coefficient estimates. The numbers in brackets are robust standard errors calculated using the method of White (1982).

Table 7: Comparison of ARIMA with GARCH Predictions of Stock Market Volatility and Their Relations to Monthly CRSP Value-Weighted Excess Holding Period Returns. GARCH-in-Mean Estimates Using Monthly Excess Holding Period Returns.^a

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t + \varepsilon_t - \theta\varepsilon_{t-1} \quad (13)$$

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t^2 + \varepsilon_t - \theta\varepsilon_{t-1} \quad (14)$$

$$\sigma_t^2 = a + b\sigma_{t-1}^2 + c_1\varepsilon_{t-1}^2 + c_2\varepsilon_{t-2}^2 \quad (10)$$

Description: This Table summarizes the results from GARCH-in-mean regressions with monthly data.

Interpretation: The relation between predicted volatility and excess returns is positive.

GARCH-in-mean equations	α	β	$a \times 10^3$	b	c_1	c_2	θ	Nyblom test for stability
<i>(A) January 1928 to December 1984, T = 684</i>								
Std. dev. (13), (10)	-0.0019 (0.0059) [0.0064]	0.215 (0.132) [0.145]	0.082 (0.034) [0.036]	0.811 (0.033) [0.040]	0.059 (0.041) [0.038]	0.109 (0.058) [0.058]	0.070 (0.038) [0.040]	1.5
Variance (14), (10)	0.0041 (0.0026) [0.0027]	1.618 (0.964) [1.019]	0.085 (0.034) [0.038]	0.809 (0.034) [0.041]	0.062 (0.041) [0.038]	0.106 (0.059) [0.058]	0.069 (0.038) [0.040]	1.4
<i>(B) January 1928 to December 1952, T = 300</i>								
Std. dev. (13), (10)	0.0110 (0.0082) [0.0073]	0.002 (0.159) [0.148]	0.062 (0.056) [0.055]	0.841 (0.037) [0.046]	0.134 (0.082) [0.070]	0.022 (0.094) [0.088]	0.077 (0.060) [0.077]	
Variance (14), (10)	0.0097 (0.0041) [0.0038]	0.517 (1.012) [0.902]	0.064 (0.057) [0.056]	0.840 (0.039) [0.047]	0.139 (0.086) [0.073]	0.018 (0.098) [0.093]	0.079 (0.060) [0.075]	
<i>(C) January 1953 to December 1984, T = 384</i>								
Std. dev. (13), (10)	-0.0132 (0.0099) [0.0103]	0.481 (0.252) [0.248]	0.165 (0.090) [0.079]	0.737 (0.076) [0.056]	0.000 (0.050) [0.056]	0.176 (0.080) [0.074]	0.054 (0.048) [0.044]	
Variance (14), (10)	-0.0049 (0.0055) [0.0065]	6.601 (3.329) [3.565]	0.163 (0.087) [0.078]	0.746 (0.073) [0.052]	0.000 (0.049) [0.054]	0.166 (0.077) [0.074]	0.052 (0.048) [0.044]	
<i>(D) January 1985 to December 2019, T = 420</i>								
Std. dev. (13), (10)	0.0030 (0.0060) [0.0060]	0.125 (0.168) [0.176]	0.067 (0.037) [0.052]	0.823 (0.047) [0.070]	0.152 (0.063) [0.077]	0.000 (0.075) [0.084]	-0.020 (0.057) [0.071]	
Variance (14), (10)	0.0053 (0.0029) [0.0029]	1.540 (1.800) [1.808]	0.067 (0.036) [0.051]	0.824 (0.046) [0.069]	0.152 (0.063) [0.077]	0.000 (0.074) [0.080]	-0.019 (0.057) [0.070]	

^a The statistical procedure used in Table 7 is the same as in Table 6, except that monthly excess holding period returns to the CRSP value-weighted portfolio are used instead of the daily excess holding period returns to the S&P composite portfolio. Asymptotic standard errors are in parentheses under the coefficient estimates. The numbers in brackets are robust standard errors calculated using the method of White (1982).

Table 8: Comparison of ARIMA with GARCH Predictions of Stock Market Volatility and Their Relations to Monthly CRSP Value-Weighted Excess Holding Period Returns. Weighted Least Squares Regressions of Monthly CRSP Value-Weighted Excess Holding Period Returns Against the Predicted Standard Deviation or Variance of Stock Returns from the Monthly GARCH-in-Mean Model.^a

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_t + \varepsilon_t \quad (15)$$

$$R_{mt} - R_{ft} = \alpha + \beta\sigma_{mt}^2 + \varepsilon_t \quad (16)$$

Description: This Table summarizes the results from regressions (15) and (16).

Interpretation: The relation between predicted volatility and risk premiums is unreliable.

Volatility measure	α	β	$S(\varepsilon)$	R^2	$Q(12)$	$SR(\varepsilon)$
<i>(A) January 1928 to December 1984, T = 684</i>						
Monthly GARCH Std. dev.	0.0034 (0.0050) [0.0063]	0.049 (0.103) [0.145]	0.0586	0.0003	21.4	8.63
Monthly GARCH Variance	0.0048 (0.0024) [0.0024]	0.341 (0.957) [0.940]	0.0585	0.0002	20.8	6.66
<i>(B) January 1928 to December 1952, T = 300</i>						
Monthly GARCH Std. dev.	0.0127 (0.0083) [0.0090]	-0.076 (0.135) [0.169]	0.0755	0.0011	16.1	7.73
Monthly GARCH Variance	0.0120 (0.0041) [0.0040]	-0.750 (1.015) [0.982]	0.0758	0.0018	13.6	6.63
<i>(C) January 1953 to December 1984, T = 384</i>						
Monthly GARCH Std. dev.	-0.0170 (0.0100) [0.0098]	0.535 (0.251) [0.253]	0.0407	0.0117	17.8	6.08
Monthly GARCH Variance	-0.0055 (0.0051) [0.0046]	5.951 (3.194) [2.826]	0.0407	0.0090	16.9	5.90
<i>(D) January 1985 to December 2019, T = 420</i>						
Monthly GARCH Std. dev.	0.0007 (0.0066) [0.0062]	0.148 (0.168) [0.167]	0.0412	0.0018	5.8	8.34
Monthly GARCH Variance	0.0043 (0.0033) [0.0029]	1.337 (2.109) [1.794]	0.0412	0.0010	5.0	8.09

^a The statistical procedure used in this Table is the same as in Table 4, except that the predicted standard deviation of the CRSP value-weighted return, σ_t , estimated in Table 7, is used to standardize each observation, instead of the prediction $\hat{\sigma}_{mt}$ from the ARIMA model in Table 1, Panel C. See the footnotes to Table 4 and Table 6 for more detailed descriptions of the statistical procedures.

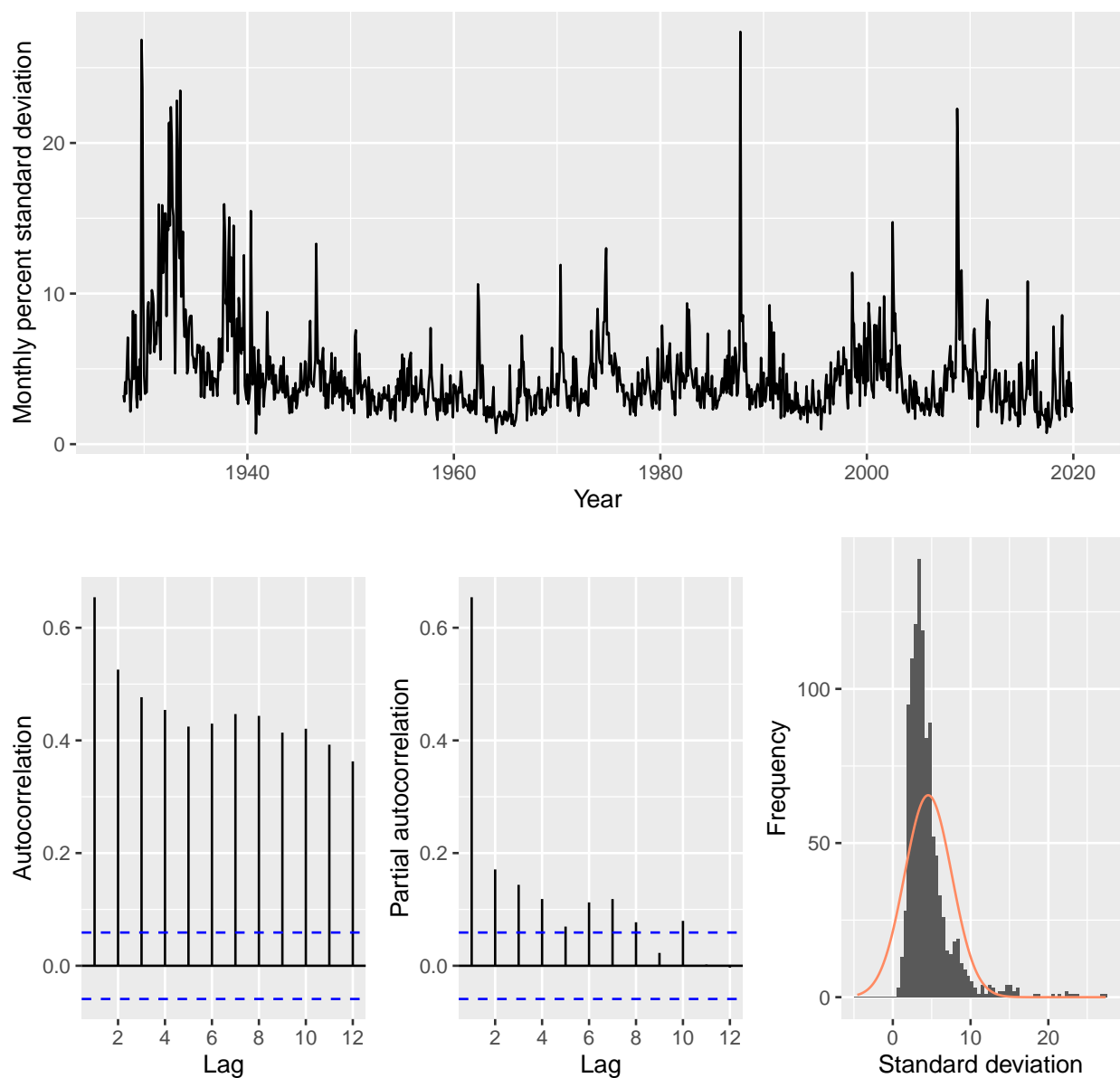


Figure 1: Monthly percent standard deviations of the returns to the Standard & Poor's composite portfolio, σ_{mt} , estimated from returns for days i within the month t , r_{it} , 1928–2019.

$$\sigma_{mt} = \left\{ \sum_{i=1}^{N_t} r_{it}^2 + 2 \sum_{i=1}^{N_t-1} r_{it} r_{i+1,t} \right\}^{1/2} \quad (2)$$

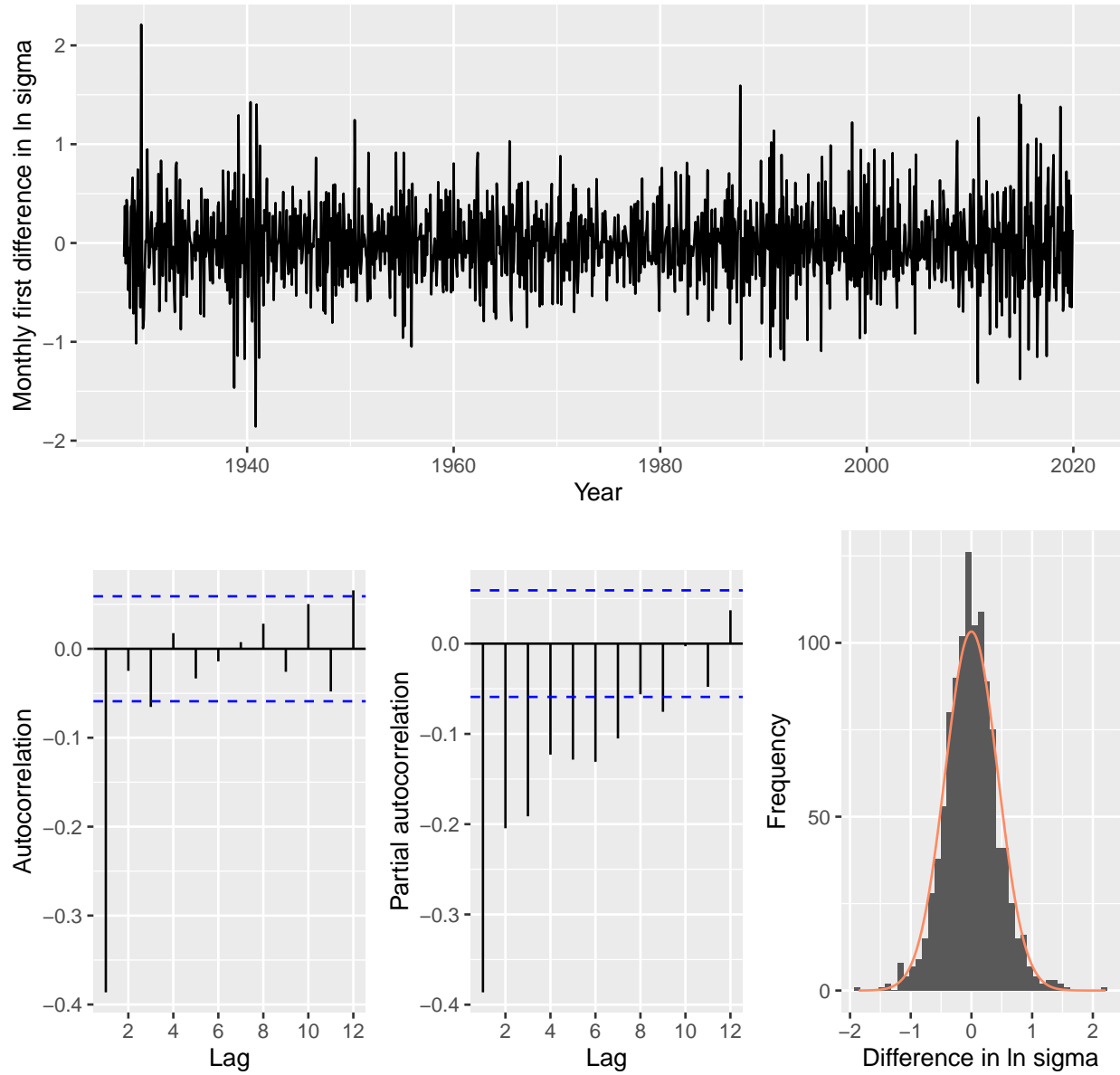


Figure 2: Monthly first difference in $\ln \sigma_{mt}$ of the returns to the Standard & Poor's composite portfolio, estimated from returns for days i within the month t , r_{it} , 1928–2019.

$$\sigma_{mt} = \left\{ \sum_{i=1}^{N_t} r_{it}^2 + 2 \sum_{i=1}^{N_t-1} r_{it} r_{i+1,t} \right\}^{1/2} \quad (2)$$

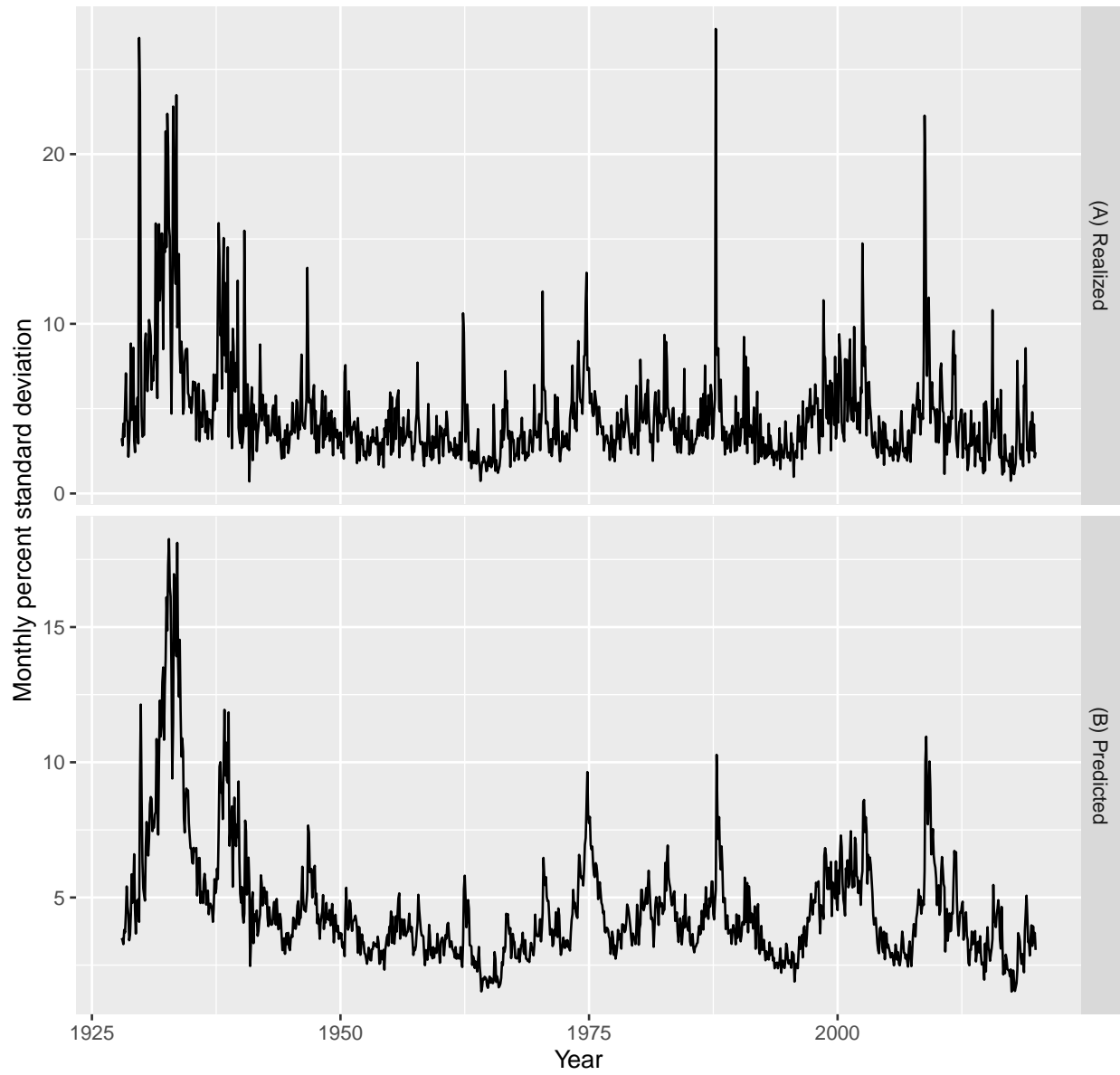


Figure 3: (A) realized, σ_{mt} , and (B) predicted, $\hat{\sigma}_{mt}$, monthly percent standard deviations of the returns to the Standard & Poor's composite portfolio estimated from the ARIMA model in Table 1, Panel C, 1928–2019.

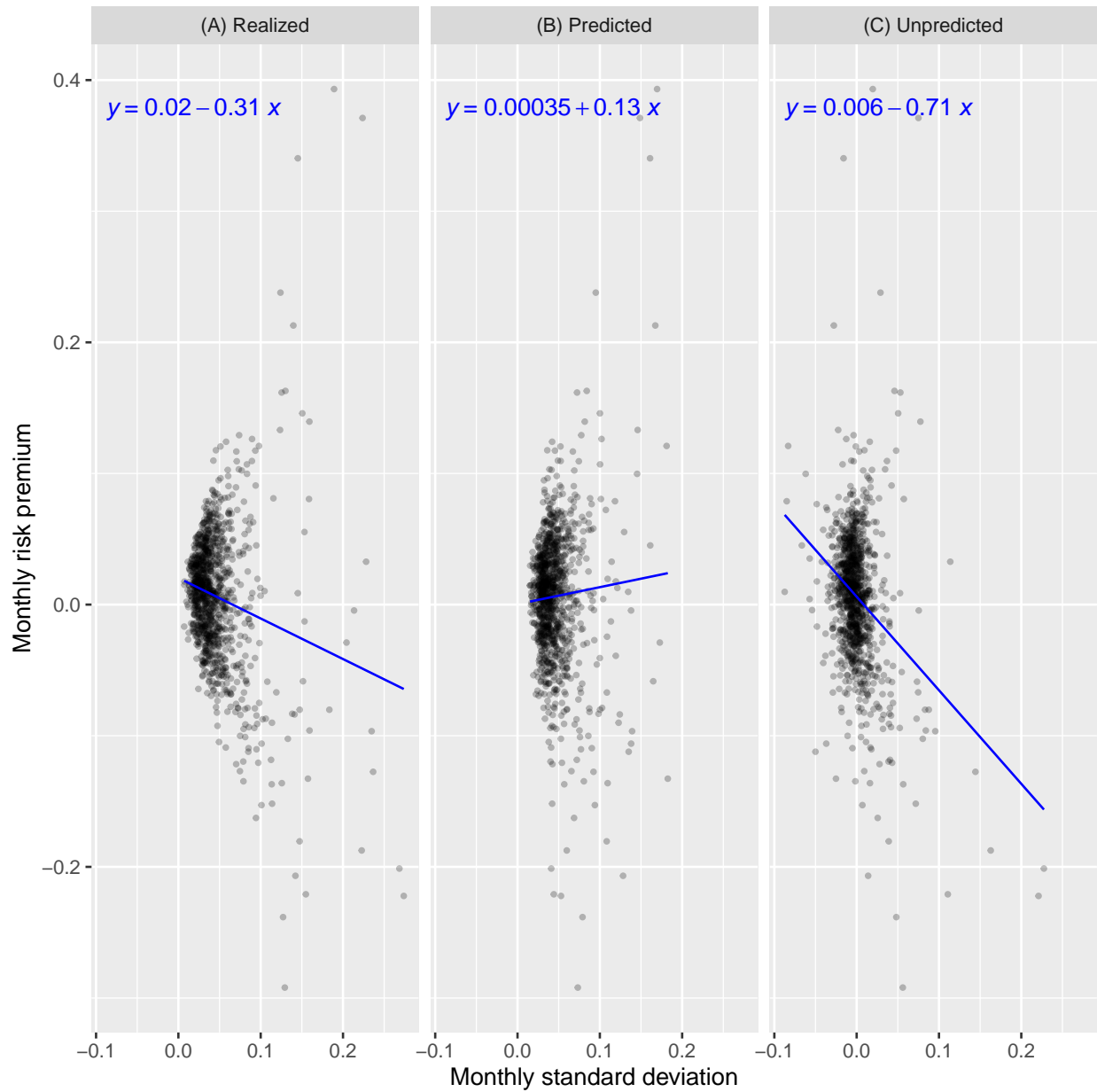


Figure 4: Scatterplots and univariate regressions of monthly risk premiums $R_{mt} - R_{ft}$ against monthly (A) realized, (B) predicted, and (C) unpredicted standard deviation. Unpredicted standard deviation is the difference between realized standard deviation from (2) and predicted standard deviation from (5), 1928–2019.

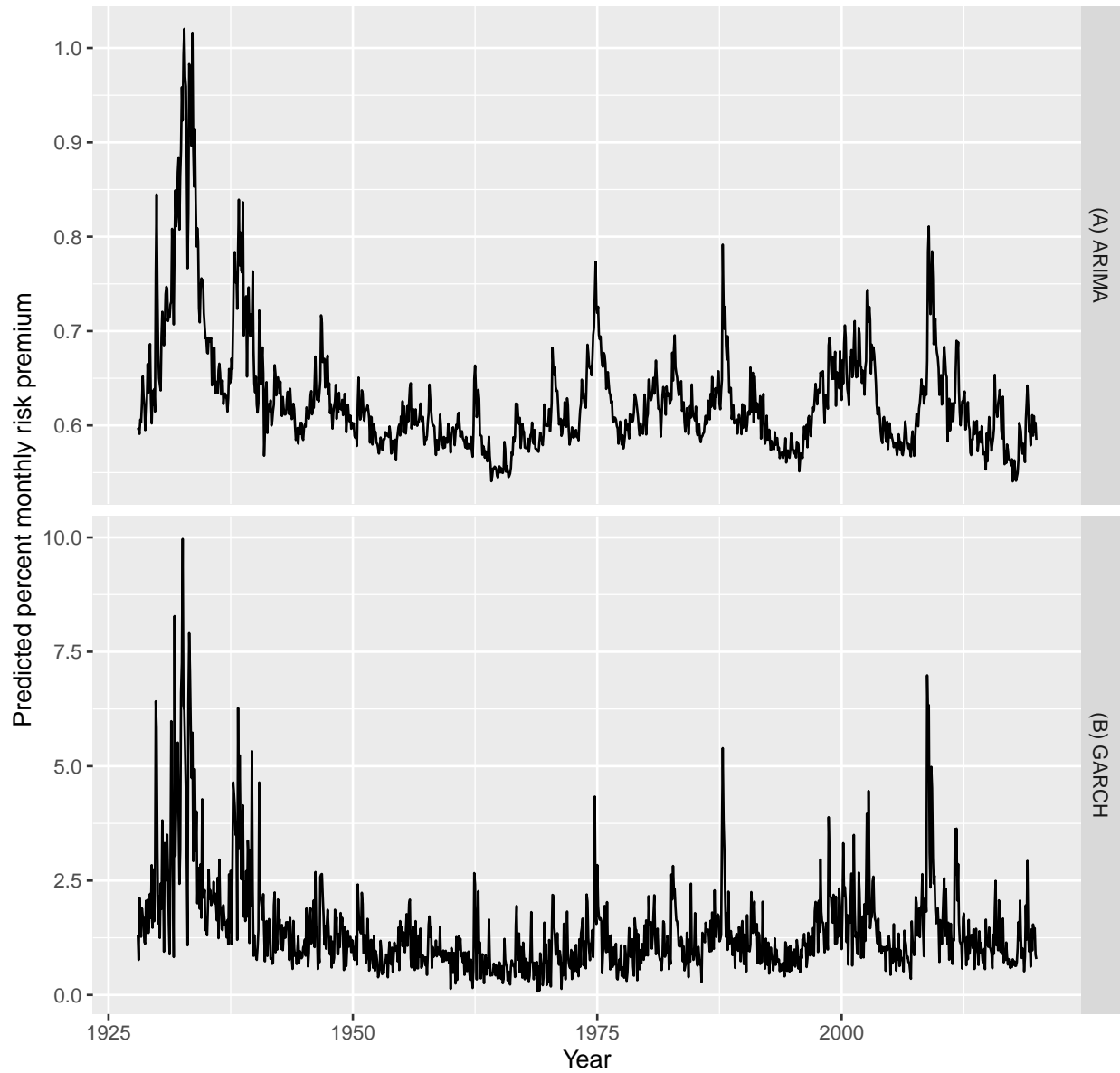


Figure 5: Predicted percent monthly risk premium to the Standard & Poor's composite portfolio from the regression on ARIMA predictions of the standard deviation, $\hat{\sigma}_{mt}$, in Table 4, and from the daily GARCH-in-mean model for the standard deviation, σ_t , in Table 6, 1928–2019.

A Appendix

Why do Other Studies Cite FSS and Which Parts of FSS Receive the Most Attention?

According to the Web of Science (WOS) database, FSS is cited in 1,189 studies published through June 2020. In this appendix, we extend our replication by text mining these studies for information on how others cite FSS and which parts of FSS receive the most attention.

Ready for use on WOS are electronic versions of the titles and abstracts for 1,189 published studies citing FSS through June 2020. Some have no abstract, so there are 1,189 titles and 1,089 abstracts available to read electronically. As with raw numeric data, raw textual data must also be cleaned before it can be read and analyzed. Starting with the raw textual data from WOS, we convert each word to lowercase so the same word with different capitalization counts the same. We then remove numbers and punctuation except for hyphens and apostrophes. We then remove words like *the*, *and*, *or*, and *is*. Known as stop words, these words are generally not informative about the meaning and content of textual data.¹⁵ After making these changes, there are 8,211 words in the titles and 82,140 words in the abstracts usable for text analysis.

With the clean textual data, we begin by counting the number of times each word is used in the titles and abstracts. This type of word frequency analysis helps us to discover the main focus of the studies that cite FSS.

Figure 6 is a plot of the top-10 words in the (A) titles and (B) abstracts of the citing study texts.

¹⁵The complete list of stop words is given in the online appendix for Lewis, Yang, Rose, and Li (2004). The current Web address is <http://www.ai.mit.edu/projects/jmlr/papers/volume5/lewis04a/a11-smart-stop-list/english.stop>. We also remove the word *elsevier*, which appears in many of the abstracts as part of the copyright notice.

[Figure 6 here]

We find that the most frequent word in the titles and abstracts is *volatility*. *Volatility* is used 443 times in the titles and 2,288 times in the abstracts. The top-five words in each of the titles and abstracts are *volatility*, *market*, *returns*, *stock*, and *risk*. These are the main focus of the studies that cite FSS.

To better appreciate the broader context in which the citing studies focus on volatility, we count every occurrence of *volatility* along with one word used immediately before and after it. Table 9 reports the results.

[Table 9 here]

We find that the most frequently used words before *volatility* include *market* and *return*. *Market* is used before *volatility* 42 times in the titles and 145 times in the abstracts, while *return* is used before *volatility* 25 times in the titles and 117 times in the abstracts. Other frequently used words before *volatility* include *realized*, *conditional*, and *stochastic*.

Closer examination confirms that studies cite FSS for realized volatility and for the stochastic properties of volatility. Studies citing FSS for realized volatility include Schwert (1989b), Bushee and Noe (2000), Campbell, Lettau, Malkiel, and Xu (2001), Alizadeh, Brandt, and Diebold (2002), Antweiler and Frank (2004), Jiang and Tian (2005), Bandi and Russell (2006), Bekaert, Harvey, and Lundblad (2007), and Jondeau, Zhang, and Zhu (2019). Andersen, Bollerslev, Diebold, and Labys (2001a) cite FSS when they construct new model-free estimates of daily exchange rate volatility, while formal justification for the realized volatility calculation in FSS is given in Andersen, Bollerslev, Diebold, and Ebens (2001b).

Studies also cite FSS for the stochastic properties of volatility. Those who cite FSS for heteroskedasticity include Fama and French (1988), Lo and MacKinlay (1988), Nelson (1990), Hsieh (1991), Hodrick (1992), McQueen and Roley (1993), Saunders (1993), Flem-

ing, Ostdiek, and Whaley (1996), Kandel and Stambaugh (1996), Edmans, García, and Norli (2007), and Liu (2007). Those who cite FSS for volatility persistence and predictability include Schwert (1989a), Bollerslev, Chou, and Kroner (1992), Easley and O'Hara (1992), Cai (1994), Christensen and Prabhala (1998), and Andersen, Bollerslev, Diebold, and Labys (2003).

The most frequently used words after *volatility* include *feedback* and *models*, suggesting that studies cite FSS for volatility feedback and to motivate conditional volatility models. Closer examination confirms this result.

Those who cite FSS for the theory of volatility feedback include Veronesi (1999), Chen, Hong, and Stein (2001), Hong and Stein (2003), Bansal et al. (2004), Hutton, Marcus, and Tehranian (2009), Kothari, Li, and Short (2009), Bibinger, Neely, and Winkelmann (2019), and Pyun (2019). Those who cite FSS for evidence on volatility feedback include Fama (1990), Chen (1991), Bessembinder and Seguin (1993), Haugen and Baker (1996), Allen and Karjalainen (1999), Harvey and Siddique (1999), Ang and Chen (2002), Lee, Jiang, and Indro (2002), Bakshi and Kapadia (2003), Lewellen and Nagel (2006), Drienko, Smith, and Reibnitz (2019), Jang and Kang (2019), and Kilic and Shaliastovich (2019). Those who compare their results on volatility feedback to the results in FSS include Harvey (1989), Baillie and DeGennaro (1990), Shanken (1990), Brock, Lakonishok, and LeBaron (1992), Goyal and Santa-Clara (2003), Ludvigson and Ng (2007), Kottimukkalur (2019), and Moreira et al. (2019).

Studies cite FSS for volatility models and for volatility modeling for two main reasons. The first is to justify volatility as an explanatory variable in their models. Among those who cite FSS to justify volatility as a variable in their models are Cutler, Poterba, and Summers (1989), Ferson et al. (1993), Andrei, Hasler, and Jeanneret (2019), and Jiang, Lee, Martin, and Zhou (2019). The second is to motivate conditional volatility models like ARCH and GARCH. Among those who cite FSS to motivate ARCH and GARCH models

are Hamao, Masulis, and Ng (1990), Lamoureux and Lastrapes (1990), Pagan and Schwert (1990), Chan, Chan, and Karolyi (1991), Nelson (1991), Tauchen and Hussey (1991), Nelson and Cao (1992), and Fu (2009). Others cite FSS when they extend and improve on ARCH and GARCH. These include Engle, Ng, and Rothschild (1990), Campbell et al. (1992), Hentschel (1995), and Engle (2002). Others compare the results from their conditional volatility models to those in FSS including Engle and González-Rivera (1991), Ghysels, Santa-Clara, and Valkanov (2005), and Yu and Yuan (2011).

Given that FSS is about “expected stock returns” and “volatility,” it would be interesting to know if the frequency of *volatility* and *return* in the titles and abstracts has evolved through time. To see how the use of these terms has evolved, we line up the titles and abstracts by publication date and divide them into 10 equally numbered groups. Because there are 1,189 titles and 1,089 abstracts, equally numbered means that there are 119 titles in the first nine groups and 118 in the 10th group and 109 abstracts in the first nine groups and 108 in the 10th group. For completeness, we count *risk* and *variance* along with *volatility* and *returns* along with *return*.

Figure 7 is a plot of the relative frequency of *volatility* and *return* by publication date group.

[Figure 7 here]

We find that studies cite FSS more for volatility than for returns, though there was not as much focus on volatility immediately after publication as there later turned out to be. There is a noticeable increase in the frequency of *volatility* in the first few groups of citing studies. The frequency of *volatility* in the titles increases from 55 in the first group to 70 in the fourth group. The frequency of *volatility* in the abstracts increases from 244 in the first group to 444 in the third group.

Table 9: Top-10 Words Before and After *volatility*
Description: This Table gives the top-10 most frequently used words before and after *volatility* in 1,189 citing study titles (A) and 1,089 citing study abstracts (B) through June 2020.
Interpretation: The most frequently used words before and after *volatility* indicate the broader context in which the citing studies focus on the topic of volatility.

(A) Titles				
Before	<i>N</i>	Keyword	After	<i>N</i>
market	42	<i>volatility</i>	evidence	15
stochastic	26		models	14
return	25		spillovers	13
implied	23		index	12
asymmetric	18		model	10
idiosyncratic	14		risk	10
returns	13		stock	10
price	12		feedback	7
realized	11		empirical	6
conditional	10		persistence	6

(B) Abstracts				
Before	<i>N</i>	Keyword	After	<i>N</i>
market	145	<i>volatility</i>	index	61
return	117		feedback	60
implied	91		risk	57
stochastic	90		model	45
realized	79		models	43
conditional	75		stock	41
asymmetric	68		persistence	36
idiosyncratic	66		spillovers	30
returns	64		spillover	26
price	44		process	24

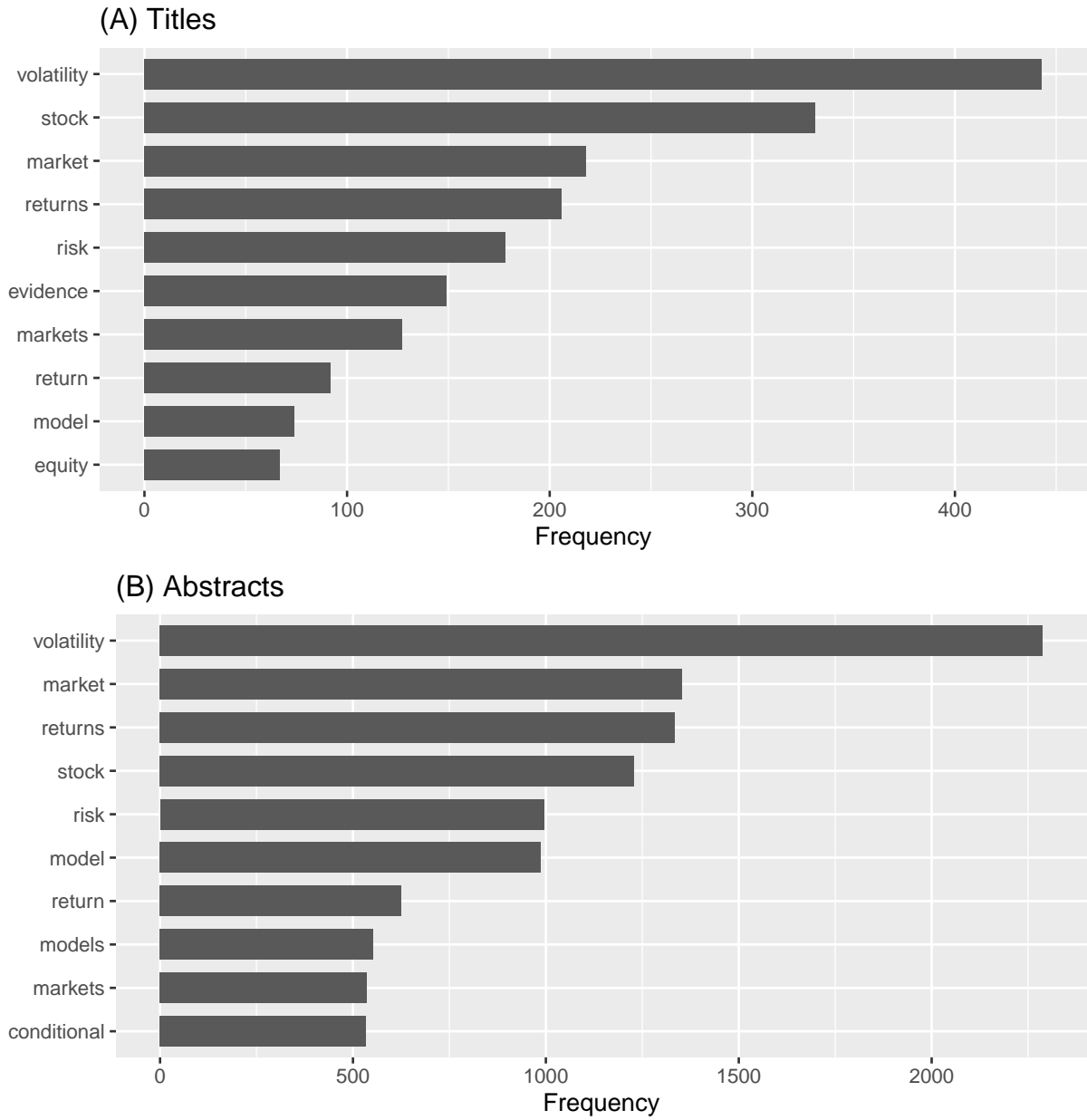


Figure 6: Frequency of the top-10 words used in 1,189 citing study (A) titles and 1,089 citing study (B) abstracts through June 2020.

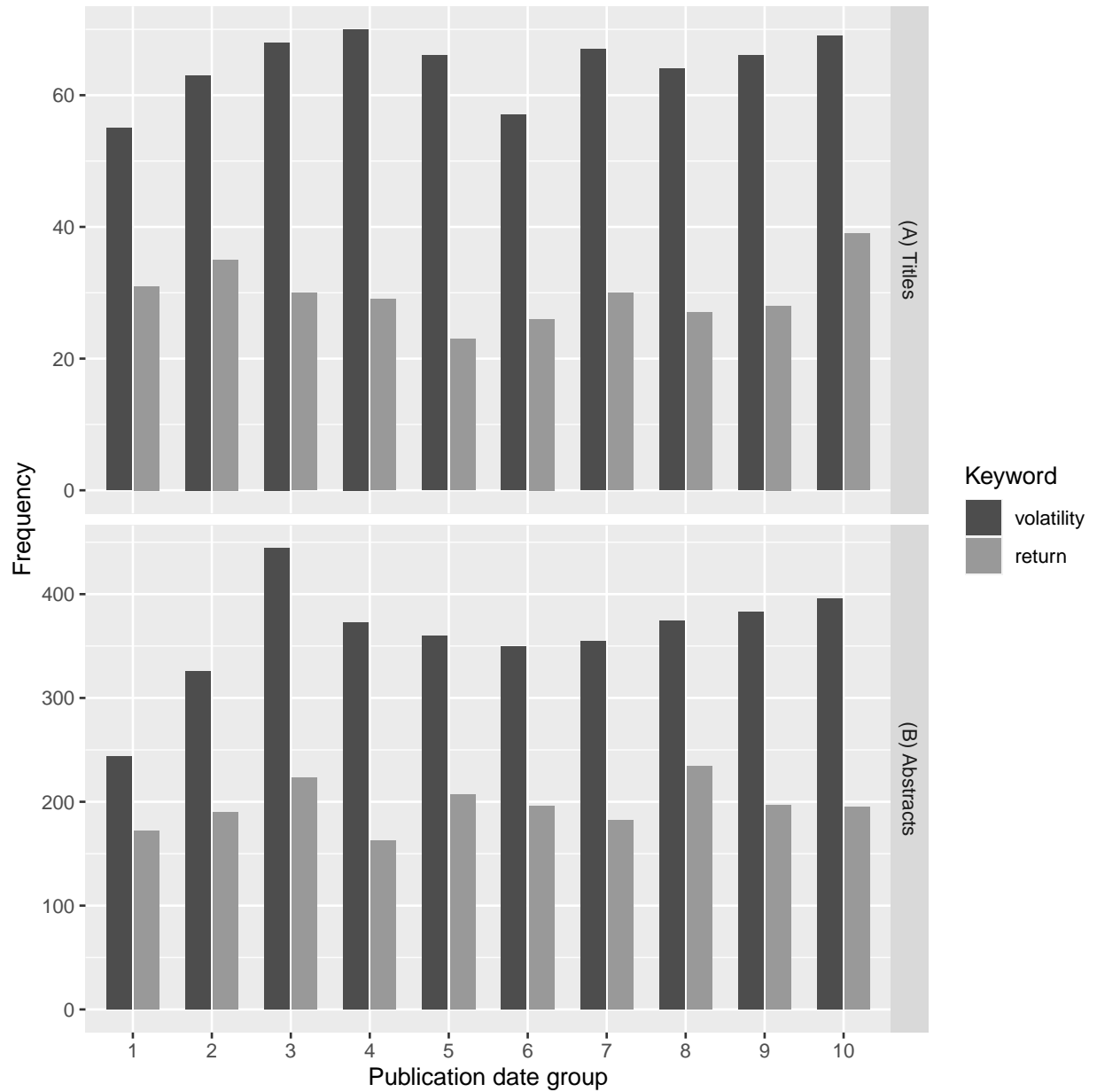


Figure 7: Frequency of *volatility* and *return* by publication date group in 1,189 citing study (A) titles and 1,089 citing study (B) abstracts through June 2020. We form the publication date groups by lining up the titles and abstracts by publication date and then dividing them into 10 approximately equally numbered groups.

B Appendix

R Computer Code

Below are the major segments of the R computer code we use for the replication. Daily data are held as `xts` objects from the `xts` package and monthly data are held as `ts` objects. The ARIMA functions are from the `forecast` package and the GARCH functions are from the `rugarch` package. Linear models with `ts` objects are estimated using the `dynlm` package to preserve time series attributes.

```
#Calculate realized monthly market volatility from daily returns

equation_2 <- function(r) {
  sum(r^2) + 2 * sum(r * lag(r, k = -1), na.rm = TRUE)
}

#Calculate ARIMA model and predicted volatility by back-transforming

equation_3 <- function(log_sigma) {
  Arima(log_sigma, order = c(0, 1, 3), include.constant = TRUE, lambda = NULL)
}

regression_3 <- lapply(log_realized_sigma, equation_3)

fitted_values_3 <- lapply(regression_3, fitted)

residuals_3 <- lapply(regression_3, residuals)

equation_4a <- function(fitted_values, residuals) {
  exp(fitted_values + 0.5 * var(residuals))
}

equation_4b <- function(fitted_values, residuals) {
  exp(2 * fitted_values + 2 * var(residuals))
}

predicted_sigma_arima <- mapply(equation_4a, fitted_values_3, residuals_3)
```

```
predicted_sigma_squared_arma <- mapply(equation_4b, fitted_values_3, residuals_3)
```

```
#GARCH model and GARCH-in-mean models
```

```
equation_5c_5e <- function(excess_returns) {  
  ugarchfit(  
    spec = ugarchspec(  
      mean.model = list(armaOrder = c(0, 1)),  
      variance.model = list(model = "sGARCH", garchOrder = c(2, 1))  
    ),  
    data = excess_returns  
  )  
}
```

```
equation_8a_5e <- function(excess_returns) {  
  ugarchfit(  
    spec = ugarchspec(  
      mean.model = list(armaOrder = c(0, 1), archm = TRUE, archpow = 1),  
      variance.model = list(model = "sGARCH", garchOrder = c(2, 1))  
    ),  
    data = excess_returns  
  )  
}
```

```
equation_8b_5e <- function(excess_returns) {  
  ugarchfit(  
    spec = ugarchspec(  
      mean.model = list(armaOrder = c(0, 1), archm = TRUE, archpow = 2),  
      variance.model = list(model = "sGARCH", garchOrder = c(2, 1))  
    ),  
    data = excess_returns  
  )  
}
```

```
#Weighted least squares
```

```
equation_mean_wls <- function(excess_returns, sigma) {  
  dynlm(excess_returns ~ 1, weights = 1 / sigma)  
}
```

```
#Linear models
```

```
equation_6 <- function(excess_returns, predicted_sigma_p) {  
  dynlm(excess_returns ~ predicted_sigma_p, weights = 1 / predicted_sigma_p)  
}
```

```

equation_7 <- function(excess_returns, predicted_sigma_p, realized_sigma_p) {
  dynlm(excess_returns ~ predicted_sigma_p +
        I(realized_sigma_p - predicted_sigma_p),
        weights = 1 / predicted_sigma_p)
}

equation_10 <- function(excess_returns, predicted_sigma_p) {
  dynlm(excess_returns ~ predicted_sigma_p, weights = 1 / predicted_sigma_p)
}

#Automatic ARIMA model selection

equation_3a <- function(sigma) {
  auto.arima(sigma, ic = "aicc", stepwise = FALSE,
             approximation = FALSE, lambda = "auto")
}

```