Journal of Financial Economics 96 (2010) 364–380

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Average correlation and stock market returns ☆

Joshua M. Pollet [a,*], Mungo Wilson [b]

[a] Department of Finance, Michigan State University, East Lansing, MI 48824, USA
[b] Saïd Business School, University of Oxford, Oxford OX1 1HP, UK

ARTICLE INFO

Article history:
Received 24 October 2008
Received in revised form
11 November 2008
Accepted 23 November 2008
Available online 24 February 2010

JEL classification:
G12

Keywords:
Correlation
Roll critique

ABSTRACT

If the Roll critique is important, changes in the variance of the stock market may be only weakly related to changes in aggregate risk and subsequent stock market excess returns. However, since individual stock returns share a common sensitivity to true market return shocks, higher aggregate risk can be revealed by higher correlation between stocks. In addition, a change in stock market variance that leaves aggregate risk unchanged can have a zero or even negative effect on the stock market risk premium. We show that the average correlation between daily stock returns predicts subsequent quarterly stock market excess returns. We also show that changes in stock market risk holding average correlation constant can be interpreted as changes in the average variance of individual stocks. Such changes have a negative relation with future stock market excess returns.

1. Introduction

The trade-off between risk and expected return is at the center of any equilibrium theory of finance. From an aggregate perspective, as systematic risk increases, risk-averse investors require a higher risk premium to hold aggregate wealth, and the equilibrium expected return must rise. The following equation for the excess log return on aggregate wealth is consistent with this intuition:

$$r_{m,t+1}-r_{f,t+1}=\beta_0+\beta_1\,Var_t[r_{m,t+1}]+\lambda x_t+\varepsilon_{t+1}, \tag{1}$$

where β_1 is positive, $Var_t[r_{m,t+1}]$ is the conditional variance of the return for aggregate wealth, x_t represents other potential sources of variation in expected returns, and $E_t[\varepsilon_{t+1}]=0$. It is often assumed that the return on aggregate wealth satisfies a variance-in-mean relationship where λ is zero in Eq. (1).

However, a large literature, including Campbell (1987), French, Schwert, and Stambaugh (1987), Glosten, Jagannathan, and Runkle (1993), and Harvey (2001), among many other studies, has attempted to identify a positive variance-in-mean relationship for *stock market* returns in the data with only limited success.[1] French et al. find a positive relationship in monthly data using a generalized autoregressive conditional heteroskedasticity (GARCH) framework, but their estimates are only marginally statistically significant. Often the evidence indicates that greater risk is actually associated with a lower expected excess return. For instance, Campbell (1987) utilizes a

☆ We thank Klaas Baks, Michael Brennan, John Campbell, Timothy Johnson, Jun Pan, Neil Pearson, Allen Poteshman, Bruno Solnik, Jiang Wang, Joshua White, and an anonymous referee as well as seminar participants at the HKUST finance symposium, the 2007 Midwest Finance Association Annual Meeting, MIT Sloan, and UIUC for their comments. Wilson acknowledges the assistance of the Research Grants Council of Hong Kong Grant no. HKUST6478/06H.

* Corresponding author.
E-mail address: pollet@msu.edu (J.M. Pollet).

[1] See Campbell, Lo, and MacKinlay (1997, Chapter 12) for a survey, and for more recent references see Harvey (2001). Ang and Liu (2007) show that plausible assumptions about the dynamics of expected returns and variance can generate a negative relation between these two variables. Campbell (1993) provides an example in an intertemporal capital asset pricing model (ICAPM) framework where the equilibrium trade-off between expected return and variance could be negative.

0304-405X/$ - see front matter
doi:10.1016/j.jfineco.2010.02.011

generalized method of moments (GMM) approach and finds a significant negative linear relationship.

Since Roll (1977), a seminal critique of empirical studies regarding the validity of the capital asset pricing model (CAPM), researchers have recognized that the return on aggregate wealth is not directly observable. Although the Roll critique is usually stated in the context of cross-sectional tests of the CAPM, it is also relevant for time-series tests of a variance-in-mean relation for the stock market return because the variance of the return for the true aggregate wealth portfolio may be only weakly related to stock market variance.

We find that changes in true aggregate risk may nevertheless reveal themselves through changes in the correlation between observable stock returns. Since the return on the aggregate wealth portfolio is a common component for most asset returns (on average, positively related to individual asset returns), an increase in aggregate risk, other things equal, is associated with an increased tendency of stock prices to move together. As a result, such increases in correlation reveal increases in true aggregate risk. If the stock market risk premium depends positively on aggregate risk, then the average correlation between stock returns should forecast future stock market excess returns.

In addition, measures of stock market variance may themselves be poor forecasts of future stock market returns even in the presence of a strong variance-in-mean relationship for the aggregate wealth portfolio. Indeed, if most changes in stock market variance are due to changes in the variance of components orthogonal to the variance of the return for aggregate wealth (i.e., holding aggregate risk constant), then such changes in stock market variance must be accompanied by offsetting changes in the covariance of the stock market with the rest of the aggregate wealth portfolio.

Our results indicate that changes in the sample variance of daily US stock market excess returns are almost completely captured by changes in the product of the average variance for the largest 500 individual stocks and the average correlation between all pairs of stocks for the largest 500 stocks. Of these two, the average correlation strongly predicts future excess stock market returns while the average variance has no discernible forecasting power. Thus, the weak ability of stock market variance to forecast stock market expected returns is due to the contamination of average correlation by average variance. Average correlation also forecasts stock market returns with an out-of-sample R^2 of more than 3%, and therefore passes the test proposed by Goyal and Welch (2008).

The large negative covariance between shocks to average correlation and shocks to realized returns may explain the phenomenon of asymmetric correlation (see, for example, Longin and Solnik, 2001; Ang and Chen, 2002; Hong, Tu, and Zhou, 2007) in which stock returns are more highly correlated when the stock market declines. If an increase in average correlation is due to an increase in aggregate risk, then the discount rate for future expected cash flows on most assets, including the stock market, should increase as well. Holding expected future cash flows constant, higher future expected returns induce an immediate fall in price, i.e., negative realized returns, or volatility feedback, as described by Campbell and Hentschel (1992). So, if stocks become more highly correlated with each other, volatility feedback generates negative stock returns on average and this relation leads to asymmetric correlation. Therefore, our findings suggest that asymmetric correlation is caused by a form of volatility feedback induced by fluctuations in the volatility of the return on aggregate wealth rather than changes in stock market volatility.

We also find a negative, albeit insignificant, relation between average variance and the subsequent market return after controlling for average correlation. In this respect, our findings are similar to those of Bali, Cakici, Yan, and Zhang (2005) and at odds with those of Goyal and Santa-Clara (2003). Bali et al. indicates that the positive forecasting power of average variance, shown in Goyal and Santa-Clara, disappears if Nasdaq and Amex stocks are excluded from market returns or if value-weighted average variance or median variance is used in place of equal-weighted average variance. Since Brown and Kapadia (2007) find that patterns in average individual stock variance are driven by equity issuance patterns, equal-weighted average variance will be affected by the fraction of companies that are recent issues.

Ghysels, Santa-Clara, and Valkanov (2005) appear to identify a variance-in-mean relationship using mixed data sampling to estimate conditional stock market variance. Their paper claims that mixed data sampling (MIDAS) provides more accurate estimates than conventional methods. However, the conventional estimator for average variance has a substantially higher R^2 compared to the MIDAS estimator for predicting realized stock market variance. Yet, in spite of this superior predictive power for stock market variance, our results indicate that average variance does not predict subsequent excess returns. If the Roll critique is correct, then an analysis of a linear relationship between the expected return and conditional variance for the stock market is flawed. Hence, we focus our analysis on the differing roles of the two components of stock market variance—average correlation and average variance.

Driessen, Maenhout, and Vilkov (2009) show that the high risk premium earned writing stock index options can be reconciled with the low (or zero) risk premium earned writing individual stock options because the price of correlation risk is much larger (in magnitude) than the price of individual stock volatility risk. This finding is consistent with our results regarding the predictability of stock market returns. Indeed, our explanation for the relationships between average correlation, average variance, and stock market returns also explains why correlation risk is priced and individual stock variance risk is not priced in options markets.

The next section presents a stylized model in which correlation between stock returns is positively related to the stock market risk premium. We also show that the average variance of stock returns can be unrelated or even negatively related to the stock market risk premium. In the same situation, average variance is likely to be negatively related to the covariance of stock market returns with the return of the non-stock component of aggregate wealth. We present the main empirical findings in Section 3. Section 4 analyzes the predictability of returns for different horizons as well as out-of-sample tests of predictability. Section 5 presents supporting evidence that average variance (average correlation) is negatively (positively) related to the covariance of bond and stock market returns and the covariance of labor income growth and stock market returns. Section 6 concludes.

2. The importance of correlation

If the portfolio of assets whose returns are observable (e.g., stocks) is the same as aggregate wealth, then aggregate risk is identical to stock market risk. Hence, the variance of the return to the market portfolio is directly observable and, under typical assumptions (e.g., Merton, 1973), the stock market risk premium is related to this observable variable. However, when the stock market portfolio is only a subset of aggregate wealth, the inability to observe the variance of the return on aggregate wealth interferes with the empirical analysis of the aggregate risk-return relationship. In certain circumstances, the correlation between stocks rather than stock market variance may better reveal changes in true aggregate risk, and hence, the stock market risk premium. We present a model where the logarithmic version of the CAPM pricing equation holds and the correlation between stocks is more strongly related to the stock market risk premium than stock market variance.

2.1. Average correlation and the Roll critique

Campbell and Viceira (2002, Chapter 2) assume that asset returns are conditionally (jointly) lognormally distributed and derive the optimal portfolio weights for investors with power utility over end-of-period wealth:

$$E_t\left[r_{i,t+1}\right]-r_{f,t+1}+\frac{\sigma_{i,t}^2}{2}=\gamma\sum_{j=1}^{N}w_{j,t}^{*}\sigma_{ij,t}. \tag{2}$$

In this equation, $w_{j,t}^{*}$ is the optimal (myopic) weight of asset j in the portfolio, $r_{i,t+1}=\log(1+R_{i,t+1})$ is the log return for asset i, $r_{f,t+1}$ is the log return on the riskless asset (known at time t), γ is the investor's coefficient of relative risk aversion, and the time t subscripts indicate that the conditional variances $\sigma_{i,t}^2$ and covariances $\sigma_{ij,t}$ for log returns are possibly time-varying.

We use the approximation in Campbell and Viceira (2002) for the log return of a portfolio in terms of its weights and the log returns for the constituent assets to analyze the asset market equilibrium for a representative agent with such preferences:

$$E_t[r_{i,t+1}]-r_{f,t+1}+\frac{\sigma_{i,t}^2}{2}\approxeq\gamma\sigma_{im,t}. \tag{3}$$

In this equilibrium, the expected excess log return for asset i (adjusted for Jensen's inequality) is proportional to the conditional covariance of the return for asset i with the return for the true market portfolio, and hence, the logarithmic version of the CAPM holds. In particular, the conditional risk premium for the observable stock market (asset s) is given by Eq. (3) where i=s.

We replace the conditional covariance of r_s with r_m with the analogous covariance in levels[2] and then decompose this covariance into the weighted sum of the variance of the stock market and the covariance of the stock market with the remainder of the aggregate wealth portfolio (asset u). Of course, the return for asset u is largely unobservable:

$$\begin{aligned}E_t[r_{s,t+1}]-r_{ft+1}+\frac{\sigma_{s,t}^2}{2}&\\ \approxeq\gamma\, Cov_t(R_{s,t+1},R_{m,t+1})&\\ =\gamma\, Cov_t(R_{s,t+1},w_{s,t}R_{s,t+1}+(1-w_{s,t})R_{u,t+1})&\\ =\gamma(w_{s,t}\, Var_t(R_{s,t+1})+(1-w_{s,t})Cov_t(R_{s,t+1},R_{u,t+1})).&\end{aligned} \tag{4}$$

To analyze the relationship between aggregate risk and correlation between stocks we consider a simple setting. Assume there are N symmetric stocks (where N is large) in the stock market such that $R_{s,t+1}=(1/N)\sum_i R_{i,t+1}$, $\beta_t=Cov_t(R_{i,t+1},R_{m,t+1})/Var_t(R_{m,t+1})$ for all i. Hence, $\lim_{N\to\infty}\sigma_{s,t}^2=\overline{\rho}_t\overline{\sigma}_t^2$ where $\overline{\rho}_t$ (average correlation) is the pairwise correlation between any two stocks and $\overline{\sigma}_t^2$ is the variance of any stock (average variance). We make additional assumptions regarding the structure of risk: shocks that affect only stock returns contain a common stock market component $\overline{\varepsilon}_{z,t+1}$ with variance $\theta_t\sigma_{z,t}^2$ and an orthogonal idiosyncratic component with variance $(1-\theta_t)\sigma_{z,t}^2$. Given these assumptions,

$$R_{i,t+1}=\beta_t R_{m,t+1}+\overline{\varepsilon}_{z,t+1}+\varepsilon_{i,t+1},$$

where $Var_t(\overline{\varepsilon}_{z,t+1}+\varepsilon_{i,t+1})=\sigma_{z,t}^2$, $Cov(\overline{\varepsilon}_{z,t+1},\varepsilon_{i,t+1})=0$, and $Var(\varepsilon_{i,t+1})=(1-\theta_t)\sigma_{z,t}^2$. The aggregate shock has variance $\sigma_{m,t}^2$ and is by definition independent from both components of the stock-specific shocks. Hence, the return for the stock market can be written in terms of the return on aggregate wealth $R_{m,t+1}$ and the common stock market component of the stock-specific shocks $\overline{\varepsilon}_{z,t+1}$:

$$R_{s,t+1}=\beta_t R_{m,t+1}+\overline{\varepsilon}_{z,t+1}.$$

The variance for each stock and pairwise covariance between any two stocks are given below. In the limit as the number of stocks becomes large, this pairwise covariance is the variance of the stock market:

$$\overline{\sigma}_t^2=Var_t(R_{i,t+1})=\beta_t^2\sigma_{m,t}^2+\sigma_{z,t}^2,$$

$$\sigma_{s,t}^2\approxeq\overline{\rho}_t\overline{\sigma}_t^2=Cov_t(R_{i,t+1},R_{j,t+1})=\beta_t^2\sigma_{m,t}^2+\theta_t\sigma_{z,t}^2.$$

Rearranging these two equations,

$$\sigma_{z,t}^2=\left(\frac{1-\overline{\rho}_t}{1-\theta_t}\right)\overline{\sigma}_t^2,$$

$$\sigma_{m,t}^2=\frac{\overline{\sigma}_t^2}{\beta_t^2}\left(\frac{\overline{\rho}_t-\theta_t}{1-\theta_t}\right).$$

Since the return on aggregate wealth is the weighted combination of the return to the stock market (asset s) and the unobserved portfolio (asset u), the return to the

[2] In a continuous-time setting these two quantities are identical.

unobserved portfolio can also be specified in terms of the return on aggregate wealth $R_{m,t+1}$ and the common stock market component of stock-specific shocks $\overline{\varepsilon}_{z,t+1}$:

$$R_{u,t+1}=\left(\frac{1-w_{s,t}\beta_t}{1-w_{s,t}}\right)R_{m,t+1}-\frac{w_{s,t}}{1-w_{s,t}}\overline{\varepsilon}_{z,t+1}.$$

Thus, the covariance between the return of the stock market and the return of the unobserved component of aggregate wealth can be expressed in terms of average correlation $\overline{\rho}_t$ and average variance $\overline{\sigma}_t^2$:

$$Cov_t(R_{s,t+1},R_{u,t+1})=\left(\frac{1-w_{s,t}\beta_t}{1-w_{s,t}}\right)\frac{\overline{\sigma}_t^2}{\beta_t}\left(\frac{\overline{\rho}_t-\theta_t}{1-\theta_t}\right)-\frac{w_{s,t}\theta_t}{1-w_{s,t}}\left(\frac{1-\overline{\rho}_t}{1-\theta_t}\right)\overline{\sigma}_t^2. \quad (5)$$

Substituting this expression into Eq. (4) and using the relation $\sigma_{s,t}^2=\overline{\rho}_t\overline{\sigma}_t^2$, the risk premium for the stock market is then given by

$$E_t[r_{s,t+1}]-r_{f,t+1}+\frac{\overline{\rho}_t\overline{\sigma}_t^2}{2}=\frac{\gamma}{\beta_t(1-\theta_t)}\overline{\rho}_t\overline{\sigma}_t^2-\frac{\gamma}{\beta_t(1-\theta_t)}\theta_t\overline{\sigma}_t^2. \quad (6)$$

Note that $w_{s,t}$ disappears completely from this expression. This expression states that the risk premium for the stock market is linearly related to stock market variance ($\sigma_{s,t}^2=\overline{\rho}_t\overline{\sigma}_t^2$), minus a correction for that part of stock market variance which is unrelated to aggregate risk $\theta_t\overline{\sigma}_t^2$.

By splitting stock market variance into two components $\overline{\rho}_t$ and $\overline{\sigma}_t^2$, we are able to explain why stock market variance is such a weak predictor of future stock market returns. Holding all else constant, a change in average variance affects the first term and the second term in opposite directions with similar magnitudes. Holding all else constant, an increase in average correlation only affects the first term. Consequently, if changes in stock market variance are mainly due to changes in average variance, then correlation will be strongly related to subsequent market returns while stock market variance will be largely unrelated.

To examine the empirical implications of this equation more concretely, we assume that the unobservable moments are parameters ($\beta_t=\beta$ and $\theta_t=\theta$) and consider a linear approximation of expression (6) around the unconditional expectation of average correlation $E[\overline{\rho}t]$ and average variance $E[\overline{\sigma}_t^2]$. We obtain the following expression for the expected excess return:

$$E_t[r_{s,t+1}]-r_{f,t+1}\approx\phi_0+\left(\frac{\gamma E[\overline{\sigma}_t^2]}{\beta(1-\theta)}-\frac{E[\overline{\sigma}_t^2]}{2}\right)\overline{\rho}_t+\left(\frac{\gamma(E[\overline{\rho}_t]-\theta)}{\beta(1-\theta)}-\frac{1}{2}\right)\overline{\sigma}_t^2, \quad (7)$$

where ϕ_0 is the constant of linearization. Both of the unobservable parameters, β and θ, affect the magnitudes of the coefficient estimates for average correlation and average variance. For plausible parameter values, the coefficient for average correlation $\overline{\rho}_t$ is positive while the coefficient for average variance $\overline{\sigma}_t^2$ is actually negative if $E[\overline{\rho}_t]$ is close to θ. Essentially, a large component of the realized stock market return must be unrelated to the realized return on aggregate wealth. Since average variance is the dominant component of stock market variance, its zero or negative risk premium explains the anemic relationship between stock market variance and subsequent market returns. This linear approximation provides the theoretical underpinnings for our empirical analysis in the next section.

Under these assumptions, the risk premium for the stock market is positively related to a weighted average of stock market variance and the covariance of the stock market with the unobserved portfolio return. If average variance is not positively related to the risk premium for the stock market, then average variance should be negatively related to $Cov_t(R_{s,t+1},R_{u,t+1})$.

To the extent that we observe a portion of $Cov_t(R_{s,t+1},R_{u,t+1})$, we can construct an additional test of the model's implications. We continue to assume that the relevant unobservable moments are parameters ($\beta_t=\beta$, $\theta_t=\theta$ and $w_{s,t}=w_s$) and consider a linear approximation for expression (5) around the unconditional expectation of average correlation $E[\overline{\rho}_t]$ and average variance $E[\overline{\sigma}_t^2]$.

$$\begin{aligned}Cov_t(R_{s,t+1},R_{u,t+1})&\\ \approx\pi_0&+\left(\frac{(1-w_s\beta(1-\theta))E[\overline{\sigma}_t^2]}{(1-w_s)\beta(1-\theta)}\right)\overline{\rho}_t\\ &+\left(\frac{(1-w_s\beta(1-\theta))E[\overline{\rho}_t]-\theta}{(1-w_s)\beta(1-\theta)}\right)\overline{\sigma}_t^2.\end{aligned} \quad (8)$$

The denominators for both coefficients of interest are positive whenever $\beta>0$ and $0\le w_s\le 1$. Hence, the directional effect for each variable will be determined by the sign of the numerator for the respective coefficients for any reasonable parameter values.

The numerator of the coefficient for average correlation will be positive, although potentially quite small, if $1-w_s\beta>0$. This condition is equivalent to assuming that the covariance between the return for the unobserved asset and the return for aggregate wealth is positive. The analysis of the coefficient for average variance is more complicated. For plausible parameter values $1-w_s\beta(1-\theta)$ is small and $E[\overline{\rho}_t]$ is close to θ. Thus, the numerator of the coefficient for average variance is likely to be negative whenever the risk premium for the stock market is not related to average variance.

2.2. An approximation for stock market variance

In the previous subsection we consider the limiting case with an infinite number of symmetric stocks, and hence, stock market variance is identical to the product of correlation $\overline{\rho}_t$ and individual stock variance $\overline{\sigma}_t^2$. In our empirical work we use an approximation.

The stock market portfolio s is the value-weighted portfolio of all stocks where $w_{j,t}$ is the market capitalization of stock j divided by the market capitalization of the entire stock market. The variance of the stock market return is given by

$$\sigma_{s,t}^2=\sum_{j=1}^{N}\sum_{k=1}^{N}w_{j,t}w_{k,t}\rho_{jk,t}\sigma_{j,t}\sigma_{k,t}.$$

A useful approximation for stock market variance is the product of average correlation between all pairs of stocks and the average variance of all individual stocks. We define $\overline{\sigma}_t^2$ to be the value-weighted cross-sectional average variance for the N stocks.[3]

$$\overline{\sigma}_t^2 = \sum_{j=1}^{N} w_{j,t}\sigma_{j,t}^2.$$

We let $\xi_{jk,t}$ denote the pairwise stock-specific deviations from the cross-sectional average for variance,

$$\xi_{jk,t} = \sigma_{j,t}\sigma_{k,t} - \overline{\sigma}_t^2,$$

and rewrite the expression for stock market variance:

$$\sigma_{s,t}^2 = \sum_{j=1}^{N}\sum_{k=1}^{N} w_{j,t}w_{k,t}\rho_{jk,t}(\overline{\sigma}_t^2 + \xi_{jk,t}) = \overline{\sigma}_t^2 \sum_{j=1}^{N}\sum_{k=1}^{N} w_{j,t}w_{k,t}\rho_{jk,t} + \sum_{j=1}^{N}\sum_{k=1}^{N} w_{j,t}w_{k,t}\rho_{jk,t}\xi_{jk,t}. \quad (9)$$

Stock market variance is the sum of two terms. The first term is the product of the value-weighted average of individual stock return variances and the value-weighted average of return correlations across all pairs of stocks in the portfolio. The second term depends on the cross-sectional relationships between weights, pairwise correlations, and the cross-products of standard deviations. When all assets have the same individual variance, the second term is equal to zero and the expression can be simplified accordingly:

$$\sigma_{s,t}^2 = \overline{\sigma}_t^2 \sum_{j=1}^{N}\sum_{k=1}^{N} w_{j,t}w_{k,t}\rho_{jk,t} = \overline{\sigma}_t^2\overline{\rho}_t. \quad (10)$$

This expression has two components: value-weighted average variance and value-weighted average correlation. We approximate stock market variance with the right-hand side of Eq. (10). In the next section we show that this approximation captures almost all of the time-series variation in stock market variance.

3. Predicting stock market returns with correlation

3.1. Proxies for the components of stock market variance

Every quarter from 1963 until 2006 we estimate the two components on the right-hand side of expression (10) directly from daily stock returns for a value-weighted portfolio of the 500 largest stocks, by market capitalization, in the Center for Research in Security Prices (CRSP) universe.

Given the number of trading days, D_t, in quarter t, the sample variance of daily returns for stock j is

$$\hat{\sigma}_{j,t}^2 = \left(\frac{1}{D_t - 1}\sum_{d=1}^{D_t}\left((1+R_{j,d}) - \frac{1}{D_t}\sum_{d=1}^{D_t}(1+R_{j,d})\right)^2\right), \quad (11)$$

where d is a trading day in quarter t. We define $w_{j,t}$ as stock j's market capitalization divided by the sum of all 500 market capitalizations observed at the end of quarter t. Our estimator of $\overline{\sigma}_t^2$, average variance, is

$$AV_t = \sum_{j=1}^{500} w_{j,t}\hat{\sigma}_{j,t}^2. \quad (12)$$

Although we estimate the conditional covariances and variances using daily returns, we are interested in the relationship between the quarterly stock market excess return and these moments for the same time interval. Therefore, we multiply our daily variance and covariance estimates by 63, the average number of trading days in a quarter.

The sample correlation for stocks j and k, denoted $\hat{\rho}_{jkt}$, is calculated in the typical fashion given the definition of $\hat{\sigma}_{jt}^2$.

$$\hat{\rho}_{jk,t} = \frac{\hat{\sigma}_{jk,t}}{\hat{\sigma}_{j,t}\hat{\sigma}_{k,t}}.$$

Our estimator of $\sum_{j=1}^{N}\sum_{k\neq j} w_{j,t}w_{k,t}\rho_{jk,t}$, average correlation, is

$$AC_t = \sum_{j=1}^{N}\sum_{k\neq j} w_{j,t}w_{k,t}\hat{\rho}_{jk,t}. \quad (13)$$

We also estimate the conditional stock market variance, $\hat{\sigma}_{s,t}^2$, from the daily index returns of the CRSP value-weighted index. Hence, Eq. (9) implies that

$$\hat{\sigma}_{s,t}^2 = b_0 + b_1(AV_t * AC_t) + \upsilon_t, \quad (14)$$

where $E[\upsilon_t|AV_t * AC_t] = 0$. The coefficient b_1 may not equal one because (i) we exclude the sum of squared weights from AC_t, (ii) there is measurement error in AV_t and AC_t, and (iii) there are differences in the composition between the CRSP market index and the value-weighted portfolio of the 500 largest stocks.

3.2. Other data and summary statistics

In our predictive regressions, we use the log stock market return minus the log three-month T-bill return based on the CRSP value-weighted portfolio (the best available proxy for the portfolio of all stocks) and the Standard and Poor's (S&P) 500 (which is arguably untainted by the illiquidity that might affect returns of small stocks in the broader CRSP index).

We also use: GARCH(1,1), the in-sample conditional variance estimates for CRSP log market excess returns generated by a GARCH(1,1) estimator using excess returns for the entire sample period; *cay* from Lettau and Ludvigson (2001) and from Martin Lettau's Web site updated through 2006; *pd*, the log ratio of the S&P 500 index level to the previous one year's dividends on its constituent stocks; and *rf*, the log three-month Treasury bill yield. The variable *cay* is the most powerful in-sample return predictor of which we are aware. Ang and Bekaert (2007) present evidence that *pd* and *rf* jointly predict stock returns.

Table 1 reports summary statistics for all of the relevant time-series. All data are observed quarterly from the first quarter of 1963 to the end of 2006, except for each excess log return, which are observed through the

[3] Our findings do not depend on the weighting scheme used to calculate $\overline{\sigma}_t^2$.

Table 1
Summary statistics.
The stock market excess return, re_{CRSP}, is the quarterly log return for the CRSP value-weighted index minus the three-month T-bill log return. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. Average variance is the value-weighted cross-sectional average of the quarterly variance of daily returns for the same 500 large stocks. The weights in both of these measures are determined by end-of-quarter market capitalizations. Stock market variance is the daily variance of the CRSP value-weighted index return, estimated quarterly. GARCH(1,1) is the in-sample estimate of the conditional variance of the quarterly excess return for the stock market based on a model with a GARCH(1,1) specification for variance. The riskfree rate, *rf*, is the log three-month T-bill return. The variable *cay* is from Lettau and Ludvigson (2001) and is estimated using their full sample from 1951 Q4 to 2006 Q4. *pd* is the log price-dividend ratio for the S&P 500 index. The variables are all observed from 1963 Q1 to 2006 Q4. The variances of daily returns are multiplied by 63, the average number of trading days in one quarter, in order to express these sample moments in terms of the same time interval as quarterly excess returns.

	re_{CRSP}	Average correlation	Average variance (%)	Stock market variance (%)	GARCH(1,1)	*cay*	*rf*	*pd*
Mean	0.012	0.237	2.217	0.482	0.743	0.000	0.056	3.568
Median	0.025	0.238	1.688	0.310	0.581	−0.001	0.052	3.505
Min	−0.300	0.034	0.642	0.029	0.335	−0.036	0.009	2.816
Max	0.202	0.646	12.036	6.297	3.739	0.036	0.143	4.486
Standard deviation	0.084	0.093	1.806	0.609	0.477	0.014	0.026	0.413
Autocorrelation	0.019	0.579	0.700	0.318	0.661	0.863	0.926	0.985
				Correlation matrix				
re_{CRSP}	1.000	−0.293	−0.275	−0.422	−0.302	−0.063	−0.128	0.081
Average correlation		1.000	0.312	0.608	0.558	0.108	0.138	−0.095
Average variance			1.000	0.836	0.533	−0.206	0.006	0.381
Stock market variance				1.000	0.662	−0.057	−0.013	0.169
GARCH(1,1)					1.000	−0.013	0.002	−0.053
cay						1.000	0.026	−0.286
rf							1.000	−0.645
pd								1.000

first quarter of 2007.[4] Average correlation between all pairs of the 500 largest stocks in the market averaged 0.237 in our sample period, with a standard deviation of 0.093. The mean for average variance of the 500 largest stocks is 2.217% while the mean for stock market variance is, not surprisingly, much lower at 0.482%. Average variance is also three times as volatile as stock market variance and strongly positively correlated with stock market variance. Average variance has a smaller but positive correlation with average correlation of 0.312.

The first autocorrelation of each time-series is also reported in Table 1. Average correlation and average variance are both fairly persistent, with autocorrelations respectively of 0.579 and 0.700. Stock market variance is much less persistent, even though the correlation between average variance and average correlation is 0.312. Stock market variance, average correlation, and average variance are negatively correlated with contemporaneous stock market excess returns. The GARCH(1,1) estimate of stock market variance, *cay*, and *rf* are also negatively correlated with excess returns while *pd* is positively correlated with returns. Compared to the quarterly estimates of stock market variance based on the sample variance of daily returns, the GARCH(1,1) estimates of stock market variance are higher (with a mean of 0.743), less volatile (with a standard deviation of 0.477), and more persistent (with an autocorrelation of 0.661).

Fig. 1 plots the time series for average correlation and average variance. Our measure of average variance has peaks and troughs that parallel the behavior of the standard deviation measure considered by Goyal and Santa-Clara (2003). However, our measure does not exhibit a pronounced upward time trend. While average variance and average correlation move together to some extent, this positive association disappears in 1999. In addition, the distribution for average variance is more positively skewed than the distribution of average correlation. The shaded regions represent periods of recession according to the National Bureau of Economic Research (NBER). There does not appear to be an obvious pattern linking recessions to the behavior of either measure. Both measures peaked in the last quarter of 1987 due to the 'Black Monday' stock market crash.

3.3. Validity of the stock market variance decomposition

Table 2 reports ordinary least squares (OLS) regressions of the sample estimate of stock market variance from $t-1$ to t on various combinations of the sample estimates of average correlation and average variance for the same time period, $t-1$ to t. The t-statistics use Newey-West standard errors with six lags. The results do not qualitatively change if we use OLS standard errors or change the lag length for the Newey-West estimator.

Column 1 reports direct estimates of (14). The R^2 of the regression is 97.69%, which indicates that variation in $\hat{\sigma}^2_{s,t}$

[4] We drop the observation of each excess log return for the first quarter of 2007 when calculating summary statistics.

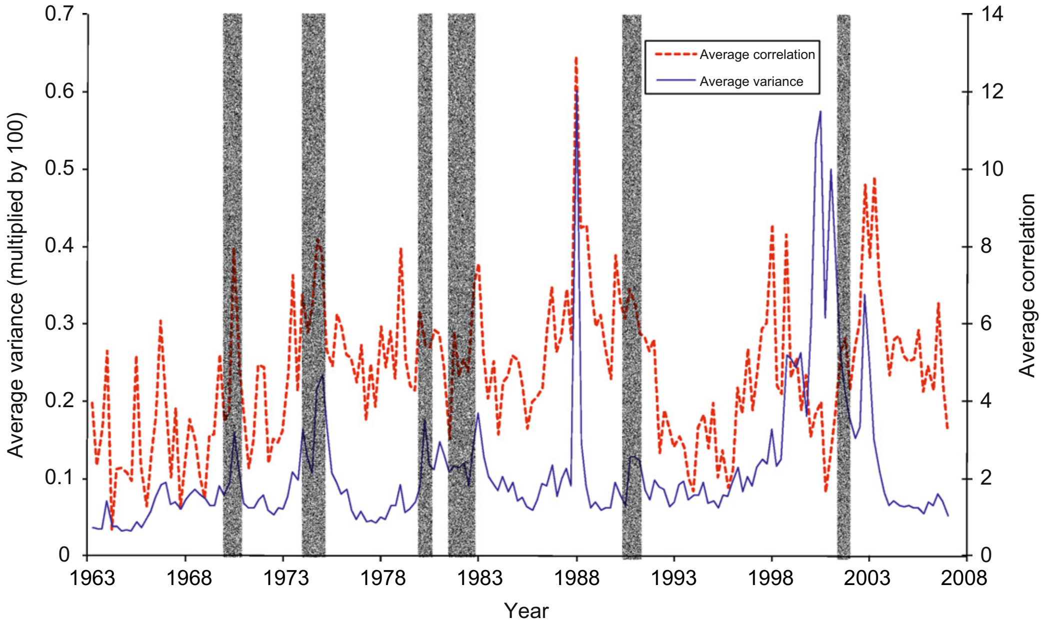

Fig. 1. Average correlation and average variance over time. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. Average variance is the value-weighted cross-sectional average of the variance of daily returns for the same 500 large stocks for the same period. The weights in both of these measures are determined by end-of-quarter market capitalizations. The shaded regions represent periods of recession according to the NBER. The variances of daily returns are multiplied by 63 in order to express these sample moments in terms of the same time interval as quarterly excess returns.

Table 2
Decomposing stock market variance.

Unless otherwise stated, the dependent variable is the variance of the CRSP value-weighted daily index return estimated quarterly from 1963 Q1 to 2006 Q4. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. Average variance is the value-weighted cross-sectional average of the quarterly variance of daily returns for the same 500 large stocks. The weights for both measures are determined by end-of-quarter market capitalizations. Newey-West *t*-statistics with six lags are reported in brackets.

	Dependent variable: variance of daily stock market returns estimated at *t*				
	(1)	(2)	(3)	(4)	(5)[a]
Constant	0.000	−0.005	−0.001	−0.007	0.000
	[0.108]	[1.810]	[1.004]	[3.741]	[2.011]
Average correlation		0.040		0.025	
		[3.719]		[4.517]	
Average variance			0.282	0.241	
			[3.513]	[4.619]	
(Avg. Var.) × (Avg. Corr.)	0.835				0.975
	[39.113]				[36.011]
R^2 (%)	97.69	36.90	69.82	83.17	98.95
N	176	176	176	176	176

[a] The dependent variable is the sample variance of the S&P 500 index daily return instead of the sample variance of the CRSP value-weighted index.

is almost entirely captured by the product of contemporaneous average variance and average correlation. Columns 2 through 4 present estimates of the relative importance of average variance and average correlation in changes in stock market variance. Column 2 shows that average correlation accounts for 36.9% of variation in stock market variance and column 3 that average variance accounts for a much larger 69.82%. Not surprisingly, since estimates of average variance and average correlation are positively correlated, the additional explanatory power provided by including average variance in a linear regression is substantial, as shown by the R^2 of 83.17% in column 4. The estimates in column 4 indicate that the linearization in Eq. (7) is reasonable because we are able to explain most of the variation in stock market variance. Column 5 uses the variance of the S&P 500 daily excess return instead of the excess return for the CRSP value-weighted index in column 1. The R^2 for this specification is even higher at 98.95%. Even though average correlation and average variance are both important components of contemporaneous realized stock market variance, Table 2 indicates that average variance is the dominant component. While this exercise indicates that stock market variance can be decomposed into average correlation and average variance, it does not offer any insights regarding subsequent realized stock market variance.

Table 3 analyzes the ability of average correlation and average variance (estimated from $t-1$ to t), to predict subsequent realized stock market variance $\hat{\sigma}^2_{s,t+1}$ (estimated from t to t+1). Column 1 shows that average correlation is, on its own, a significant predictor of $\hat{\sigma}^2_{s,t+1}$, with a t-statistic of 2.73 and an R^2 of 4.29%. However, average variance is a much more powerful predictor of future realized stock market variance, as shown in column 2, with an R^2 of 18.31% and a t-statistic of 5.65.

Column 3 of Table 3 shows that the product of average correlation and average variance, which captures almost all the variation in contemporaneous realized variance, predicts future realized variance with a t-statistic of 2.02 and an R^2 of 9.72%. Column 4 is a regression of $\hat{\sigma}^2_{s,t+1}$ on its own lag $\hat{\sigma}^2_{s,t}$. Although stock market variance predicts $\hat{\sigma}^2_{s,t+1}$ with an R^2 of 10.1%, it only just outperforms the product of average variance and average correlation. Column 5 indicates that GARCH(1,1) has a higher t-statistic (3.66) than either $\hat{\sigma}^2_{s,t}$ or the product of average variance and average correlation, but a lower R^2 than both (8.96%). In these univariate regressions average variance is a superior predictor of subsequent realized stock market variance compared to any of the three more traditional measures of variance. Column 6 shows that if average correlation and average variance are both included in the regression, each has a positive coefficient, but the R^2 only increases marginally to 18.91% compared to the specification that only uses average variance and the coefficient on average correlation is only marginally significant with a t-statistic of 1.76. Controlling for average variance in columns 7 through 9, we find that stock market variance, the product of average variance and correlation, and the GARCH(1,1) variance estimator do not have any significant power to predict future realized stock market variance.

We conclude that average variance is the dominant predictor for subsequent stock market variance and

Table 3
Predicting stock market variance.
The dependent variable is the variance of the CRSP value-weighted daily index return estimated quarterly from 1963 Q2 to 2007 Q1. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. Average variance is the value-weighted cross-sectional average of the quarterly variance of daily returns for the same 500 large stocks. The weights for both measures are determined by end-of-quarter market capitalizations. Stock market variance is the daily variance of the CRSP value-weighted index return, estimated quarterly. GARCH(1,1) is the in-sample estimate of the conditional variance of the quarterly excess return for the stock market based on a model with a GARCH(1,1) specification for variance. Newey-West t-statistics with six lags are reported in brackets.

	Dependent variable: variance of daily stock market returns estimated at t+1								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.002	0.002	0.003	0.003	0.002	0.001	0.002	0.002	0.001
	[1.125]	[2.055]	[3.454]	[3.510]	[2.281]	[0.505]	[2.256]	[2.227]	[1.088]
Average correlation	0.014					0.005			
	[2.727]					[1.759]			
Average variance		0.144				0.136	0.167	0.181	0.126
		[5.650]				[5.450]	[7.047]	[6.411]	[6.112]
(Avg. Var.) $\times$ (Avg. Corr.)			0.263				-0.073		
			[2.018]				[0.793]		
Stock market variance				0.318				-0.132	
				[2.124]				[1.116]	
GARCH(1,1)					0.382				0.127
					[3.660]				[1.499]
R^2 (%)	4.29	18.31	9.72	10.10	8.96	18.91	18.58	18.83	19.02
N	176	176	176	176	176	176	176	176	176

the dominant source of variation in contemporaneous realized stock market variance. Thus, if average variance is not related to aggregate risk or to the stock market risk premium, it should not be surprising that stock market variance is not strongly related to subsequent stock market returns.

3.4. Predictive regressions

In Table 4, we present quarterly stock market return forecasting regressions using the two major components of stock market variance $\hat{\sigma}^2_{s,t}$, average variance AV_t and average correlation AC_t. In column 5 and column 6 we also include various controls. Usually, the dependent variable is the excess log return of the CRSP value-weighted index. In column 4 the excess log return for the CRSP index is replaced by the excess log return for the S&P 500 index.

Column 1 shows that average correlation is a strong and highly economically significant forecaster of excess log stock market returns. A one-standard deviation increase in the average correlation of the daily returns of the 500 largest stocks in the market forecasts a 1.86% additional stock market excess return over the following quarter. The estimate is highly statistically significant with a robust t-statistic of 3.11. Variation in average correlation between stock returns accounts for 4.8% of variation in quarterly stock market return, comparing favorably with any other forecasting variable proposed to date.

Column 2 indicates that average variance has no forecasting power for subsequent excess returns in spite of the relative strength of average variance, as compared to average correlation, as a predictor for stock market variance in Table 3. The estimate is not significant and has the 'wrong' sign: an increase in stock market risk due to higher average individual variance forecasts a low future return.

In column 3, including both measures of risk in a linear regression further reduces the sign of the coefficient on average variance and boosts the forecasting power and significance of average correlation. As discussed in Section 2, it is the component of average correlation that is unrelated to average individual variance that predicts future stock market returns. The coefficient estimates in this regression indicate that average correlation has a positive risk premium while average variance has, if anything, a negative risk premium. Eq. (7) allows us to interpret the coefficients in column 4 in terms of the parameters of our model with symmetric stocks. For example, if the beta for each stock with the true market portfolio equals one ($\beta = 1$), then these estimates imply that the relative risk aversion coefficient γ is equal to 8.25.

The specification in column 4 replaces the excess log return of the CRSP index with the excess log return of the S&P 500 index, but is otherwise identical to the specification in column 3. The stocks in the S&P 500 index are all heavily traded and any return predictability that we observe is unlikely to be due to illiquid assets. Our linearization forecasts the S&P 500 excess return almost as well as it forecasts the broader CRSP index, with an R^2 of 5.22% and a t-statistic of 3.37 for average correlation. We conclude that average correlation is not merely a proxy for the changing liquidity of lightly traded assets in the broader stock market index.

In columns 5 and 6, our objective is to control for alternative predictors and ensure that average correlation is not simply replicating other well-known predictability results. The extensive literature about stock market predictability includes many variables that might forecast returns either singly or in combination with other variables. In our sample, most of these variables lack significant predictive power in either univariate forecasting regressions or controlling for other well-known forecasting variables. In unreported results, no other potential controls are significant once we include *cay*, *pd*, and *rf* in any regression specification for quarterly or monthly predictive regressions during our sample period. Therefore, we only consider *cay*, *pd*, *rf*, the lagged excess log return, and two measures of the variance of the excess log stock market return in our analysis. Column 5 shows that average correlation is significant with a t-statistic of 2.42 when we include controls for the lagged realized stock market variance, the lagged excess log return, *cay*, *pd*, and *rf*. Column 6 replaces the lag of realized stock market variance with the GARCH(1,1) estimate of conditional stock market variance and otherwise contains the same controls as Column 5. This specification yields similar results and average correlation is significant with a t-statistic of 2.40. In addition to average correlation, only *cay* and *rf* are statistically significant regressors.

To ensure that the strength of the predictability evidence for average correlation does not depend heavily on the frequency selected, we include a second panel in Table 4 (Panel B) in which we predict log excess returns at a monthly, rather than a quarterly, frequency. Since *cay* is only calculated on a quarterly basis, we simply use the same quarterly observation for the three constituent months within each quarter. We also generate a monthly GARCH(1,1) estimator based on monthly CRSP excess returns for the sample period. The other regressors, including average correlation and average variance, can be constructed every month, and hence, every third observation in the monthly data is identical to the contemporaneous observation in the quarterly data. The monthly regressors are observed from January 1963 to December 2006.

The patterns of statistical significance for average correlation in Table 4, Panel B are similar to those in Panel A. The coefficient for average correlation is positive and statistically significant. This result is stronger when average variance is included in the regression specification. The findings are similar if we replace the monthly excess log return for the CRSP value-weighted index with the excess log return of the S&P 500. In addition, average correlation is a significant predictor when controlling for lagged stock market variance or GARCH(1,1), together with lagged excess log returns, *cay*, *pd*, and *rf*.

3.5. Empirical proxies and measurement error

In unreported results, we do not find qualitatively different findings if we use an equal-weighted measure of

Table 4
Predicting stock market excess returns.

In Panel A, the dependent variable is the quarterly log return for the CRSP value-weighted index minus the log return for the three-month T-bill from 1963 Q2 to 2007 Q1. In Panel B, the dependent variable is the monthly log return for the CRSP value-weighted index minus the log return for the one-month T-bill from February 1963 to January 2007. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. Average variance is the value-weighted cross-sectional average of the quarterly variance of daily returns for the same 500 large stocks for the same period. Stock market variance is the daily variance of the CRSP index, estimated quarterly. GARCH(1,1) is the in-sample estimate of the conditional variance of the quarterly excess return for the stock market based on a model with a GARCH(1,1) specification for variance. re_m is the lagged dependent variable. The riskfree rate, *rf*, is the log three-month T-bill return. *cay* is from Lettau and Ludvigson (2001). *pd* is the log price-dividend ratio for the S&P 500 index. Newey-West *t*-statistics with six lags are reported in brackets.

Panel A: predicting the quarterly excess log return for the stock market at $t+1$

	(1)	(2)	(3)	(4)[a]	(5)	(6)
Constant	−0.036	0.013	−0.032	−0.027	0.130	0.098
	[2.162]	[1.301]	[2.090]	[1.848]	[1.223]	[0.924]
Average correlation	0.200		0.227	0.204	0.185	0.175
	[3.105]		[3.543]	[3.374]	[2.418]	[2.404]
Average variance		−0.083	−0.445	−0.428	−0.072	−0.055
		[0.184]	[1.633]	[1.663]	[0.125]	[0.140]
Stock market variance					1.076	
					[0.532]	
GARCH(1,1)						1.762
						[1.240]
re_m					0.103	0.098
					[1.409]	[1.413]
cay					1.032	1.073
					[2.131]	[2.261]
pd					−0.035	−0.029
					[1.310]	[1.111]
rf					−0.721	−0.658
					[2.284]	[2.349]
R^2 (%)	4.84	0.03	5.67	5.22	13.01	13.37
N	176	176	176	176	176	176

Panel B: predicting the monthly excess log return for the stock market at $t+1$

	(1)	(2)	(3)	(4)[a]	(5)	(6)
Constant	−0.007	0.005	−0.006	−0.005	0.072	0.057
	[1.635]	[1.840]	[1.415]	[1.151]	[2.290]	[1.679]
Average correlation	0.061		0.078	0.069	0.090	0.060
	[2.655]		[3.433]	[3.172]	[3.085]	[2.172]
Average variance		−0.055	−0.192	−0.173	0.257	0.005
		[0.477]	[1.864]	[1.679]	[1.098]	[0.031]
Stock market variance					−0.862	
					[1.280]	
GARCH(1,1)						0.735
						[0.264]

Table 4. (*continued*)

Panel B: predicting the monthly excess log return for the stock market at $t+1$						
	(1)	(2)	(3)	(4)[a]	(5)	(6)
re_m					0.039	0.046
					[0.840]	[0.973]
cay					0.306	0.321
					[2.271]	[2.370]
pd					−0.019	−0.014
					[2.252]	[1.634]
rf					−0.338	−0.282
					[3.200]	[2.930]
R^2 (%)	1.26	0.05	1.79	1.49	5.25	5.08
N	528	528	528	528	528	528

[a] The dependent variable is the excess log return on the S&P 500 index instead of the excess log return of the CRSP value-weighted index.

average correlation measure instead of the value-weighted version; use the previous one month (or six months) of daily returns instead of the previous three months to construct our estimates of average correlation and average variance; use the largest 50 or 100 stocks by market capitalization to construct average correlation instead of the largest 500; use the historical constituents of the S&P 500 index in the CRSP universe instead of the 500 largest stocks by market capitalization; use log average correlation instead of average correlation; or replace the dependent variable with the excess return in levels. We could also estimate average correlation and average variance using the 48 Fama-French industry portfolios (equal-weighted across industries) or using the 25 Fama-French size and book-to-market sorted portfolios (equal-weighted across portfolios). Both such measures of average correlation predict excess stock market returns with similar magnitudes and statistical significance. None of the corresponding average variance measures demonstrate any predictive power.

In the variance-in-mean framework, it is difficult to explain why average correlation predicts stock market returns while average variance does not predict stock returns when both average variance and average correlation are persistent components of stock market variance. In unreported variance-in-mean simulations we find that there is no simple explanation based on measurement error for either average correlation or average variance. First, these estimation errors appear to be small, and are much smaller for average variance than average correlation. Second, average variance should be the dominant predictor of subsequent stock market variance according to these simulations. Third, the simulations indicate that average variance and stock market variance should both be better predictors of subsequent stock market returns than average correlation.

3.6. Biased estimators in predictive regressions

We estimate a restricted vector autoregression (VAR(1)) for excess log stock market returns and average correlation. First, we are interested in the persistence of average correlation shocks. Second, we are interested in the volatility of shocks to stock market returns, the volatility of shocks to average correlation, and the correlation of these shocks. Third, we need to estimate the bias in the OLS estimate from the forecasting regressions due to persistence in average correlation and the correlation between the shocks to average correlation and realized returns.

The first equation in the VAR is

$$r_{s,t+1}-r_{f,t+1}=a+bAC_t+u_{t+1},$$

and the second equation is

$$AC_{t+1}=c+dAC_t+v_{t+1},$$

where the variances and covariance of the residuals are written σ_u^2, σ_v^2, and σ_{uv}. The coefficient b is our estimate of the ability of average correlation to forecast returns, but, as Stambaugh (1999) discusses, if σ_{uv} is negative then the estimate of b is biased upwards in small samples. If d, the persistence of shocks to average correlation, is large (close to one), this bias can be severe.

If average correlation forecasts returns, there are strong theoretical reasons to believe that σ_{uv} should be negative. Campbell (1991) shows that shocks to log returns are approximately equal to shocks to the present value of expected future log dividend growth minus shocks to future expected returns:

$$r_{t+1}-E_t[r_{t+1}]\approx\Delta E_{t+1}\sum_{s=0}^{\infty}\eta^s\Delta d_{t+1+s}-\Delta E_{t+1}\sum_{s=1}^{\infty}\eta^s r_{t+1+s}, \tag{15}$$

where the log-linearization constant η is slightly less than one. This approximate accounting identity shows that, holding expected future cash flows constant, a positive and persistent shock to expected returns must result in a negative shock to current realized returns, since the same expected future cash flows are discounted at higher rates. Expected returns whose innovations are negatively correlated with shocks to realized returns are said to be mean-reverting.

When expected returns are positively related to systematic risk, as captured by market variance or average correlation, persistent increases in risk cause increases in required or expected returns, inducing negative realized returns. Campbell and Hentschel (1992) refer to this special case of mean reversion as volatility feedback. The more persistent the shock to risk, the larger the volatility feedback. Therefore, volatility feedback generates an estimation bias of the form described by Stambaugh (1999).

Table 5 reports estimates of the restricted VAR. The parameter estimates for a and b in the top panel are familiar from Tables 1 and 4. Newey-West t-statistics with the usual six lags are reported in square brackets. The estimate of d in Table 5 is 0.579, the standard deviation of average correlation shocks σ_v is 7.5%, the standard deviation of the stock market return shocks σ_u is slightly more volatile than average correlation shocks and is 8.2%, and the correlation between the shocks ρ_{uv} is -0.53.

To estimate the bias under the null hypothesis of no predictive power, we carry out a simulation exercise following that in Baker, Taliafero, and Wurgler (2006). The OLS estimate of d is biased downwards by approximately $(1+3d)/T$, where T is the sample length. Our bias-adjusted estimate of d in the full sample is therefore 0.596. We use our estimates of the constant terms in each equation and of the residual covariance matrix but also impose different values of ρ_{uv} (the correlation of the residuals) and adjust σ_{uv} accordingly. We then simulate data for average correlation and for the excess log return for the stock market under the assumption that they are generated by the restricted VAR with b equal to zero. We use the actual starting variables and discard the first 100 simulated observations, then calculate OLS estimates of b for the remaining sample.

The lower panel of Table 5 reports choices of d and ρ_{uv} for the simulation, together with the mean implied estimate of b, its standard deviation, and the probability that the simulated estimate is greater than the actual estimate. This probability is our Monte Carlo estimate of the correct small-sample bias-corrected p-value for our estimate of b under the null that true b is zero.

For the full sample, using d equal to 0.596 we should estimate b of 0.200 less than 0.5% of the time under the null according to the simulation using the sample estimate of ρ_{uv}. In this case the bias in the estimator for b is about 0.01. For the worst-case scenario of $\rho_{uv}=-1$, we estimate a slightly higher p-value of 0.7% but we still reject the null at conventional levels of significance. Increasing d to 0.749, three standard errors above its OLS estimate, and setting $\rho_{uv}=-1$ does not materially change the p-value. Although the bias in $\hat{b}$ increases to 0.021 because average correlation is more persistent, the dispersion of the simulated estimates is lower. Similarly, choosing a high persistence of 0.95, while increasing the bias to 2.6%, actually reduces the p-value to 0.1%. These results indicate that any bias due to the persistence of average correlation is not a serious problem for the estimate of b in the full sample.

4. Alternative specifications

4.1. Subsample analysis

In Table 6 we estimate predictive regressions for the excess log return for the CRSP value-weighted index on average correlation in different subsamples. We report estimates for the full sample (from the end of the second

Table 5
Analysis of bias in predictive regressions with a persistent regressor.

The table reports results from simulations designed to address the potential for bias caused by a persistent regressor in predictive regressions with a small sample (Stambaugh, 1999). The dependent variable is the quarterly log return for the CRSP value-weighted index minus the log return for the three-month T-bill from 1963 Q2 to 2007 Q1. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. ρ_{uv} is the correlation of the residuals from the two regressions. Following Baker, Taliafero, and Wurgler (2006) we select the following parameters for the simulation: (1) d = 0.596 is the bias-corrected estimate of d based on Proposition 4 in Stambaugh (1999), and (2) d=0.749 is the bias-corrected estimate plus three standard errors. Using the relevant OLS estimates and the assumed values for d and ρ_{uv}, we generate pseudodata for the full sample under the null that b=0 for each of 50,000 simulations. For every simulation we estimate b. The mean and standard deviation of the resulting distribution are reported in columns 3 and 4 of the bias simulations. Column 5 reports the probability that a simulated estimate is larger than the original OLS for each set of simulation parameter values under the null hypothesis that b=0. Since the simulated distribution is slightly positively skewed, these probabilities are larger than those for a normal distribution. Newey-West t-statistics with six lags are reported in brackets.

Excess log stock market return at t+1=a + b ∗ (Average correlation at t) + $u(t$+1)
Average correlation at t+1=c + d ∗ (Average correlation at t) + $v(t$+1)

Original OLS coefficient estimates			a	b	c	d	
			−0.036	0.200	0.099	0.579	
			[2.162]	[3.105]	[7.243]	[11.572]	
Covariance and correlation estimates		Covariance matrix			Correlation (standard deviation)		
		u	v		u	v	
	u	0.007	−0.003	u	(0.082)	−0.528	
	v		0.006	v		(0.075)	
Bias simulations		d	ρ_{uv}	Mean($b_{sim}\|b=0$)	Std($b_{sim}\|b=0$)	P($b_{sim}>0.200\|b=0$)	
		0.596	−0.528	0.009	0.068	0.005	
		0.596	−1.000	0.017	0.067	0.007	
		0.749	−1.000	0.021	0.058	0.004	
		0.950	−1.000	0.026	0.035	0.001	

Table 6
Predicting stock market excess returns during different sample periods.

The dependent variable is the quarterly log return for the CRSP value-weighted index minus the log return for the three-month T-bill for each particular sample period. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. This cross-sectional average is value-weighted based on the end-of-quarter market capitalizations. Newey-West t-statistics with six lags are reported in brackets.

Dependent variable: quarterly excess log return for the CRSP value-weighted index at $t+1$									
Sample period	Full sample	1963 Q2–1985 Q1	1985 Q2–2007 Q1	1963 Q2–1974 Q1	1974 Q2–1985 Q1	1985 Q2–1996 Q1	1996 Q2–2007 Q1	1976 Q1–2007 Q1	1963 Q2–1989 Q4
Constant	−0.036	−0.049	−0.023	−0.036	−0.136	0.024	−0.105	−0.037	−0.035
	[2.162]	[2.037]	[0.930]	[1.633]	[2.418]	[1.389]	[3.592]	[1.525]	[1.898]
Average correlation	0.200	0.246	0.158	0.226	0.538	−0.013	0.458	0.208	0.188
	[3.105]	[2.350]	[1.854]	[1.887]	[2.424]	[0.240]	[3.769]	[2.376]	[2.572]
R^2(%)	4.84	5.85	3.51	6.07	10.56	0.03	19.01	5.12	4.10
N	176	88	88	44	44	44	44	125	107

Table 7
Predicting stock market excess returns at different frequencies.

The dependent variable is the sum of one-month log returns for the CRSP value-weighted index minus the one-month T-bill for k months. The sample period of the monthly log excess returns is from February 1963 to January 2007. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. This cross-sectional average is value-weighted based on the end-of-quarter market capitalizations. All coefficient estimates are annualized by multiplying them by $12/k$. We use overlapping monthly observations with Hodrick (1992) t-statistics reported in brackets.

Dependent variable: Sum of monthly excess log returns for the CRSP index from $t+1$ to $t+k$								
Forecast horizon (months)	1	3	6	9	12	18	24	30
Constant	−0.084	−0.098	−0.0645	−0.038	−0.031	−0.027	−0.025	−0.013
	[1.551]	[2.062]	[1.576]	[0.960]	[0.806]	[0.747]	[0.700]	[0.380]
Average correlation	0.734	0.806	0.620	0.469	0.429	0.408	0.394	0.326
	[2.467]	[3.220]	[2.905]	[2.299]	[2.223]	[2.293]	[2.424]	[2.152]
R^2(%)	1.26	4.33	5.14	4.36	4.90	6.82	9.04	8.33
N	528	526	523	520	517	511	505	499

quarter of 1963 to the end of the first quarter 2007), both halves of the sample, all four quarters of the sample, from 1976 until 2007, and from 1963 until 1989. The period from 1976 until 2007 begins after the first oil price shock in the early 1970s. Goyal and Welch (2008) find that many candidate forecasting variables have very little forecasting power in this subsample and they conclude that these candidates rely heavily on the oil price shock of the early 1970s for in-sample forecasting power. The period 1963 to 1989 excludes the 1990s during which several of the predictive relationships proposed by the literature are quite weak (e.g., price-dividend ratio, price-earnings ratio, etc.).

In the periods 1963–1974, 1974–1985, and 1996–2007 average correlation has predictive power for stock market returns. For the 30 years since 1976 it is also a strongly economically and statistically significant predictor, with an R^2 of 5.12% for the stock market return. The results are also significant during the earlier period ending in 1989, with an R^2 of 4.1% and a t-statistic of 2.57. Average correlation is less powerful in the second half of the sample, from 1985–2007, with a t-statistic of 1.85. Column 6 shows that from 1985–1996 average correlation has no ability to forecast market returns. This is not due to the stock market crash of 1987: in unreported results we estimate the same equation excluding the quarters around the crash and the results remain essentially unchanged. In Fig. 1 it can be seen that average correlation fell to very low levels in the late 1980s even though returns remained high. With the exception of the decade 1985–1996, the coefficients for average correlation are between 0.15 and 0.55.

4.2. Different horizons

Table 7 reports predictive regression estimates of CRSP log market excess returns for return intervals of one to 30 months using overlapping data. We regress the subsequent annualized k-month sum of monthly log stock market returns in excess of the one-month T-bill log return, $r^e_{t+k}=(12/k)\sum_{j=1}^{k}(r_{CRSP,t+j}-r_{ft+j})$ on average correlation at date t. Because the residuals of this regression have an $MA(k-1)$ structure, we report t-statistics based on Hodrick (1992) standard errors rather than Newey-West standard errors.[5]

The coefficient for the monthly horizon in Table 6 is identical to the coefficient in column 1 of Table 4, Panel B except that the coefficient in this table has been annualized by multiplying by 12. Average correlation forecasts one-month-ahead stock market returns with a Hodrick t-statistic of 2.47 and an R^2 of 1.3%. Average correlation predicts excess log returns for longer return horizons up to 30 months with a significance level of more than 5%. The R^2 of the regression grows with the return horizon up to six months (to 5.14%), and is again higher at horizons of 18 months (6.82%) and 24 months (9.04%) falling slightly at horizons of 30 months (8.33%). The annualized coefficient declines monotonically with horizon from three months to 30 months.

Since average correlation is not a highly persistent variable, it is not surprising its ability to predict excess returns does not grow substantially with the return horizon. However, our results suggest that average correlation has information relevant for one-year and two-year excess returns.

4.3. Out-of-sample findings

Goyal and Welch (2008) suggest that most forecasting variables with in-sample forecasting power do not demonstrate ability to forecast returns out-of-sample. They propose a test in which forecasting regressions use only data available at t and the one-period-ahead squared forecast error is compared to the squared difference between year t+1 realized return and the sample mean return for time t. They then compare either the difference in root mean squared error ("$\Delta RMSE$") or out-of-sample R^2 (one minus the ratio of the sum of squared forecast errors over the sum of squared differences from the time t sample mean). Both these measures will be positive if the proposed variable has superior forecasting ability out-of-sample relative to an average of past returns, zero if there is no difference, and negative otherwise. With the exception of *cay*, they fail to find superior out-of-sample performance for any variable. This failure suggests that the estimates from forecasting regressions are unstable and quite sensitive to sample period choice.

Campbell and Thompson (2005) indicate that short in-sample estimation periods could be responsible for poor out-of-sample performance. In spite of this potential concern, we do not use the full sample available for all other variables and restrict our analysis to the period from 1963 until 2007. Since our focus is on average correlation and average variance, this approach ensures that the critical values for the formal test statistics and the period of assessment will be the same for each predictor.

Clark and McCracken (2001) propose a formal test for out-of-sample predictability. They show how the asymptotic distribution of *Enc_New*, the forecast encompassing test statistic, depends on the ratio of R, the number of observations used for the first out-of-sample forecast, to P, the number of pseudo-out-of-sample forecasts. This ratio is denoted $\pi=\lim_{R,P\to\infty}P/R$. McCracken (2007) presents a different test statistic, *MSE_F*, based on reduction of out-of-sample mean square error whose asymptotic distribution also depends on π.

The derivations in these papers are valid for heteroskedasticity of unknown form in the forecast error if there is one additional forecasting variable relative to a baseline forecast. Accordingly, we test whether a number of different variables have forecasting power relative to a forecast based only on the sample mean of returns. We use the recursive forecasting scheme in which one-step-ahead forecasts are based on estimates using the sample available up to the date of the forecast.

Table 8 presents results regarding the out-of-sample R^2, $\Delta RMSE$, *Enc_New* (from Clark and McCracken), and *MSE_F* (from McCracken) for average correlation, average variance, sample stock market variance, GARCH(1,1), *cay*, *rf*, and *pd*. In the upper panel, we allow only the first 80 observations of returns to be used to make the first prediction, so that there are 95 forecasts and $\pi=1.19$. In the lower panel, we allow more observations to be used so that R=120, P=55, and $\pi=0.46$. We obtain asymptotic critical values from tables on Clark's Web site, for the nearest reported values of π and k=1.

The findings in Table 8 indicate that forecasts based on average correlation strongly increase predictive power relative to forecasts based only on the sample mean. For R=80, the out-of-sample R^2 is 3.34%, and 8.22% when 120 observations are used to make the first forecast. The corresponding reductions in root mean squared forecast error are 0.137 and 0.342. For R=120, average correlation is a statistically significant predictor at the 1% level for both formal tests. For R=80 there is a significant reduction in out-of-sample mean squared error at the 5% level and rejection of the null of encompassing at the 1% significance level.

Average variance, market variance, GARCH(1,1), and *pd* all have negative $\Delta RMSE$ and out-of-sample R^2 during both out-of-sample assessment periods.[6] For R=80, *cay* has a negative out-of-sample R^2 while *rf* has a slightly positive one. For R=120 this pattern is reversed. With the exception of *cay*, we find little formal evidence for substantial out-of-sample forecasting power for the other variables, save that we reject the hypothesis of encompassing for stock market variance at the 5% level when R=80.[7] This failure suggests that the results from standard predictive regressions are quite sensitive to sample period choice.

[5] Ang and Bekaert (2007) show that inferences based on these standard errors are superior to those based on Newey-West standard errors when returns are overlapping.

[6] It is possible for a forecasting variable to have both a negative out-of-sample R^2 and yet to be able to reject the hypothesis of forecast encompassing.

[7] In unreported results many other univariate predictors commonly considered in the literature do not have predictive power in-sample or out-of-sample in quarterly data during our sample period.

Table 8
Out-of-sample predictability of the quarterly excess log stock market return.

Each column analyzes the forecasting properties of a variable for the quarterly log excess return on the CRSP index from 1963 Q2 to 2007 Q1. Following Goyal and Welch (2008), the forecasting power of each variable is compared to the forecasting power of the sample mean of the return using only information available at the date of the forecast. A positive out-of-sample R^2 and a positive difference in root mean squared forecast error (ΔRMSE) indicate superior forecasting ability compared to the historical sample mean. Following Clark and McCracken (2001) and McCracken (2007), R is the number of observations used to make the first pseudo-out-of-sample forecast and P is the number of such forecasts. We conduct two formal out-of-sample predictability tests. Enc_New is the F-test value for forecast encompassing from Clark and McCracken (2001) and MSE_F is the F-test value for forecasting power from McCracken (2007). A significant value of either statistic (compared to the relevant asymptotic distribution for $\pi = P/R$ and $k=1$) indicates that the variable adds information to the forecast relative to the historical mean. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. Average variance is the value-weighted cross-sectional average of the quarterly variance of daily returns for the same 500 large stocks. The weights in both of these measures are determined by end-of-quarter market capitalizations. Stock market variance is the daily variance of the CRSP value-weighted index return, estimated quarterly. GARCH(1,1) is the in-sample estimate of the conditional variance of the quarterly excess return for the stock market based on a model with a GARCH(1,1) specification for variance. The riskfree rate, *rf*, is the log three-month T-bill return. *cay* is from Lettau and Ludvigson (2001). *pd* is the log price-dividend ratio for the S&P 500 index. * and ** denote significance at the 5% and 1% levels, respectively.

Forecasting variable:	Average correlation	Average variance	Stock market variance	GARCH(1,1)	*cay*	*rf*	*pd*
			Test parameters: R=80, P=95, and $\pi = 1.19$				
Out-of-sample R^2	3.337	−8.796	−27.071	−1.435	−0.795	0.048	−6.126
ΔRMSE	0.137	−0.351	−1.036	−0.058	−0.032	0.002	−0.246
Enc_New	4.123**	0.247	2.627*	1.289	5.897**	0.467	−1.018
MSE_F	3.487*	−7.620	−20.186	−1.278	−0.683	0.112	−5.421
			Test parameters: R=120, P=55, and $\pi = 0.46$				
Out-of-sample R^2	8.218	−9.920	−2.870	−4.005	3.514	−0.593	−11.056
ΔRMSE	0.342	−0.395	−0.116	−0.162	0.145	−0.024	−0.439
Enc_New	3.588**	−0.846	−0.147	−0.146	3.471**	0.139	−1.614
MSE_F	4.958**	−4.936	−1.505	−2.089	2.035*	−0.294	−5.448

5. Predicting covariances

In this section, we test the second implication of the model. Because average variance explains a substantial portion of the variation of conditional stock market variance according to Table 3, the failure to forecast returns with average variance is a puzzle for conventional models. Eq. (8) describes the relationship between average correlation, average variance, and the covariance of the stock market return with the return for non-stock assets in our model. Since the net effect of an increase in average variance on the stock market risk premium is small or negative, the model parameters consistent with this result indicate that average variance should be negatively related to the conditional covariance of the stock market with the return on the non-stock remainder of the aggregate wealth portfolio. By contrast, the same parameter values indicate that average correlation should be positively related to the conditional covariance of the stock market return and the return on the non-stock component of wealth.

Accordingly, there should be classes of non-stock assets whose subsequent conditional covariance with the stock market is positively related to average correlation and negatively related to the average variance of stocks. Our results indicate that both government bonds and labor income are asset classes whose covariance properties are largely consistent with this prediction. The findings for bonds as well as labor income suggest that investors who possess bonds and/or human capital in addition to stocks will not necessarily be exposed to greater portfolio risk when there is an increase in average variance for stocks, but will likely face greater portfolio risk when there is an increase in average correlation between stocks.

While the return on the non-stock component of aggregate wealth, R_u, is largely unobservable, it is relatively straightforward to estimate the covariance of the return for government bonds with the return for the stock market. Government bonds are an important component of aggregate wealth to the extent that they are not in zero net supply (i.e., to the extent that Ricardian equivalence does not hold). Moreover, the returns for government bonds are presumably highly correlated with those of corporate bonds, which are in positive net supply even if Ricardian equivalence does hold. We estimate the required covariance with the within-quarter sample covariance of daily returns on the CRSP value-weighted index and the 10-year US Treasury bond and we term this variable the stock-bond covariance.

Human capital is another important non-stock component of aggregate wealth. Aggregate labor income can be interpreted as the dividend on human capital following Jagannathan and Wang (1996) or as its non-stationary component following Lettau and Ludvigson (2001). Although the return on human capital is probably more volatile than labor income growth for the same reasons that the return on the stock market is more volatile than aggregate dividend growth, labor income growth is approximately proportional to a noisy measure of the return on human capital under plausible assumptions. Aggregate labor income for the US is observable only at monthly or lower frequencies. We construct monthly log real per capita labor income y_t from Table 2.6 of the National Income and Product Accounts (NIPA). The first difference of y_t is monthly labor income growth, Δy_t.

Given the evidence that average correlation predicts stock market returns at a monthly frequency, the product of the stock market return and labor income growth should be predictable as well. Hence, we construct an estimate of the conditional covariance of the quarterly stock market return and quarterly labor income growth using summed residuals from a one-month-ahead forecast of the *monthly* excess log stock return and log labor income growth on *quarterly* average correlation, average variance, log excess stock market return, and log labor income growth, all lagged one month. We define the conditional stock-labor covariance estimate for the quarter ending in month t as

$$\hat{\sigma}_{ry,t} = \sum_{s=1}^{3} (\varepsilon_{r,t+1-s})(\varepsilon_{y,t+1-s}),$$

where $\varepsilon_{r,t+1-s}$ and $\varepsilon_{y,t+1-s}$ are the monthly residuals for stock market returns and labor income growth, respectively.

The first four columns of Table 9 present evidence on stock-bond covariance. In column 1 average correlation has a positive coefficient but is not statistically significant, while the coefficient for average variance is negatively related to the subsequent stock-bond covariance in column 2 (marginally significant with a t-statistic of 1.8). The specification including both average correlation and average variance in column 3 indicates that average variance is significantly negatively related to the stock-bond covariance (t-statistic of 2.5). This evidence is consistent with the hypothesis that variation in average variance, and therefore stock market variance, is due to changes in risk common to all stocks but unrelated to changes in aggregate risk.

In columns 5 through 8 we investigate the relation between average correlation, average variance, and the subsequent stock-labor covariance. In column 5 the coefficient for average correlation is positive and marginally statistically significant. In column 2, average variance is negatively related to stock-labor covariance, although not significantly so. Controlling for both, an increase in average correlation leads to a statistically significant increase for stock-labor covariance (t-statistic of 2.1), whereas an increase in average variance predicts a marginally significant decrease (t-statistic of 1.7). While the direction and statistical significance are consistent with the predictions of the model in Section 2, the non-observable nature of returns on human capital implies that the magnitudes of the coefficients are particularly difficult to interpret. In addition, the difficulties associated with estimating the conditional covariance of stock market returns and aggregate labor income indicates that the findings about the covariance between stock market returns and the return to human capital in Table 9 are only suggestive.

6. Conclusion

The absence of an easily detectable relationship between stock market risk and subsequent stock market returns poses difficulties for mainstream asset pricing

Table 9
Predicting the covariance of the stock market with other assets.
For columns 1 to 4, the dependent variable is the stock-bond covariance estimated quarterly from 1963 Q2 until 2007 Q1. It is the sample covariance of CRSP value-weighted daily index return and the daily return on 10-year US government bonds during quarter $t+1$. For columns 5 to 8, the dependent variable is the stock-labor covariance estimated quarterly from 1963 Q2 until 2007 Q1. It is the sum of products of the CRSP value-weighted index log monthly excess return residuals and log monthly labor income growth residuals for the three months in quarter $t+1$. These residuals are estimated from a one-step-ahead linear forecast (additional details in text). The table reports coefficient estimates for regressions of the two dependent variables on information known at the end of quarter t. Average correlation is the value-weighted cross-sectional average of the pairwise correlation of daily returns during each quarter for all pairs of the 500 largest (by market capitalization) exchange-traded stocks in the United States at the end of that quarter. Average variance is the value-weighted cross-sectional average of the quarterly variance of daily returns for the same 500 large stocks. The weights in both of these measures are determined by end-of-quarter market capitalizations. Stock market variance is the daily variance of the CRSP value-weighted index return, estimated quarterly. Newey-West t-statistics with six lags are reported in brackets.

	Stock-bond covariance				Stock-labor covariance			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.036	0.072	0.054	0.050	−0.0001	0.0001	0.0001	0.0000
	[0.967]	[4.508]	[3.481]	[1.226]	[0.783]	[2.799]	[2.953]	[0.5317]
Average correlation	−0.003			0.107	0.0005			0.0007
	[0.014]			[0.640]	[1.7805]			[2.1020]
Average variance		−1.628		−1.799		−0.0013		−0.0024
		[1.795]		[2.505]		[1.3938]		[1.7421]
Stock market variance			−3.841				−0.0080	
			[0.919]				[1.3301]	
R^2(%)	0.00	5.62	3.55	6.19	1.46	0.33	1.33	2.45
N	176	176	176	176	176	176	176	176

models. We claim that the lack of evidence of such a relationship is a manifestation of the Roll critique: if the aggregate wealth portfolio contains important non-stock assets, then changes in stock market variance are not necessarily closely related to changes in aggregate risk. Therefore, even if aggregate risk drives the risk premium, stock market variance need not be related to excess returns.

We propose an explanation based on the observation that aggregate risk should be a source of interdependence among observable stock returns. Hence, an increase in aggregate risk is partially reflected in an increase in average correlation between stock returns whenever the stock market as a whole has a positive sensitivity to aggregate market shocks. Therefore, average correlation should forecast stock market returns whenever there is an aggregate risk-expected return trade-off. Furthermore, changes in stock market variance that are unrelated to changes in average correlation could have a zero or even negative relationship with future stock market returns. We show that such unpriced stock market risk is, for practical purposes, captured by the average variance of individual stocks.

We present new results indicating that average correlation and average variance together account for almost all variation in stock market variance and that average variance is the dominant component. We show that average correlation forecasts future stock market returns while average variance is negatively related to future returns.

In addition, average correlation is positively related to the subsequent covariance between labor income growth and the stock market and (weakly) positively related to the subsequent covariance between stock market and bond returns. On the other hand, average variance is negatively related to the subsequent covariance between stock market and bond returns as well as the subsequent covariance between labor income growth and the stock market. Both of these findings are consistent with the hypothesis that changes in average variance unrelated to average correlation do not necessarily reflect changes in aggregate risk.

The main alternative explanation in the existing literature assumes some degree of deviation from a typical representative agent rational-choice framework. Such deviations imply that linear or log-linear relations between the risk premium and aggregate risk do not hold. By contrast, our paper assumes that the risk premium for aggregate wealth is related in a straightforward way to aggregate risk. We apply the Roll critique to the risk-return trade-off for the stock market and we exploit the common component of observable asset returns to reveal aggregate risk.

References

Ang, A., Bekaert, G., 2007. Stock return predictability: Is it there? Review of Financial Studies 20, 651–707

Ang, A., Chen, J., 2002. Asymmetric correlation of equity portfolios. Journal of Financial Economics 63, 443–494.

Ang, A., Liu, J., 2007. Risk, return and dividends. Journal of Financial Economics 85, 1–38.

Baker, M., Taliafero, R., Wurgler, J., 2006. Predicting returns with managerial decision variables: Is there a small-sample bias? Journal of Finance 61, 1711–1730

Bali, T., Cakici, N., Yan, X., Zhang, Z., 2005. Does idiosyncratic risk really matter? Journal of Finance 60, 905–929

Brown, G., Kapadia, N., 2007. Firm-specific risk and equity market development. Journal of Financial Economics 84, 358–388.

Campbell, J., 1987. Stock returns and the term structure. Journal of Financial Economics 18, 373–399.

Campbell, J., 1991. A variance decomposition for stock returns. Economic Journal 101, 157–179.

Campbell, J., 1993. Intertemporal asset pricing without consumption data. American Economic Review 83, 487–512.

Campbell, J., Hentschel, L., 1992. No news is good news: an asymmetric model of changing volatility in stock returns. Journal of Financial Economics 31, 281–318.

Campbell, J., Lo, A., MacKinlay, A., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ.

Campbell, J., Thompson, S., 2005. Predicting the equity premium out of sample: Can anything beat the historical average? NBER Working Paper No. 11468.

Campbell, J., Viceira, L., 2002. Strategic Asset Allocation. Oxford University Press, New York, NY.

Clark, T., McCracken, M., 2001. Tests of equal forecast accuracy and encompassing for nested models. Journal of Econometrics 105, 85–110.

Driessen, J., Maenhout, P., Vilkov, G., 2009. The price of correlation risk: Evidence from equity options. Journal of Finance 64, 1377–1406.

French, K., Schwert, G., Stambaugh, R., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–29.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2005. There is a risk-return trade-off after all. Journal of Financial Economics 76, 509–548.

Glosten, L., Jagannathan, R., Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779–1801.

Goyal, A., Santa-Clara, P., 2003. Idiosyncratic risk matters!. Journal of Finance 58, 975–1008.

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

Harvey, C., 2001. The specification of conditional expectation. Journal of Empirical Finance 8, 573–638.

Hodrick, R., 1992. Dividend yields and expected stock returns: alternative procedures for inference and measurement. Review of Financial Studies 5, 357–386.

Hong, Y., Tu, J., Zhou, G., 2007. Asymmetric correlation of stock returns: statistical tests and economic evaluation. Review of Financial Studies 20, 1547–1581.

Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. Journal of Finance 51, 3–53.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth and expected stock returns. Journal of Finance 56, 815–849.

Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. Journal of Finance 56, 649–676.

McCracken, M., 2007. Asymptotics for out-of-sample tests of Granger causality. Journal of Econometrics 140, 719–752.

Merton, R., 1973. An intertemporal capital asset pricing model. Econometrica 41, 867–887.

Roll, R., 1977. A critique of the asset pricing theory's tests: part I. Journal of Financial Economics 4, 129–176.

Stambaugh, R., 1999. Predictive regressions. Journal of Financial Economics 54, 375–421.