Is There a Relation Between Downside Risk and Expected Stock Returns?•

Turan G. Bali, K. Ozgur Demirtas and Haim Levy*

Abstract

This paper examines the intertemporal relation between downside risk and expected stock returns. Value at risk (VaR), expected shortfall, and tail risk are used as measures of downside risk to determine the existence and significance of a risk-return tradeoff for several stock market indices. We find a positive and significant relation between downside risk and the portfolio returns on the NYSE/AMEX/Nasdaq stocks. This result also holds for the NYSE/AMEX, NYSE, Nasdaq, and S&P 500 index portfolios. Moreover, VaR remains to be a superior measure of risk even when it is compared to the traditional risk measures which have significant predictive power for market returns. These results are robust across different measures of downside risk, loss probability levels, and after controlling for macroeconomic variables associated with business cycle fluctuations.

Key words: ICAPM, downside risk, value at risk, expected shortfall, tail risk, stock returns, risk-return tradeoff.

JEL classification: G10, G11, C13

•We thank Pierluigi Balduzzi, David Chapman, Wayne Ferson, Maureen O'Hara, Jeff Pontiff, and David Weinbaum for their extremely helpful comments and suggestions. We also benefited from discussions with Linda Allen, George Aragon, Armen Hovakimian, Salih Neftci, Lin Peng, Robert Schwartz, Jonathan Wang, and Liuren Wu. An earlier version of this paper was presented at Baruch College, Boston College, Cornell University, and the Graduate School and University Center of the City University of New York. Turan Bali gratefully acknowledges the financial support from the Eugene Lang Research Foundation of the Baruch College. K. Ozgur Demirtas gratefully acknowledges the financial support from the PSC-CUNY Research Foundation of the City University of New York. Haim Levy acknowledges the financial support from the Krueger Center of Finance. All errors remain our responsibility.

*Turan Bali is an associate professor of finance at the Department of Economics and Finance, Zicklin School of Business, Baruch College, City University of New York, One Bernard Baruch Way, Box 10-225, New York, New York 10010. Phone: (646) 312-3506, Fax: (646) 312-3451, E-mail: Turan_Bali@baruch.cuny.edu; K. Ozgur Demirtas is an assistant professor at the Department of Economics and Finance, Zicklin School of Business, Baruch College, City University of New York, One Bernard Baruch Way, Box 10-225, New York, New York 10010. Phone: (646) 312-3484, Fax: (646) 312-3451, E-mail: Ozgur_Demirtas@baruch.cuny.edu; Haim Levy is a professor of finance at the Department of Finance, Jerusalem School of Business Administration, The Hebrew University of Jerusalem, 91905, Israel. Phone: 972 2 5883101, Fax: 972 2 5881341, E-mail: mshlevy@mscc.huji.ac.il

Is There a Relation Between Downside Risk and Expected Stock Returns?

Abstract

This paper examines the intertemporal relation between downside risk and expected stock returns. Value at risk (VaR), expected shortfall, and tail risk are used as measures of downside risk to determine the existence and significance of a risk-return tradeoff for several stock market indices. We find a positive and significant relation between downside risk and the portfolio returns on the NYSE/AMEX/Nasdaq stocks. This result also holds for the NYSE/AMEX, NYSE, Nasdaq, and S&P 500 index portfolios. Moreover, VaR remains to be a superior measure of risk even when it is compared to the traditional risk measures which have significant predictive power for market returns. These results are robust across different measures of downside risk, loss probability levels, and after controlling for macroeconomic variables associated with business cycle fluctuations.

Introduction

The conditional mean and variance of return on the market portfolio play a central role in Merton's (1973) intertemporal capital asset pricing model (ICAPM). Although theoretical models suggest a positive relation between risk and return for the aggregate stock market, the existing empirical literature fails to agree on the intertemporal relation between expected return and volatility. There is a long literature that has tried to identify the existence of such a tradeoff between risk and return, but the results are far from being conclusive.[1]

This paper examines the intertemporal relation between downside risk and expected return on the market. Value at risk, expected shortfall, and tail risk are used as measures of downside risk to determine the existence and significance of a risk-return tradeoff for several stock market indices. There are several reasons why we consider downside risk in determining the existence of a positive risk-return tradeoff.

First, there is a long literature about safety-first investors who minimize the chance of disaster (or the probability of failure). The portfolio choice of a safety-first investor is to maximize expected return subject to a downside risk constraint. Roy's (1952), Baumol's (1963), Levy and Sarnat's (1972) and Arzac and Bawa's (1977) safety-first investor uses a downside risk measure which is a function of VaR. Roy (1952) indicates that most investors are principally concerned with avoiding a possible disaster and that the principle of safety plays a crucial role in the decision-making process. Roy (1952, p.432) states that:

> *"Decisions taken in practice are less concerned with whether a little more of this or of that will yield the largest net increase in satisfaction than with avoiding known rocks of uncertain position or with deploying forces so that, if there is an ambush round the next corner, total disaster is avoided. If economic survival is always taken for granted, the rules of behavior applicable in an uncertain and ruthless world cannot be covered."*

Thus, the idea of a disaster exists and a risk averse safety-first investor will seek to reduce as far as is possible the chance of such a catastrophe occurring.

Second, we believe optimal portfolio selection under limited downside risk to be a practical problem. Even if agents are endowed with standard concave utility functions such that to a first-order approxi-

[1] See Ghysels, Santa-Clara and Valkanov (2004) and the references therein.

mation they would be mean-variance optimizers, practical circumstances often impose constraints that elicit asymmetric treatment of upside potential and downside risk.

Third, commercial banks, investment banks, insurance companies, and nonfinancial firms hold portfolios of assets that may include stocks, bonds, currencies, and derivatives. Each institution needs to quantify the amount of risk its portfolio may incur in the course of a day, week, month, or year. For example, a bank needs to assess its potential losses in order to put aside enough capital to cover them. Similarly, a company needs to track the value of its assets and any cash flows resulting from losses on its portfolio. In addition, credit-rating and regulatory agencies must be able to assess likely losses on portfolios as well, since they need to set capital requirements and issue credit ratings. These institutions can judge the likelihood and magnitude of potential losses on their portfolios using value at risk. Regulatory concerns require commercial banks to report a single number, the so-called VaR, which measures the maximum loss on their trading portfolio if the lowest 1% quantile return would materialize. Capital adequacy is judged on the basis of the size of this expected loss. Likewise, pension funds are often required by law to structure their investment portfolio such that the risk of underfunding is kept low, e.g., equity investment may be capped.

Fourth, there is a wealth of experimental evidence for loss aversion (see Markowitz (1952b) and Kahneman et al. (1990)). Other evidence is provided through consumption behavior. As Deaton (1991) shows, consumption responds asymmetrically to good and bad states. Similarly, within the mean-variance setup there is a range of returns such that consumption is too low for survival. Over this range, modeling risk by the expectation of squared or absolute returns (i.e., conditional variance or standard deviation) may not be useful. An alternative way may be to collapse all returns below the survival threshold as being equally risky.

Fifth, asset returns have been modeled in continuous time as diffusions by Black and Scholes (1973), as pure jump processes by Cox and Ross (1976), and as jump-diffusions by Merton (1976). The rationale usually given for describing asset returns as jump-diffusions is that diffusions capture frequent small moves, while jumps capture rare large moves. Carr, Geman, Madan, and Yor (2002) develop a continuous time model that allows for both diffusions and for jumps of both finite and infinite activity. They find that market index returns tend to be pure jump processes of infinite activity and finite variation, and thus the index return processes appear to have effectively diversified away any diffusion risk. They indicate that the jump components account for significant skewness levels that may statistically be either positive or negative but that risk-neutrally are negative. They report significantly greater skewness and kurtosis in the risk-neutral process than in the statistical process.[2] The results presented in Carr et al. suggest that extreme movements in stock returns can be interpreted as signal whereas the frequent small fluctuations can be viewed as noise which may not have power to explain time-series variation in excess market returns.

Finally, the mean-variance analysis developed by Markowitz (1952a, 1959) critically relies on two assumptions: either the investors have a quadratic utility or the asset returns are jointly normally distributed (see Levy and Markowitz (1979), Chamberlain (1983) and Berk (1997)). Both assumptions are not required, just one or the other: (1) If an investor has quadratic preferences, she cares only about the mean and variance of returns; and the skewness and kurtosis of returns have no effect on expected utility, i.e., she will not care, for example, about extreme losses. (2) Mean-variance optimization can be justified if the asset returns are jointly normally distributed since the mean and variance will completely describe the distribution. However, the empirical distribution of stock returns is typically skewed, peaked around the mode, and has fat-tails, implying that extreme events occur much more frequently than predicted by the normal distribution (see Post and Vliet (2004a)). Therefore, the traditional measures of market risk (e.g., variance or standard deviation) cannot be used to approximate the maximum likely loss that a firm can expect to lose under normal or highly volatile periods.[3]

Although the mean-variance criterion has been basis for many academic papers and has had significant impact on the academic and non-academic financial community, it is still subject to theoretical and empirical criticism (see, e.g., Post and Vliet (2004b) and Ang, Chen and Xing (2005)). Arditti (1967), Levy (1969), Arditti and Levy (1975) and Kraus and Litzenberger (1976) extend the standard portfolio theory to incorporate the effect of skewness on valuation. They present a three-moment model with unconditional skewness. Harvey and Siddique (2000) present an asset-pricing model with conditional co-skewness, where risk-averse investors prefer positively skewed assets to negatively skewed assets. Their results imply preference for positive skewness: investors should prefer stocks that are right-skewed to stocks that are left-skewed. Assets that decrease a portfolio's skewness (i.e., that make the portfolio returns more left-skewed) are less desirable and should command higher expected

[2] This implies that the significance of skewness and kurtosis of stock returns found in empirical distributions is not only a time-series statistical property, but it exists in risk-neutral distributions as well.

[3] Longin (2000) and Bali (2003) find that VaR provides good predictions of catastrophic market risks, and performs surprisingly well in capturing both the rate of occurrence and the extent of extreme events in financial markets. However, the traditional measures of market risk such as the conditional variance and standard deviation yield an inaccurate characterization of extreme movements in financial markets.

returns. Similarly, assets that increase a portfolio's skewness should have lower expected returns. Dittmar (2002) extends the three-moment asset-pricing model using the restriction of decreasing absolute prudence (see Pratt and Zeckhauser (1987) and Kimball (1993)). He examines the co-kurtosis coefficient, and argues that investors with decreasing absolute prudence dislike co-kurtosis. His findings suggest preference for lower kurtosis. Investors are averse to kurtosis, and prefer stocks with lower probability mass in the tails of the distribution to stocks with higher probability mass in the tails of the distribution. Assets that increase a portfolio's kurtosis (i.e, that make the portfolio returns more leptokurtic) are less desirable and should command higher expected returns. Similarly, assets that decrease a portfolio's kurtosis should have lower expected returns.

Since the magnitude of VaR becomes larger for negatively skewed and thicker-tailed asset distributions, the findings of the three-moment and four-moment asset-pricing models indicate a positive relation between value at risk and expected stock returns, i.e, the more a market index can potentially fall in value the higher should be the expected return.

We consider VaR as a reliable measure of downside market risk and investigate the presence and significance of a positive risk-return tradeoff over the sample period of July 1962 to December 2002. We find a positive and significant relation between VaR and the value-weighted and equal-weighted portfolio returns on the NYSE/AMEX/Nasdaq stocks. This result also holds for the NYSE/AMEX, NYSE, Nasdaq, and S&P 500 index portfolios.

Ghysels, Santa-Clara, and Valkanov (2004) have recently shown that the choice of window size (from 1 month to 6 months) in the estimation of realized variance has tremendous impact on the significance of risk-return tradeoff. When they compute the realized variance as the sum of squared daily returns over the past one month, they find no evidence of a significant link between realized variance and future market returns. However, they report a significantly positive relation between the excess market return and the realized variance obtained from the past 3 to 6 months of daily data. In this paper, confirming their findings, we also obtain a significantly positive relation between the excess market return and the realized variance for window sizes larger than 1 month. Finally, we compare the relative performance of various VaR and realized variance measures computed over different horizons in predictive regressions. VaR remains to be a superior measure of risk even when it is compared to the traditional risk measures which have significant predictive power for market returns. These results are robust across different loss probability levels and after controlling for macroeconomic variables associated with business cycle fluctuations.

As an alternative measure of downside risk, we also consider expected shortfall and tail risk, which measure the mean and variance of losses beyond some value at risk level, respectively. We show that the strong positive relation between downside risk and excess market return is robust across different left-tail risk measures. Furthermore, to accommodate skewness and leptokurtosis in the empirical return distribution and more accurately identify the conditional mean of index returns, we model the time-series variation in monthly returns using the skewed t distribution. The parameter estimates from the maximum likelihood methodology based on the skewed t density reiterate the central finding of the paper that there exists a positive and significant relation between downside risk and expected returns.

The paper is organized as follows. Section I presents a framework that relates value at risk to expected returns. Section II presents alternative measures of market risk, and describes our investigation of the risk-return tradeoff. Section III presents the descriptive statistics of the data. Section IV discusses the empirical results from time-series regressions. Section V runs a battery of robustness checks. Section VI concludes the paper.

I Economic Framework

The standard theory of portfolio choice determines the optimum asset mix by maximizing (1) the expected risk premium per unit of risk in a mean-variance framework or (2) the expected value of a utility function approximated by the expected return and variance of the portfolio. In both cases, market risk of the portfolio is defined in terms of the variance (or standard deviation) of the portfolio's returns. Modeling portfolio risk with the traditional volatility measures implies that investors are concerned only about the average variation (and co-variation) of individual stock returns, and they are not allowed to treat the negative and positive tails of the return distribution separately.

In what follows, we consider an investor who allocates her portfolio in order to maximize the expected utility of end-of-period wealth $U(W)$. We assume that the distribution of returns on the investor's portfolio of risky assets is nonsymmetrical and fat-tailed. The expected value of end-of-period wealth can be written as $\overline{W} = \sum_{i=1}^{n} q_i \overline{R}_i + q_f R_f$, where $\overline{R}_i$ is unity plus the expected rate of return on the i th risky asset, R_f is unity plus the rate of return on the riskless asset, q_i is the fraction of wealth allocated to the i^{th} risky asset, and q_f is the fraction of wealth allocated to the riskless asset. Since our objective is to measure the effect of higher moments on the standard asset pricing models, we now approximate the expected utility by a Taylor series expansion around the expected wealth. For this purpose, the utility function is expressed in terms of the wealth distribution, so that $E[U(W)] = \int U(W)f(W)dW$, where $f(W)$ is the probability density function of the end-of-period wealth, that depends on the multivariate distribution of returns and on the vector of weights q. We now consider the infinite-order Taylor series expansion of the utility function

$$U(W) = \sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})(W - \overline{W})^k}{k!} \tag{1}$$

where $\overline{W} = E(W)$ denotes the expected end-of-period wealth. Under rather mild conditions (see Loistl (1976)), the expected utility is given by:

$$E[U(W)] = E\left[\sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})(W - \overline{W})^k}{k!}\right] = \sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})}{k!} E\left[(W - \overline{W})^k\right]. \tag{2}$$

Therefore, the expected utility depends on all central moments of the distribution of the end-of-period wealth.

It should be noticed that the approximation of the expected utility by a Taylor series expansion is related to the investor's preference (or aversion) towards all moments of the distribution, that are directly given by derivatives of the utility function. Scott and Horvath (1980) indicate that, under the assumptions of positive marginal utility, decreasing absolute risk aversion at all wealth levels together with strict consistency for moment preferences, one has $U^{(k)}(W) > 0$, $\forall W$ if k is odd and $U^{(k)}(W) > 0$, $\forall W$ if k is even. Further discussion on the conditions that yield such moment preferences or aversion can be found in Pratt and Zeckhauser (1987), Kimball (1993), and Dittmar (2002). Focusing on terms up to the fourth one, we obtain:

$$\begin{aligned} E[U(W)] &= U(\overline{W}) + U^{(1)}(\overline{W})E\left[(W - \overline{W})\right] + \frac{1}{2}U^{(2)}(\overline{W})E\left[(W - \overline{W})^2\right] \\ &\quad + \frac{1}{3!}U^{(3)}(\overline{W})E\left[(W - \overline{W})^3\right] + \frac{1}{4!}U^{(4)}(\overline{W})E\left[(W - \overline{W})^4\right] + O(W^4), \end{aligned} \tag{3}$$

where $O(W^4)$ is the Taylor remainder. We define the expected return, variance, skewness, and kurtosis of the end-of-period return, R_p, as

$$\mu_p = E[R_p] = \overline{W} \tag{4}$$

$$\sigma_p^2 = E\left[(R_p - \mu_p)^2\right] = E\left[(W - \overline{W})^2\right] \tag{5}$$

$$s_p^3 = E\left[(R_p - \mu_p)^3\right] = E\left[(W - \overline{W})^3\right] \tag{6}$$

$$k_p^4 = E\left[(R_p - \mu_p)^4\right] = E\left[(W - \overline{W})^4\right]. \tag{7}$$

Hence, the expected utility is simply approximated by the following preference function:

$$E[U(W)] \approx U(\overline{W}) + \frac{1}{2}U^{(2)}(\overline{W})\sigma_p^2 + \frac{1}{3!}U^{(3)}(\overline{W})s_p^3 + \frac{1}{4!}U^{(4)}(\overline{W})k_p^4. \tag{8}$$

Under conditions established by Scott and Horvath (1980), the expected utility depends positively on expected returns and skewness and negatively on variance and kurtosis (Also see Berkelaar, Kouwenberg, and Post (2004)). Based on the CARA and CRRA utility functions, we now show that an increase in VaR reduces the expected utility of wealth.

We first consider the CARA (Constant Absolute Risk Aversion) utility function. The CARA utility function is defined by: $U(W) = -\exp(-\theta W)$, where θ measures the investor's constant absolute risk aversion. The approximation for the expected utility is given by

$$E[U(W)] \approx -\exp(-\theta\overline{W})\left[1 + \frac{\theta^2}{2}\sigma_p^2 - \frac{\theta^3}{3!}s_p^3 + \frac{\theta^4}{4!}k_p^4\right]. \tag{9}$$

Equation (9) indicates aversion to variance and kurtosis and preference for (positive) skewness since $\frac{\partial E[U(W)]}{\partial \sigma_p^2} = -\exp(-\theta\overline{W})\frac{\theta^2}{2} < 0$, $\frac{\partial E[U(W)]}{\partial s_p^3} = \exp(-\theta\overline{W})\frac{\theta^3}{3!} > 0$, and $\frac{\partial E[U(W)]}{\partial k_p^4} = -\exp(-\theta\overline{W})\frac{\theta^4}{4!} < 0$.

These results imply aversion to value at risk (VaR), $\frac{\partial E[U(W)]}{\partial VaR_p} < 0$, since VaR for long positions (defined by the left tail of the return distribution) increases with variance and kurtosis and decreases with positive skewness.

Similar results are obtained from the CRRA (Constant Relative Risk Aversion) utility function given by: $U(W) = \frac{W^{1-\theta}}{1-\theta}$, $(\theta > 0, \theta \neq 1)$. Since $\theta > 0$, the expected utility of wealth decreases with variance and kurtosis, whereas it increases with positive skewness, i.e., $\frac{\partial E[U(W)]}{\partial \sigma_p^2} = -\frac{\theta}{2}\overline{W}^{-(1+\theta)} < 0$, $\frac{\partial E[U(W)]}{\partial s_p^3} = -\frac{\theta(1+\theta)}{3!}\overline{W}^{-(2+\theta)} > 0$, and $\frac{\partial E[U(W)]}{\partial k_p^4} = -\frac{\theta(1+\theta)(2+\theta)}{4!}\overline{W}^{-(3+\theta)} < 0$.

Since $\frac{\partial VaR_p}{\partial \sigma_p^2} > 0$, $\frac{\partial VaR_p}{\partial s_p^3} < 0$, and $\frac{\partial VaR_p}{\partial k_p^4} > 0$ investors dislike VaR, i.e., an increase in VaR reduces the expected utility of wealth, $\frac{\partial E[U(W)]}{\partial VaR_p} < 0$.

In summary, since VaR is a function of higher-order moments of the return distribution (variance, skewness, and kurtosis) in a certain way, investors have an aversion to VaR. Consequently, there is an implied positive relation between VaR of a portfolio and the portfolio's expected return.

II Measuring Risk-Return Relationship

A Alternative Risk Measures

Realized variance: Following French et al. (1987) and Ghysels et al. (2004), we calculate the variance of a market portfolio using various window sizes of return data:

$$\sigma_{k,t}^2 = \sum_{d=1}^{D_k} r_{k,d}^2 + 2\sum_{d=2}^{D_k} r_{k,d} \cdot r_{k,d-1} \tag{10}$$

where $\sigma_{k,t}^2$ is the variance of index returns, D_k is the number of trading days over the past k months[4], and $r_{k,d}$ is the portfolio's return on day d which resides within k months. The second term on the right hand side adjusts for the autocorrelation in daily returns using the approach of French et al. (1987). Note that the realized variance measure given in equation (10) is not, strictly speaking, a variance measure since daily returns are not demeaned before taking the expectation. However, as pointed out by French et al. (1987) and Goyal and Santa-Clara (2003), the impact of subtracting the means is trivial for short holding periods.

Nonparametric Value-at-Risk: VaR determines how much the value of a portfolio could decline over a given period of time with a given probability as a result of changes in market rates. For example, if the given period of time is one day and the given probability is 1%, the VaR measure would be an estimate of the decline in the portfolio value that could occur with a 1% probability over the next trading day. In other words, if the VaR measure is accurate, losses greater than the VaR measure should occur less than 1% of the time. In this paper, we use different confidence levels to check the robustness of VaR measures as an explanatory variable for the expected return on the market. The estimation is based on the lower tail of the actual empirical distribution. We use 100 daily returns to estimate the 1% VaR level from the empirical distribution. The 1% VaR is defined as the minimum index return observed during the last 100 days as of the end of month t.[5] It should be noted that the original VaR measures are multiplied by -1 before running our regressions. The original maximum likely loss values are negative since they are obtained from the left tail of the distribution, but the downside risk measure, VaR_t, used in our regressions is defined as $-1\times$the maximum likely loss. Therefore, the slope coefficients turn out to be positive, which gives the central result of the paper that there is a positive

[4] As in Ghysels, Santa-Clara and Valkanov (2004), we use 1 month to 6 months of past daily data to compute the rolling window variance estimates.

[5] We also use the past one month to six months of daily returns to estimate alternative VaR measures from the empirical distribution.

and statistically significant relation between VaR and the excess return on the market, i.e, the more a market index can potentially fall in value the higher should be the expected return.

Parametric Value-at-Risk: In continuous time diffusion models, (log)-stock price movements are described by the following stochastic differential equation,

$$d \ln P_t = \mu dt + \sigma dW_t \tag{11}$$

where W_t is a standard Wiener process with zero mean and variance of dt, μ and σ are the drift and diffusion parameters of the geometric Brownian motion, respectively. In discrete time, equation (11) yields a return process:

$$\ln P_{t+\Delta} - \ln P_t = R_t = \mu\Delta t + \sigma z\sqrt{\Delta t} \tag{12}$$

where Δt is the length of time interval in which the discrete time data are recorded and $\Delta W_t = z\sqrt{\Delta t}$ is the Wiener process with zero mean and variance of Δt since z is a random variable drawn from the standard normal density, i.e., $E(z) = 0$ and $E(z^2) = 1$.

The critical step in calculating VaR measures is the estimation of the threshold point defining what variation in returns R_t is considered to be extreme. Let Φ be the probability that R_t is less than the threshold Γ. That is,

$$\Pr(R_t < \Gamma) = \Pr\left(z < a = \frac{\Gamma - \mu\Delta t}{\sigma\sqrt{\Delta t}}\right) = \Phi \tag{13}$$

where $\Pr(.)$ is the underlying probability distribution. In the traditional VaR models with the assumption of normality we have $\Phi = 1\%, a = -2.326$,

$$\Gamma_{Normal} = \mu\Delta t - 2.326\sigma\sqrt{\Delta t}. \tag{14}$$

However, there is substantial empirical evidence showing that the distribution of financial returns is typically skewed to the left, is peaked around the mean (leptokurtic) and has fat tails. The fat tails and negative skewness suggest that extreme outcomes happen much more frequently than would be predicted by the normal distribution, and the negative returns of a given magnitude have higher probabilities than positive returns of the same magnitude. This also suggests that the normality assumption can produce VaR numbers that are inappropriate measures of the true risk faced by individual firms. To account for skewness and excess kurtosis in the data, we use the skewed t distribution of Hansen (1994) that accounts for the non-normality of returns and relatively infrequent events.

Hansen (1994) introduces a generalization of the Student t distribution where asymmetries may occur, while maintaining the assumption of a zero mean and unit variance. The skewed t density that provides a flexible tool for modeling the empirical distribution of stock market returns exhibiting skewness and leptokurtosis is given by equation (15):

$$f(z_t;\mu,\sigma,\upsilon,\lambda)=\begin{cases} bc\left(1+\frac{1}{\upsilon-2}\left(\frac{bz_t+a}{1-\lambda}\right)^2\right)^{-\frac{\upsilon+1}{2}} & if\ z_t<-\frac{a}{b} \\ bc\left(1+\frac{1}{\upsilon-2}\left(\frac{bz_t+a}{1+\lambda}\right)^2\right)^{-\frac{\upsilon+1}{2}} & if\ z_t\geqslant-\frac{a}{b}\end{cases} \tag{15}$$

where $z_t=\frac{R_t-\mu}{\sigma}$ is the standardized market return, the constants a, b, and c are given by

$$a=4\lambda c\left(\frac{\upsilon-2}{\upsilon-1}\right),\ b^2=1+3\lambda^2-a^2,\ c=\frac{\Gamma\left(\frac{\upsilon+1}{2}\right)}{\sqrt{\pi(\upsilon-2)}\Gamma\left(\frac{\upsilon}{2}\right)} \tag{16}$$

Hansen shows that this density is defined for $2<\upsilon<\infty$ and $-1<\lambda<1$. This density has a single mode at $-a/b$, which is of opposite sign with the parameter λ. Thus, if $\lambda>0$, the mode of the density is to the left of zero and the variable is skewed to the right, and vice versa when $\lambda<0$. Furthermore, if $\lambda=0$, Hansen's distribution reduces to the standardized t distribution. If $\lambda=0$ and $\upsilon=\infty$, it reduces to a normal density.[6]

A parametric approach to calculating VaR is based on the lower tail of the skewed t distribution. Specifically, we estimate the parameters of the skewed t density (μ, σ, υ, λ) using 100 daily returns and then find the one percentile of the estimated distribution. Assuming that $R_t\sim f_{\upsilon,\lambda}(z)$ follows a skewed t (ST) density, parametric VaR is the solution to:

$$\int_{-\infty}^{\Gamma_{ST}(\Phi)} f_{\upsilon,\lambda}(z)dz=\Phi \tag{17}$$

where $\Gamma_{ST}(\Phi)$ is the VaR threshold based on the skewed t density with a loss probability of Φ. Equation (17) indicates that value at risk can be calculated by integrating the area under the probability density function of the skewed t distribution.

[6] The parameters of the skewed t density are estimated by maximizing the log-likelihood function of R_t with respect to the parameters μ, σ, υ, and λ:

$$LogL=n\ln b+n\ln\Gamma\left(\frac{\upsilon+1}{2}\right)-\frac{n}{2}\ln\pi-n\ln\Gamma(\upsilon-2)-n\ln\Gamma\left(\frac{\upsilon}{2}\right)-n\ln\sigma-\left(\frac{\upsilon+1}{2}\right)\sum_{t=1}^{n}\ln\left(1+\frac{d_t^2}{(\upsilon-2)}\right)$$

where $d_t=\frac{bz_t+a}{(1-\lambda s)}$ and s is a sign dummy taking the value of 1 if $bz_t+a<0$ and s=-1 otherwise.

B Time-Series Regressions

We investigate the intertemporal relation between downside risk and excess market return at the monthly frequency. The downside risk-return relationship we analyze in the paper takes the following form:

$$R_{t+1} = \alpha + \beta E_t(VaR_{t+1}) + \varepsilon_{t+1}, \tag{18}$$

where R_{t+1} is the monthly excess return of the market portfolio, $E_t(VaR_{t+1})$ is the conditional value at risk of the market portfolio obtained from the daily index returns, and ε_{t+1} is the residual term. We use various measures of the lagged realized VaR as a proxy for the expected conditional downside risk for the current period.[7] The slope coefficient β in equation (18) is expected be positive and statistically significant.

We also test the usual form of the risk-return tradeoff by examining whether the relation between the conditional variance and the expected excess return is positive. We use the following discrete-time specification of Merton (1980):

$$R_{t+1} = \alpha + \gamma E_t(\sigma^2_{t+1}) + \varepsilon_{t+1}, \tag{19}$$

where the coefficients α and γ according to Merton's ICAPM, should be zero and equal to the relative risk aversion coefficient, respectively. Positive values of γ imply the existence of a risk-return tradeoff, indicating that the expected returns are higher as the risk level for the market increases. Following Ghysels, Santa-Clara and Valkanov (2004), we use various measures of the lagged realized variance as a proxy for $E_t(\sigma^2_{t+1})$.

Campbell (1987) and Scruggs (1998) point out that the approximate relationship in equation (19) may be misspecified if the hedging term in ICAPM is important. To make sure that our results from estimating equations (18) and (19) are not due to model misspecification, we added to the regressions a set of control variables that have been used in the literature to capture the state variables that determine changes in the investment opportunity set. Specifically, we use a set of macroeconomic variables proxying for business cycle fluctuations, the lagged excess return and a dummy variable for October 1987 crash.

[7] As discussed later in the paper, we conduct robustness checks where VaR measures that conditionally change over time are used in regressions. The results from the lagged realized VaR and from the conditional forecasts of VaR are found to be similar.

III Data

To capture the U.S. stock market returns, we use the monthly returns on the NYSE/AMEX/Nasdaq index. As a robustness check we also repeat our analysis for NYSE/AMEX, NYSE, Nasdaq, and S&P 500 indices. These index returns are available from July 1962 to December 2002, except for the Nasdaq sample that covers the period from January 1973 to December 2002. In predictive regressions, we use the excess market return defined as the difference between the index return and the risk free rate. We use the one-month Treasury bill return as the risk free rate, which is available on Kenneth French's online data library.

Panel A of Table 1 provides descriptive statistics for the value-weighted index returns. Panel A shows that the average monthly return is in the range of 0.94% for the NYSE/AMEX/Nasdaq and 0.99% for the Nasdaq, which correspond to annualized returns of 11.28% and 11.88%, respectively. The unconditional standard deviations of monthly returns are in the range of 4.3% for the NYSE and 6.7% for the Nasdaq index. The skewness and kurtosis statistics are reported for testing the distributional assumption of normality. The skewness statistics for monthly returns are negative and significant at the 1% level. The kurtosis statistics are greater than three and statistically significant at the 1% level. Furthermore, the Jarque-Bera statistics strongly reject the distributional assumption of normality.[8]

Panel B shows summary statistics for value at risk computed using rolling window estimation over various months denoted by k. VaR is defined as -1 times the minimum NYSE/AMEX/Nasdaq index return observed during the last k months of daily data as of the end of each month t. Furthermore, each month is assumed to have 21 trading days. For example, value at risk for the past 4 months is computed as the lowest return observed during the last 84 days. Observe that the distributions of various VaRs are skewed to the right and have fatter tails than the normal distribution. Moreover, the first-order autoregressive coefficients are high indicating persistency in VaR measures.

A series of papers argue that the stock market can be predicted by macroeconomic variables associated with business cycle fluctuations. The commonly chosen variables include default spread (DEF_t), term spread (TERM_t), dividend price ratio (DP_t), and the detrended riskless rate (RREL_t). We

[8] Jarque-Bera, $JB = n[(S^2/6) + (K-3)^2/24]$, is a formal test statistic for testing whether the returns are normally distributed, where n denotes the number of observations, S is skewness and K is kurtosis. The test statistic distributed as Chi-square with two degrees of freedom measures the difference of the skewness and kurtosis of the series with those from the normal distribution.

investigate how incorporating these variables into the predictive regressions affects the intertemporal relation between downside risk and expected stock returns. We define DEF_t as the difference between the yields on BAA- and AAA-rated corporate bonds, TERM_t as the difference between the yields on the 10-year Treasury bond and the three-month Treasury bill. RREL_t is defined as the difference between 3-month T-bill rate and its 12-month backward moving average.[9]

IV Empirical Results

A Downside Risk and Expected Returns

In Table II, we present the first set of empirical results from the time-series regressions of the value-weighted excess market return of NYSE/AMEX/Nasdaq index on nonparametric and parametric value at risk (VaR_t and VaR_t^p). At this stage, our downside risk measure is 1% VaR computed using the returns observed over the past 100 days. The dependent variable is the one-month ahead value-weighted excess market return, R_{t+1}. The independent variables are the 1% VaR, a dummy variable that takes the value of one in October 1987 and zero otherwise, the lagged excess market return, R_t, and the macroeconomic variables. For each parameter estimate we present the Newey-West (1987) adjusted t-statistic in parenthesis.

Panel A of Table II uses the nonparametric VaR. The first regression shows that the coefficient on VaR_t is positive and 1.92 standard deviations away from zero. When the lagged return and the dummy is added to the regression, the coefficient estimate of VaR_t gets higher and its t-statistic becomes 2.91. Note that the coefficient estimate of the lagged return is not significant. This is not surprising given the low autocorrelation in the value-weighted index returns. The dummy variable, on the other hand, controls for the October 1987 crash. The coefficient on dummy variable is negative and highly significant indicating that there would be a specification error if we have not used it as a right hand side variable.[10,11]

[9] The time series data on monthly three-month T-bill rates, 10-year Treasury bond yields, and BAA- and AAA-rated corporate bond yields are available from the Federal Reserve statistics release website. We obtain the dividend price ratio from Robert Shiller's website: http://aida.econ.yale.edu/~shiller/. The data are available from January 1871 to June 2002. When analyzing the role of the dividend price ratio, we use the common sample from July 1962 to June 2002.

[10] To check whether our results are driven by extreme return or VaR realizations, we also exclude the observations lower (higher) than 1 percentile (99 percentile) of the return and VaR distributions and re-run our time-series regressions. The t-statistics of the estimated coefficients on VaR with the new sample turn out to be even higher. Thus, we conclude that the main findings of this paper are not driven by outliers.

[11] We also repeat our analysis by eliminating the month of October 1987. Since the qualitative results turn out to be very similar to those reported in our tables, we do not present them here. They are available upon request.

We further investigate the relation between downside risk and expected returns after controlling for macroeconomic variables known to forecast the stock market. As shown by Merton (1973), and subsequently pointed out by Campbell (1987) and Scruggs (1998), equation (19) omits the hedging component that captures the investor's motive to hedge for future investment opportunities. Thus, we include macroeconomic variables that have been shown in the literature to capture state variables that determine the investment opportunity set. We control for term premia TERM_t and default premia DEF_t both of which are shown to covary with business conditions. We also control for the detrended riskless rate RREL_t and dividend price ratio DP_t, which are used to forecast market returns. The last regression in Panel A shows that even after using the macroeconomic control variables, VaR_t has a positive and significant coefficient indicating that there is a robust and significantly positive relationship between downside risk and expected returns. On the other hand, the R^2 values are small, in the range of 0.58% to 3.28%, but they are consistent with the earlier research (e.g., Goyal and Santa-Clara (2003), Ghysels et al. (2004) and Bali et al. (2005)).[12]

Panel B of Table II shows the parameter estimates from regressions of the value-weighted CRSP index return on the lagged VaR which is calculated parametrically based on the lower tail of the skewed t distribution, VaR_t^p. The magnitude and statistical significance of the slope coefficient turn out to be very similar to our findings in Panel A. Although there is marginal significance at the univariate level, after controlling for the autocorrelation in market returns and macroeconomic variables, VaR_t^p has a coefficient of 0.310 with a t-statistic of 2.67. Furthermore, as will be discussed in the next subsection, even the univariate regression shows strong statistical significance once the skewed t distribution is assumed to govern the error process.[13]

B Maximum Likelihood Estimation Using the Skewed t Distribution

To examine the intertemporal relation between downside risk and expected returns, following most of the previous studies, we have so far used the ordinary least square (OLS) estimation where the statistical significance of estimated coefficients is established using the Newey-West adjusted t-statistics. This estimation procedure implicitly assumes that the monthly returns on stock market indices follow

[12] See Appendix for parameter estimates obtained by using returns of other indices.

[13] At an earlier stage of the study, we have repeated our analyses for the equal-weighted indices. Stocks with smaller market capitalization are weighted more heavily in the equal-weighted index than in the value-weighted index. We find VaR to be even more powerful in forecasting the future equal-weighted returns because the return distribution of small stocks exhibits higher peaks, fatter tails and more outliers on the left or right tail than the distribution of bigger stocks. We report conservative results in our tables. The findings from the equal-weighted indices are available upon request.

a normal distribution. However, the residuals from the time-series regressions are found to be non-Gaussian: the empirical distribution of ε_{t+1} exhibits skewness and excess kurtosis, and the fat-tail property is more dominant than skewness in the sample. This implies that the estimated conditional mean, $\alpha + \beta VaR_t$, of equation (18) based on the normal distribution may not reflect the true value and thus the magnitude and statistical significance of the slope coefficient, β, may not be accurate.

To accommodate skewness and leptokurtosis in the empirical return distribution and more accurately identify the conditional mean of index returns, we model the time-series variation in monthly returns using the skewed t distribution of Hansen (1994) given in equation (15). Then, we estimate the intertemporal relation between VaR and excess market return using the maximum likelihood methodology described in Section II.A. The maximum likelihood estimation of the risk-return tradeoff yields the intercept (α) and the slope coefficient (β) in the conditional mean equation as well as the standard deviation (σ), skewness (λ), and tail-thickness (v) parameters of the skewed t density.

Panel A of Table III presents the maximum likelihood parameter estimates and the asymptotic t-statistics when the nonparametric value at risk is used to explain the excess return on the CRSP value-weighted index. With or without the lagged return and dummy variable, the slope coefficient is found to be in the range of 0.25 to 0.27 and highly significant. Specifically, the asymptotic t-statistics obtained from the maximum likelihood estimation are in the range of 2.49 to 2.59. When macroeconomic control variables are added to the regression, the coefficient estimate of VaR slightly decreases to 0.23, but still significant with a t-statistic of 2.14.[14]

Panel B of Table III presents the maximum likelihood parameter estimates and the asymptotic t-statistics when the parametric VaR is used as a measure of downside risk. Again, we find similar set of results. VaR has a positive and significant coefficient in all three regressions. The coefficient estimates range between 0.239 to 0.279, where the t-statistics are in the range of 2.16 to 2.27. Overall, we conclude that the parametric VaR is as good as the nonparametric VaR in terms of predicting the excess market return.

The parameter estimates in Table III reiterate the central result of the paper that the more a market index can potentially fall in value the higher should be the expected return. Furthermore, A notable point in Table III is that the estimated coefficient on dummy variable is not significant, indicating that the October 1987 crash is captured properly by the skewed fat-tailed distribution.

[14] Also note that the skewness (λ) parameter is estimated to be negative and highly significant, and the tail-thickness (v) parameter is finite (i.e., $\frac{1}{v} \neq 0$) and highly significant, confirming our projections of a non-normal return distribution.

Another notable point is that the magnitude of the coefficient estimates on VaR presented in Table III, is very robust and it is close to the parameter estimates reported in the last two rows of Panels A-B of Table II. These results indicate that the OLS estimation of the risk-return tradeoff actually yields similar results to the maximum likelihood estimation with skewed t density when the October 1987 crash dummy is included in OLS regressions.

C Various Measures of Realized Variance

As mentioned earlier, theoretical models suggest a positive relation between conditional mean and variance of returns for the aggregate stock market. One of the most commonly used estimator of the conditional variance is the sum of squared daily returns over the previous month (see French, Schwert, and Stambaugh (1987)). Although this measure of market variance have been used extensively in tests of risk-return tradeoff, there is no evidence of a positive and significant relation between this measure of conditional variance and expected returns.

Recently Ghysels, Santa-Clara and Valkanov (2004) present new and interesting evidence regarding the risk-return tradeoff. Like French, Schwert, and Stambaugh (1987), they use the rolling window approach and use the sum of squared daily return as a proxy for the monthly conditional variance. Additionally, they argue that since the realized variance is very persistent, it ought to be a good proxy for the conditional variance. On the other hand, they rightfully point out that it is not clear why the researchers should confine themselves to using data from the last one month only to estimate the conditional variance. Therefore, they use larger window size (from 1 month to 6 months) when they sum the past squared returns to acquire the conditional variance measure. Interestingly, this choice has a tremendous impact on the significance of risk-return tradeoff. Ghysels, Santa-Clara and Valkanov (2004) report that the variance measures which are computed by using the daily returns over the previous 3 to 6 months significantly forecast the market return. These findings are striking because they also confirm the findings of Ghysels et al. on MIDAS (Mixed Data Sampling) estimator of variance which is also found to be a statistically significant predictor of market returns. Since MIDAS is a weighted average of past squared returns, its forecasting power is effectively driven by the strength of rolling window estimates of market variance computed using the past 3 to 6 months of daily data. Thus, Ghysels, Santa-Clara and Valkanov interpret their rolling window approach as a robustness check of the MIDAS regressions since it is such a simple estimator of conditional variance with no parameters to estimate as in MIDAS.

These recent findings are important for us because we would like to compare our measure of downside risk with the traditional risk measures which are shown to have a statistically significant predictive power for the expected market returns. Thus, in the light of these recent findings, we have to focus on realized variance that is computed using larger than 1-month window size. Simply because the realized variance in the previous month is not a significant predictor of expected returns.

In Table IV, using our sample from 1962 to 2002, we reexamine the findings of Ghysels, Santa-Clara and Valkanov (2004) on the rolling window estimates. We assume that each month has 21 trading days and compute the realized variance as the sum of squared daily returns on the value weighted NYSE/AMEX/Nasdaq index plus an adjustment term for the first-order serial correlation in daily returns. We generate different variance measures for horizons of 1 month to 6 months (i.e., past 21 days to 126 days). Panel A of Table IV present parameter estimates from time-series regressions. The first column shows the number of months used in the estimation of the conditional variance proxy. Similar to the very early literature, we find that the realized variance in the previous month has no forecasting power, but starting from month 2 going through month 6, realized variance is a significant forecaster of market returns. Indeed at 4-month horizon, the variance has a positive coefficient estimate of 0.538 with a t-statistic of 2.78. In Panel B we find similar evidence when we control for the macroeconomic variables. Although at 6-month horizon the variance loses its significance, the t-statistics range from 1.92 to 2.85 for all the other horizons.[15] Thus, we conclude that both the significance levels and patterns are very similar to the findings of Ghysels, Santa-Clara and Valkanov (2004), such that the 1-month variance estimator is not significant in regressions, whereas the rolling window variance estimators computed using larger window sizes turn out to be significant in tests of risk-return tradeoff.[16]

D Value at Risk versus Realized Variance

We have so far used the 1% VaR as a measure of downside risk. In this section, we first generate alternative measures of VaR based on the past 1 month to 6 months of daily returns, and then investigate the predictive power of these alternative VaR measures in forecasting future market returns.

[15] Similar to the findings of Ghysels et al., we find that as the window size increases beyond 6 months (not shown in the table), the magnitude and statistical significance of the slope coefficient decrease. This suggests that there is an optimal window size to estimate the risk-return tradeoff if one would like to use the lagged realized variance as a proxy for the current conditional variance.

[16] Note that our coefficient estimates are not exactly comparable to those of Ghysels, Santa-Clara and Valkanov (2004) simply because we do not convert our variance measures to monthly figures before we use them as RHS variables.

Finally, we compare the relative performance of various VaR and realized variance measures computed over different horizons in predictive regressions.

In Table V, we present the estimates of risk-return tradeoff using the rolling window estimators of VaR. The first column in each panel of Table V shows the number of months used to compute VaR. As before, we assume that each month has 21 trading days. Therefore, at 1-month horizon, VaR is defined as the minimum daily return observed during the past 21 days, hence it corresponds to 4.76% VaR. At 2-month horizon, VaR is defined as the minimum daily return observed during the past 42 days, hence it can be viewed as 2.38% VaR. Similarly, at 5-month horizon, VaR is defined as the minimum daily return observed during the past 105 days, hence it is 0.95% VaR measure etc.

Panel A of Table V shows that VaR has a positive and statistically significant coefficient estimate at every horizon. Similar to our findings for variance estimators, statistical significance varies with the window used to estimate VaR. When VaR is computed using the past 1 month of daily data, the t-statistic of VaR is 1.98, whereas it goes up to 4.12 at 4-month horizon. Although not presented in the paper to save space, the statistical significance of the coefficient of VaR survives up to 9 months. Another interesting point in Panel A of Table V is that, not only VaR is significant at all horizons (which is not the case for rolling window estimators of variance), but also the t-statistics and R-squares from these regressions are much higher than the corresponding statistics in Panel A of Table IV. For example, at 4-month horizon, VaR together with the lagged return and dummy variable explains 2.55% of the monthly return variation whereas this ratio is only 1.82% when we use variance as the risk proxy. In Panel B of Table V, we use additional control variables and find similar evidence. Although the significance and explanatory power of VaR vary with the window size, VaR is significant at all horizons and the control variables do not affect our main findings.

We have so far shown that there is a significant relation between VaR and expected stock returns. However, our ultimate goal is to compare the predictive power of VaR with the predictive power of traditional risk proxies which are shown to forecast market returns. In Panel A of Table VI, for horizons 1 month to 6 months, we compute VaR and variance measures and use them in the same regressions. For example, at 2-month horizon, we compute VaR as the minimum return observed during the past 42 days and we compute the variance as the sum of squared daily returns during the past 42 days plus the autocorrelation adjustment term. We further include the control variables to run a full specification. Observe that, at all horizons, VaR measure has a positive and significant coefficient estimate. Statistical significance decreases slightly, but the t-statistics are still in the range of 2.05 and

3.30 when the past 1 month to 4 months of daily returns are used in estimations. On the other hand, the rolling window estimate of variance is not significant at any horizon. Therefore, measuring variance by using longer than 1-month window size has a substantial effect on the risk-return tradeoff, but that impact is fully captured by VaR. We conclude that VaR is not only a good measure of downside risk which is related to expected returns, it also captures information about expected returns which cannot be explained by the traditional measure of market risk, realized variance, even if it is computed by using a larger number of observations.

It is important to note that the findings in Table IV suggest that there is an optimal window size to estimate the risk-return tradeoff if one would like to use the lagged realized variance as a proxy for the current conditional variance. Based on this evidence, Ghysels, Santa-Clara, and Valkanov (2004) develop a new volatility estimator, MIDAS, which is a weighted average of rolling window variance estimators. Specifically, MIDAS estimator is defined as $MIDAS_t = 22\sum_{d=1}^{260} w_d r_{t-d}^2$, where $w_d = \frac{\exp(k_1 d + k_2 d^2)}{\sum_{i=1}^{260} \exp(k_1 i + k_2 i^2)}$ and k_1 and k_2 are parameters in the weight function. Since MIDAS estimator is simply a weighted average of past squared daily returns, we expect a significant relation between MIDAS and expected returns in the light of our findings in Table IV.

In Panel B of Table VI, we first generate the MIDAS estimator using our sample from 1962 to 2002 and test its performance in predicting excess market returns. Observe that the coefficient on MIDAS is highly significant with a t-statistic of 2.68. This t-statistic compares well with the findings of Ghysels, Santa-Clara, and Valkanov (2004) who report a t-statistic of 2.64 for the MIDAS estimator during the period 1946 to 2000. Our next goal is compare the predictive power of our VaR estimate with the predictive power of MIDAS. Since MIDAS is a weighted average of rolling window variance estimators, we generate a weighted average VaR measure. Specifically, we form an equal weighted VaR measure (VaR^{ew}), calculated by equally weighting six value at risk measures computed by using windows of 1 month to 6 months. The second row of Panel B reports results using this equal weighted measure. As expected VaR^{ew} is highly significant with a t-statistic of 3.47. Finally, the third row of Panel B compares the predictive powers. Although MIDAS estimator is highly significant when used alone, our VaR measure is superior in the tests of risk-return tradeoff, such that VaR^{ew} has a t-statistic of 2.03 whereas MIDAS is insignificant. This result is not surprising given that VaR measures over various horizons win convincingly against rolling window estimates of variance that drives the forecasting power of MIDAS.

V Robustness Checks

A Conditional Value at Risk

As mentioned earlier, we approximate the conditional value at risk by the lagged VaR, i.e., $E_t(VaR_{t+1}) = VaR_t$, to test the intertemporal relation between downside risk and excess market return as shown in equation (18). Similar approximations are used in the literature when the realized variance is used as a proxy for the conditional variance, simply because the realized variance is very persistent. Our approximation is also justified by the fact that the VaR is highly persistent. For example, as mentioned before, the VaR obtained from the past 5 months of daily data, which roughly corresponds to a 1% VaR, has a first-order autocorrelation, AR(1), coefficient of 0.82 (see Panel B of Table I). However, the measures of VaR are supposed to be changing conditionally over time, hence we have to use more accurate conditional measures of downside risk. In this section, we use two different methods to obtain conditional time-varying VaR.

The first method generates conditional VaR based on the conditional mean and conditional standard deviation of market returns along with the skewness and tail-thickness parameters of a skewed fat-tailed density. In this framework, we assume that the discrete time version of the geometric Brownian motion governing financial price movements is:

$$R_t = \ln(P_{t+\Delta t}) - \ln(P_t) = \mu^* \Delta t + z\sigma^* \sqrt{\Delta t}, \tag{20}$$

where P_t is the price level at time t, R_t is the log-return from time t to $t + \Delta t$, Δt is the length of time interval between two successive prices, μ^* and σ^* are the annualized mean and standard deviation of R_t, and $\Delta W_t = z\sqrt{\Delta t}$ is a discrete approximation of the Wiener process. In the case of time-varying mean and variance, equation (20) can be modified to reflect these dependencies as:

$$R_t = \mu_t^* \Delta t + z\sigma_t^* \sqrt{\Delta t}, \tag{21}$$

where μ_t^* and σ_t^* are annualized measures of conditional mean and conditional standard deviation of R_t. For simplicity, equation (21) can be rewritten as $R_t = \mu_t + z\sigma_t$, where $\mu_t = \mu_t^* \Delta t$ and $\sigma_t = \sigma_t^* \sqrt{\Delta t}$ are, respectively, the conditional mean and conditional standard deviation of returns. Note that the standardized return $z = \frac{R_t - \mu_t}{\sigma_t}$ preserves the properties of zero mean and unit variance.

The conditional VaR threshold for R_t at a given coverage probability Φ, denoted by θ_t, is obtained from the solution of the following cumulative distribution of returns,

$$\Pr(R_t \leq \theta_t \mid \Omega_{t-1}) = \int_{-\infty}^{\theta_t} f(R_t \mid \Omega_{t-1}) dR_t = \Phi, \tag{22}$$

where Pr(.) denotes the probability and $f(R_t \mid \Omega_{t-1})$ is the conditional probability density function for R_t. The above probability function can be written in terms of the standardized returns as follows:

$$\Pr(R_t \leq \theta_t \mid \Omega_{t-1}) = \Pr\left(\frac{R_t - \mu_t}{\sigma_t} \leq \frac{\theta_t - \mu_t}{\sigma_t} \mid \Omega_{t-1}\right) = \Pr\left(z \leq a = \frac{\theta_t - \mu_t}{\sigma_t}\right) = \int_{-\infty}^{a} f(z)dz = \Phi, \tag{23}$$

where the density $f(z)$ and the threshold a with the coverage probability Φ do not depend on the information set Ω_{t-1}. The latter is a by-product of the assumption that the series of standardized returns z is identically and independently distributed. The latter assumption is consistent with empirical evidence related to the GARCH models for stock returns (see, e.g., Bollerslev, Chou, and Kroner (1992) and Bollerslev, Engle, and Nelson (1994)).

Given the probability density function of standardized returns $f(z)$, the threshold a can be easily obtained from the solution of the equation $\int_{-\infty}^{a} f(z)dz = \Phi$. That is, by finding the numerical value of a that equalizes the area under $f(z)$ to the coverage probability Φ. For the skewed t density given in equation (15), the value of a is a function of the skewness and tail-thickness parameters (λ and υ).

Given the estimated threshold a, the conditional (time-varying) VaR for the returns can be computed using the equation

$$\theta_t = \mu_{t|t-1} + a\sigma_{t|t-1}, \tag{24}$$

where the values of $\mu_{t|t-1}$ and $\sigma_{t|t-1}$ are based on some GARCH specification with time-varying conditional mean and standard deviation. Equation (24) originally proposed by Bali and Theodossiou (2004) can be used to compute the conditional VaR thresholds for excess market returns.[17]

Earlier empirical analyses are based on a sound statistical theory, but do not yield VaR measures that reflect the current mean-volatility background. In light of the fact that serial correlation and conditional heteroscedasticity are present in most financial data, unconditional VaR models cannot provide an accurate characterization of the actual VaR thresholds. In this section, we extend the unconditional VaR approach by taking into account the dynamic behavior of the conditional mean

[17] The conditonal VaR has so far not been used in the tests of risk-return tradeoff. However, the econometric approach to estimate conditional VaR is originally developed by Engle and Manganelli (2004) and Bali and Theodossiou (2004).

and volatility of financial returns. More specifically, we use the following ARMA(1,1)-GARCH(1,1) specification with the skewed t density:

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \alpha_2 \varepsilon_{t-1} + \varepsilon_t, \quad \varepsilon_t = z_t \sigma_{t|t-1} \tag{25}$$

$$E(\varepsilon_t^2 \mid \Omega_{t-1}) = \sigma_{t|t-1}^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{26}$$

where R_t is the return for period t, z_t is a random variable drawn from the skew t distribution, $\widehat{\mu}_{t|t-1} = \widehat{\alpha}_0 + \widehat{\alpha}_1 R_{t-1} + \widehat{\alpha}_2 \varepsilon_{t-1}$ is the conditional mean, and $\widehat{\sigma}_{t|t-1} = \sqrt{\widehat{\sigma}_{t|t-1}^2}$ is the conditional standard deviation of returns based on the information set up to time $t-1$. The conditional VaR is computed by substituting $\widehat{\mu}_{t|t-1}$, $\widehat{\sigma}_{t|t-1}$, and a into $VaR_{t|t-1} = \widehat{\mu}_{t|t-1} + a\widehat{\sigma}_{t|t-1}$, where a is obtained from the estimated skewed t distribution.

After obtaining the conditional VaR measure based on the above mentioned methodology, we use it as a right hand side variable in Panel A of Table VII. Observe that in the univariate regression it has a t statistic of 1.98, and when we control for the crash month, the t-statistic goes up to 3.51. Even after controlling for the macroeconomic variables, the t-statistic is 3.22. These significance levels are even higher than those we obtain when using the realized VaR as a proxy for the conditional downside risk. Hence, we conclude that there is a significantly positive relation between VaR measures that conditionally change over time and the expected market returns.

Our second measure of conditional variance is obtained from a simpler method. We consider the following regression:

$$VaR_{t+1} = \lambda + \delta VaR_t + \phi H_t + \zeta_t, \tag{27}$$

where H_t denotes a vector that consists of the lagged macroeconomic variables, i.e., RREL_t, TERM_t, DEF_t, and DP_t.[18] Therefore, assuming that investors' information set as of time t consists of the lagged VaR and a set of macroeconomic variables, we estimate the conditional VaR as the explained portion of the regression shown in equation (27), i.e., $E_t(VaR_{t+1}) = \widehat{\lambda} + \widehat{\delta} VaR_t + \widehat{\phi} H_t$. After running this first stage regression, we use the explained part in our second stage regressions designed to test risk-return tradeoff.

Panel B of Table VII presents results from the time-series regressions of the one-month-ahead value-weighted excess market return, R_{t+1}, on the conditional VaR, $E_t(VaR_{t+1})$, obtained from various horizons from 1 month to 6 months, dummy variable, lagged excess market return, and macroeconomic

[18] We test the statistical significance of the first-order serial correlation in alternative measures of VaR obtained from the past 1 month to 6 months of daily data. The monthly VaR measures are found to be highly persistent (as shown in Panel B Table I), which justify our use of the lagged VaR in equation (27).

variables. The results indicate that, similar to our findings in Panel A, there exists a significant relation between expected returns and conditional VaR at every horizon.

B Alternative Measures of Downside Risk

VaR provides information about the left tail of the empirical return distribution, however it is not the only measure of downside risk. If downside risk is an important determinant of expected returns, we expect other proxies of downside risk to perform well in predictive regressions too. In this section, we conduct an empirical analysis of the various left-tail risk measures.

An important example for a risk measure of this kind is Expected Shortfall originally proposed by Artzner et al. (1999). Expected Shortfall (ES) is defined as the conditional expectation of loss given that the loss is beyond the VaR level. That is, when the distributions of losses are continuous, expected shortfall at the 100(1–Φ) percent confidence level is defined by

$$ES_{\Phi}(R_t) = E\left[R_t | R_t \leq VaR_{\Phi}(R_t)\right]. \tag{28}$$

Equation (28) can be viewed as a mathematical transcription of the concept "average loss in the worst 100Φ% cases".

In addition to expected shortfall that measures the mean of losses larger than VaR, we also compute the variance of losses larger than VaR, and call it tail risk (TR)

$$TR_{\Phi}(R_t) = E\left[\left(R_t - E\left(R_t | R_t \leq VaR_{\Phi}(R_t)\right)\right)^2 | R_t \leq VaR_{\Phi}(R_t)\right]. \tag{29}$$

We consider the 2.5% and 5% tail risk ($TR_t^{2.5\%}$ and $TR_t^{5\%}$) and the 2.5% and 5% expected shortfall ($ES_t^{2.5\%}$ and $ES_t^{5\%}$) as alternative proxies for downside risk. In our empirical analysis we define the 2.5% (5%) tail risk as the sum of squared deviations of the lowest 2.5 percentile (5 percentile) of the NYSE/AMEX/Nasdaq index returns from the mean of index returns during the last 100 days. Similarly, we define the 2.5% (5%) expected shortfall as the average of the lowest 2.5 percentile (5 percentile) of the NYSE/AMEX/Nasdaq index returns observed during the last 100 days as of the end of month t.

Table VIII presents results from the regressions of the value-weighted index return on the lagged tail risk and expected shortfall measures. Panel A and B report results for $TR_t^{2.5\%}$ and $TR_t^{5\%}$. Both $TR_t^{2.5\%}$ and $TR_t^{5\%}$ have positive coefficient estimates and their t-statistics are around 2.3. In Panels C and D, we show results for expected shortfall measures, $ES_t^{2.5\%}$ and $ES_t^{5\%}$. The results are similar to those reported in Panels A and B. When we control for the lagged return, dummy variable and macroeconomic variables, $ES_t^{2.5\%}$ and $ES_t^{5\%}$ have positive and significant coefficient estimates. Overall, the parameter estimates in Table VIII indicate that alternative measures of left-tail risk measures predict the one-month-ahead market returns almost as well as VaR.

VI Conclusion

We examine the intertemporal relation between downside risk and expected stock returns. We use value at risk as a measure of downside risk and find a positive and significant relation between value at risk and expected return on the market. Moreover, we generate alternative measures of VaR based on the past 1 month to 6 months of daily data, and show that there is a significantly positive relation between VaR and expected market return for all horizons considered in the paper. Finally, we test the relative performance of various VaR and realized variance measures computed over different horizons in predictive regressions. The results indicate that VaR wins convincingly even when it is compared to the conditional variance proxies which have significant predictive power for market returns. These findings are robust across different measures of market return, loss probability levels, and after controlling for macroeconomic variables associated with business cycle fluctuations.

If downside risk is an important determinant of expected returns, we expect other proxies of downside risk to perform well in predictive regressions too. Therefore, we use expected shortfall and tail risk, both of which inform us about the left tail of the return distribution, as alternative measures of downside risk. We show that, regardless of the left-tail measure we use, our qualitative results from predictive regressions do not change.

The results provide strong evidence that there exists a positive and significant relation between downside risk and expected returns, implying that the more a market index can potentially fall in value the higher should be the expected return. Our findings also suggest that rare large moves in the market or relatively infrequent return observations can be interpreted as signal whereas the frequent small fluctuations can be viewed as noise which do not have power to explain time-series variation in excess market returns.

References

[1] Ang, Andrew, Joe Chen and Yuhang Xing, "Downside Risk," *Review of Financial Studies* (2005), forthcoming.

[2] Arditti Fred D., "Risk and the Required Return on Equity," *Journal of Finance* 22 (1967), 19-36.

[3] Arditti, Fred D., and Haim Levy, "Portfolio Efficiency Analysis in Three Moments: The Multi-Period Case," *Journal of Finance* 30 (1975), 797-809.

[4] Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath, "Coherent Measures of Risk," *Mathematical Finance* 9 (1999), 203-228.

[5] Arzac, Enrique R., and Vijas S. Bawa, "Portfolio Choice and Equilibrium in Capital Markets with Safety-First Investors," *Journal of Financial Economics* 4 (1977), 277-288.

[6] Bali, Turan G., "An Extreme Value Approach to Estimating Volatility and Value at Risk," *Journal of Business* 76 (2003), 83-108.

[7] Bali, Turan G., Nusret Cakici, Xeumin Yan, and Zhe Zhang, "Does Idiosyncratic Risk Really Matter?" *Journal of Finance* 60 (2005), 905-929.

[8] Bali, Turan G., and Panayiotis Theodossiou, , "A Conditional SGT VAR Approach with Alternative GARCH Models," *Annals of Operations Research* (2004), forthcoming.

[9] Baumol, William J., "An Expected Gain-Confidence Limit Criterion for Portfolio Selection," *Management Science* 10 (1963), 174-182.

[10] Berk, Jonathan B., "Necessary Conditions for the CAPM," *Journal of Economic Theory* 73 (1997), 245-257.

[11] Berkelaar, Arjan B., Roy Kouwenberg, and Thierry Post, "Optimal Portfolio Choice Under Loss Aversion," *Review of Economics and Statistics* 86 (2004), 973–987.

[12] Black, Fisher, and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81 (1973), 637-654.

[13] Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner, "ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics* 52 (1992), 5-59.

[14] Bollerslev, Tim, Robert F. Engle, and Daniel B. Nelson, "ARCH Models," in Robert Engle, and Daniel L. McFadden (Eds.), *Handbook of Econometrics*, Vol 4. (Amsterdam: Elsevier, 1994).

[15] Carr, Peter, Helyette Geman, Dilip B Madan, and Marc Yor, "The Fine Structure of Asset Returns: An Empirical Investigation," *Journal of Business* 75 (2002), 305-332.

[16] Campbell, John Y., "Stock Returns and the Term Structure," *Journal of Financial Economics* 18 (1987), 373-399.

[17] Chamberlain, Gary, "A Characterization of the Distributions that Imply Mean-Variance Utility Functions," *Journal of Economic Theory* 29 (1983), 185-201.

[18] Cox, John C., and Stephen A. Ross, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics* 3 (1976), 145-166.

[19] Deaton, Angus, "Saving and Liquidity Constraints," *Econometrica* 59 (1991), 1221-1248.

[20] Dittmar, Robert F., "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns," *Journal of Finance* 57 (2002), 369-403.

[21] Engle, Robert F., and Simone Manganelli, "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," *Journal of Business and Economic Statistics* 22 (2004), 367-381.

[22] French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, "Expected Stock Returns and Volatility," *Journal of Financial Economics* 19 (1987), 3-29.

[23] Ghysels, Eric, Pedro Santa-Clara and Rossen Valkanov, "There is a Risk-Return Tradeoff After All," *Journal of Financial Economics* (2004), forthcoming.

[24] Goyal, Amit, and Pedro Santa-Clara, "Idiosyncratic Risk Matters!" *Journal of Finance* 58 (2003), 975-1008.

[25] Hansen, Bruce E., "Autoregressive Conditional Density Estimation," *International Economic Review* 35 (1994), 705-730.

[26] Harrison, Paul, and Harold H. Zhang, "An Invetigation of the Risk and Return Relation at Long Horizons," *Review of Economics and Statistics* 81 (1999), 399-408.

[27] Harvey, Campbell R., and Akhtar Siddique, "Conditional Skewness in Asset Pricing Tests," *Journal of Finance* 55 (2000), 1263-1295.

[28] Kahneman, Daniel, Jack L. Knetsch, and Richard H., Thaler, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy* 99 (1990), 1325-1350.

[29] Kimball, Miles S., "Standard Risk Aversion," *Econometrica* 61 (1993), 589-611.

[30] Kraus, Alan, and Robert H. Litzenberger, "Skewness Preference and the Valuation of Risk Assets," *Journal of Finance* 31 (1976), 1085-1100.

[31] Levy, Haim, "A Utility Function Depending on the First Three Moments," *Journal of Finance* 24 (1969), 715-721.

[32] Levy, Haim, and Harry Markowitz, "Approximating Expected Utility by a Function of Mean and Variance," *American Economic Review* 69 (1979), 308-317.

[33] Levy, Haim, and Marshall Sarnat, "Safety First: An Expected Utility Principle," *Journal of Financial and Quantitative Analysis* 7 (1972), 1829-1834.

[34] Loistl, Otto, "The Erroneous Approximation of Expected Utility by Means of a Taylor's Series Expansion: Analytic and Computational Results," *American Economic Review* 66 (1976), 904-910.

[35] Longin, Francois M., "From Value at Risk to Stress Testing: The Extreme Value Approach," *Journal of Banking and Finance* 24 (2000), 1097-1130.

[36] Markowitz, Harry, "Portfolio Selection," *Journal of Finance* 7 (1952a), 77-91.

[37] Markowitz, Harry, "The Utility of Wealth," *Journal of Political Economy* 60 (1952b), 151-158.

[38] Markowitz Harry, *Portfolio Selection*: *Efficient Diversification of Investments*, 2nd Edition, (Malden: Blackwell, 1959)

[39] Merton, Robert C., "An Intertemporal Capital Asset Pricing Model," *Econometrica* 41 (1973), 867-887.

[40] Merton, Robert C., "Option Pricing When Underlying Stock Returns are Discontinuous," *Journal of Financial Economics* 3 (1976), 125-144.

[41] Merton, Robert C., "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics* 8 (1980), 323-361.

[42] Newey, Whitney K., and Kenneth D. West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55 (1987), 703-708.

[43] Post, Thierry, and Pim van Vliet, "Downside Risk and Asset Pricing," *Journal of Banking and Finance (*2004a), forthcoming.

[44] Post, Thierry, and Pim van Vliet, "Conditional Downside Risk and the CAPM," Working paper (2004b), Erasmus Research Institute of Management.

[45] Pratt, John W., and Richard J. Zeckhauser, "Proper Risk Aversion," *Econometrica* 55 (1987), 143-154.

[46] Roy, A. D., "Safety First and the Holding of Assets," *Econometrica* 20 (1952), 431-449.

[47] Scott, Robert C., and Philip A. Horvath, "On the Direction of Preference for Moments of Higher Order than the Variance," *Journal of Finance* 35 (1980), 915-919.

[48] Scruggs, John T., "Resolving the Puzzling Intertemporal Relation Between the Market Risk Premium and Conditional Market Variance: A Two-Factor Approach," *Journal of Finance* 52 (1998), 575-603.

Table I. Descriptive Statistics

Panel A shows summary statistics for the monthly return on the value-weighted NYSE/AMEX/Nasdaq, NYSE/AMEX, NYSE, Nasdaq, and S&P 500 index. Panel B shows summary statistics for Value at Risk (VaR_t) computed using rolling window estimation over various months (k). VaR_t is defined as -1 times the minimum NYSE/AMEX/Nasdaq index return observed during the last k months as of the end of each month t. Each month is assumed to have 21 trading days. We report the mean, median, standard deviation, maximum, minimum, skewness, kurtosis, and Jarque-Bera statistics. We report the first-order autoreggressive coefficients for VaR measures. Statistics in Panel B are computed after eliminating the month of October 1987.

Panel A. Monthly Index Returns

	NYSE/AMEX/Nasdaq	NYSE/AMEX	NYSE	Nasdaq	SP500
Obs	486	486	486	360	486
Mean	0.0094	0.0095	0.0095	0.0099	0.0096
Median	0.012	0.011	0.011	0.013	0.011
Std. Dev.	0.045	0.043	0.043	0.067	0.044
Maximum	0.166	0.165	0.168	0.220	0.170
Minimum	-0.225	-0.218	-0.216	-0.271	-0.216
Skewness	-0.451	-0.369	-0.345	-0.471	-0.317
Kurtosis	4.907	5.077	5.050	4.628	4.793
Jarque-Bera	90.092	98.394	94.709	53.092	73.239

Panel B. Value at Risk over Various Horizons

k	1	2	3	4	5	6
Obs	483	483	481	481	480	478
Mean	0.015	0.018	0.021	0.023	0.025	0.026
Median	0.013	0.015	0.017	0.019	0.020	0.022
Std. Dev.	0.009	0.012	0.015	0.017	0.018	0.020
Maximum	0.066	0.171	0.171	0.171	0.171	0.171
Minimum	0.002	0.004	0.005	0.005	0.008	0.008
Skewness	2.084	4.999	5.131	5.028	4.793	4.594
Kurtosis	10.220	52.412	47.158	41.801	36.314	32.412
AR(1)	0.36	0.48	0.70	0.77	0.82	0.86

Table II. Parametric and Nonparametric Value at Risk

This table presents results from the time-series regressions of the value-weighted excess market return on nonparametric and parametric value at risk (VaR and VaR_t^p). VaR_t is defined as the minimum index return observed during the last 100 days as of the end of month t. VaR_t^p is the parametric value at risk calculated based on the lower tail of the skewed t distribution. The original VaRs are multiplied by –1 before running our regressions. Value-weighted excess return, R_t, is defined as the return on the value-weighted NYSE/AMEX/Nasdaq index minus the one-month Treasury bill rate. A dummy takes the value of one in October 1987 and zero otherwise. DEF_t is the default spread calculated as the difference between the yields on BAA- and AAA-rated corporate bonds. TERM_t is the term spread calculated as the difference between the yields on the 10-year Treasury bond and the three-month Treasury bill. RREL_t is the stochastically detrended riskless rate defined as the three-month Treasury bill rate minus its 12-month backward moving average. DP_t is the dividend yield on S&P 500 index. In each regression, the dependent variable is the one-month-ahead excess market return, R_{t+1}. In each regression, the first row gives the estimated coefficients. The second row gives the Newey-West adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. Nonparametric 1% VaR

Constant	VaR_t	R_t	dummy	RREL_t	TERM_t	DEF_t	DP_t	R^2
0.000	0.177							0.58%
(-0.03)	(1.92)							
-0.002	0.274	0.029	-0.114					1.98%
(-0.76)	(2.91)	(0.67)	(-6.74)					
-0.015	0.296	0.028	-0.111	-0.246	-0.069	-0.002	0.387	3.28%
(-1.88)	(2.70)	(0.61)	(-6.14)	(-0.94)	(-0.31)	(0.00)	(1.37)	

Panel B. Parametric 1% VaR

Constant	VaR_t^p	R_t	dummy	RREL_t	TERM_t	DEF_t	DP_t	R^2
0.001	0.164							0.37%
(0.27)	(1.65)							
-0.002	0.286	0.031	-0.112					1.71%
(-0.59)	(2.79)	(0.71)	(-6.51)					
-0.014	0.310	0.029	-0.109	-0.261	-0.077	0.058	0.358	3.02%
(-1.81)	(2.67)	(0.62)	(-6.16)	(-0.99)	(-0.35)	(0.08)	(1.29)	

Table III. Maximum Likelihood Estimation with Skewed t Distribution

This table presents the maximum likelihood parameter estimates and the asymptotic t-statistics in parentheses based on the skewed t density. VaR$_t$, VaR$_t^p$, R_t, dummy and control variables are defined in Table II. In each regression, the dependent variable is the one-month-ahead excess market return, R_{t+1}. The maximum likelihood estimation of the risk-return tradeoff yields the intercept (α) and the slope coefficient (β) in the conditional mean equation as well as the standard deviation (σ), skewness (λ), and tail-thickness (v) parameters of the skewed t density.

Panel A. Nonparametric 1% VaR

$$R_{t+1} = \alpha + \beta VaR_t + \delta R_t + \varphi dummy + \Omega F_t + \varepsilon_{t+1};\ \varepsilon_{t+1} \sim skewed\ t(\sigma, \lambda, v)$$

$Constant$	VaR_t	R_t	$dummy$	$RREL_t$	$TERM_t$	DEF_t	DP_t	σ	λ	v
0.025	0.267							0.045	-0.204	6.751
(3.13)	(2.59)							(20.84)	(-3.35)	(3.28)
0.022	0.254	0.001	-0.247					0.044	-0.179	8.321
(2.58)	(2.49)	(0.03)	(-0.10)					(21.41)	(-2.73)	(2.37)
0.021	0.232	-0.027	-0.244	-0.297	-0.210	0.513	0.136	0.043	-0.216	8.287
(2.03)	(2.14)	(0.60)	(-0.11)	(-1.50)	(-1.14)	(0.86)	(0.66)	(20.25)	(-3.19)	(2.27)

Panel B. Parametric 1% VaR

$$R_{t+1} = \alpha + \beta VaR_t^p + \delta R_t + \varphi dummy + \Omega F_t + \varepsilon_{t+1};\ \varepsilon_{t+1} \sim skewed\ t(\sigma, \lambda, v)$$

$Constant$	VaR_t^p	R_t	$dummy$	$RREL_t$	$TERM_t$	DEF_t	DP_t	σ	λ	v
0.026	0.279							0.045	-0.203	6.794
(3.20)	(2.27)							(20.94)	(-3.33)	(3.27)
0.023	0.262	0.003	-0.247					0.044	-0.177	8.418
(2.76)	(2.17)	(0.08)	(-0.03)					(21.59)	(-2.69)	(2.36)
0.022	0.239	-0.027	-0.245	-0.309	-0.218	0.580	0.108	0.043	-0.216	8.393
(2.13)	(2.16)	(0.61)	(-0.09)	(-1.55)	(-1.18)	(0.97)	(0.53)	(20.44)	(-3.20)	(2.26)

Table IV. Various Measures of Realized Variance

This table shows estimates of the risk-return tradeoff with the rolling window estimators of realized variance. Each month is assumed to be 21 days. For each horizon the realized variance is computed as the sum of the squared daily returns on the value weighted NYSE/AMEX/Nasdaq index plus an adjustment term for the first-order serial correlation in daily returns. First column in each panel shows the number of months (k) used to compute the variance. R_t, dummy and control variables are defined in Table II. In each regression, the dependent variable is the one-month-ahead excess market return, R_{t+1}. Panel A presents the parameter estimates without the control variables and Panel B presents the parameter estimates with the control variables. In each regression, the first row gives the estimated coefficients. The second row gives the Newey-West adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. Mean-Variance Tradeoff

k	$Constant$	σ_t^2	R_t	$dummy$	R^2
1	0.004	0.381	0.032	-0.097	0.76%
	(1.38)	(0.39)	(0.73)	(-1.54)	
2	0.001	0.688	0.038	-0.116	1.28%
	(0.60)	(1.71)	(0.83)	(-4.39)	
3	0.000	0.643	0.033	-0.114	1.71%
	(0.19)	(2.67)	(0.75)	(-5.95)	
4	0.000	0.538	0.029	-0.108	1.82%
	(-0.01)	(2.78)	(0.65)	(-6.12)	
5	0.000	0.367	0.027	-0.097	1.45%
	(0.18)	(2.08)	(0.61)	(-5.77)	
6	0.000	0.294	0.027	-0.092	1.32%
	(0.20)	(1.81)	(0.62)	(-5.63)	

Panel B. Mean-Variance Tradeoff with Macrovariables

k	$Constant$	σ_t^2	R_t	$dummy$	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	-0.006	0.349	0.027	-0.091	-0.264	-0.081	0.272	0.253	1.93%
	(0.83)	(0.29)	(0.57)	(-1.19)	(-1.04)	(-0.38)	(0.36)	(0.91)	
2	-0.009	0.840	0.037	-0.121	-0.229	-0.039	0.054	0.304	2.62%
	(-1.29)	(2.18)	(0.75)	(-5.13)	(-0.90)	(-0.18)	(0.07)	(1.09)	
3	-0.009	0.683	0.031	-0.112	-0.225	-0.036	-0.048	0.326	2.85%
	(-1.34)	(2.85)	(0.66)	(-6.38)	(-0.87)	(-0.17)	(-0.06)	(1.16)	
4	-0.010	0.558	0.028	-0.104	-0.220	-0.037	0.105	0.343	2.89%
	(-1.40)	(2.82)	(0.60)	(-6.46)	(-0.85)	(-0.17)	(-0.14)	(1.21)	
5	-0.009	0.370	0.026	-0.092	-0.228	-0.042	0.007	0.312	2.54%
	(-1.31)	(1.92)	(0.55)	(-5.81)	(-0.90)	(-0.20)	(0.01)	(1.09)	
6	-0.009	0.285	0.026	-0.087	-0.231	-0.046	0.048	0.304	2.40%
	(-1.29)	(1.59)	(0.55)	(-5.59)	(-0.91)	(-0.22)	(0.06)	(1.05)	

Table V. Value at Risk at Various Horizons

This table shows estimates of the risk-return tradeoff with the rolling window estimators of Value-at-Risk. Each month is assumed to be 21 days. First column in each panel shows the number of months (k) used to compute the VaR. For each horizon the VaR is computed as the lowest daily returns on the value weighted NYSE/AMEX/Nasdaq index. R_t, dummy and control variables are defined in Table II. In each regression, the dependent variable is the one-month-ahead excess market return, R_{t+1}. Panel A presents the parameter estimates without the control variables and Panel B presents the parameter estimates with the control variables. In each regression, the first row gives the estimated coefficients. The second row gives the Newey-West adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. Mean-VaR Tradeoff

k	$Constant$	VaR_t	R_t	$dummy$	R^2
1	-0.002	0.456	0.068	-0.136	1.42%
	(-0.66)	(1.98)	(1.37)	(-4.17)	
2	-0.004	0.474	0.054	-0.140	2.35%
	(-1.50)	(3.99)	(1.20)	(-7.45)	
3	-0.003	0.380	0.038	-0.129	2.27%
	(-1.15)	(3.66)	(0.87)	(-7.26)	
4	-0.004	0.367	0.033	-0.127	2.55%
	(-1.32)	(4.12)	(0.77)	(-7.76)	
5	-0.002	0.267	0.029	-0.113	1.91%
	(-0.73)	(2.93)	(0.67)	(-6.82)	
6	-0.002	0.225	0.030	-0.106	1.71%
	(0.55)	(2.57)	(0.68)	(-6.53)	

Panel B. Mean-VaR Tradeoff with Macrovariables

k	$Constant$	VaR_t	R_t	$dummy$	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	-0.016	0.545	0.075	-0.143	-0.228	-0.024	-0.096	0.396	2.73%
	(-1.89)	(2.06)	(1.33)	(-4.10)	(-0.89)	(-0.11)	(-0.13)	(1.36)	
2	-0.018	0.531	0.058	-0.142	-0.231	-0.026	-0.146	0.430	3.76%
	(-2.37)	(4.04)	(1.17)	(-7.19)	(-0.89)	(-0.12)	(-0.21)	(1.55)	
3	-0.016	0.420	0.039	-0.128	-0.246	-0.050	-0.130	0.425	3.65%
	(-2.13)	(3.63)	(0.83)	(-6.88)	(-0.94)	(-0.23)	(-0.18)	(1.52)	
4	-0.017	0.401	0.035	-0.126	-0.235	-0.062	-0.149	0.440	3.91%
	(-2.17)	(3.74)	(0.75)	(-6.89)	(-0.89)	(-0.27)	(-0.21)	(1.57)	
5	-0.015	0.291	0.028	-0.110	-0.247	-0.068	-0.009	0.390	3.23%
	(-1.87)	(2.69)	(0.61)	(-6.22)	(-0.95)	(-0.31)	(-0.01)	(1.38)	
6	-0.014	0.247	0.029	-0.103	-0.249	-0.073	0.042	0.377	3.03%
	(-1.80)	(2.40)	(0.61)	(-6.06)	(-0.96)	(-0.34)	(0.06)	(1.32)	

Table VI. Comparing VaR with Realized Variance

This table compares the relative performance of rolling window estimates of the VaR and realized variance, and MIDAS estimator in predicting future market returns. Each month is assumed to be 21 days. First column in each panel shows the number of months (k) used to compute the VaR and Variance. For each horizon the VaR is computed as the lowest daily returns on the value weighted NYSE/AMEX/Nasdaq index. Similarly, for each horizon, the realized variance is computed as the sum of squared daily returns on the value weighted NYSE/AMEX/Nasdaq index plus an autocorrelation adjustment term. R_t, dummy, and macroeconomic variables are defined in Table II. VaR_t^{ew} is the equal-weighted VaR measure. In each regression, the dependent variable is the one-month-ahead excess market return, R_{t+1}. In each regression, the first row gives the estimated coefficients. The second row gives the Newey-West adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. Relative Performance of VaR and Variance

k	$Constant$	VaR_t	σ_t^2	R_t	$dummy$	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	-0.018	0.926	-2.162	0.088	-0.058	-0.258	-0.069	-0.076	0.432	3.16%
	(-2.14)	(2.69)	(-1.47)	(1.59)	(-0.81)	(-1.01)	(-0.33)	(-0.10)	(1.54)	
2	-0.020	0.766	-0.741	0.061	-0.129	-0.252	-0.046	-0.116	0.458	3.96%
	(-2.49)	(3.30)	(-1.23)	(1.23)	(-6.03)	(-0.98)	(-0.21)	(-0.16)	(1.65)	
3	-0.018	0.546	-0.329	0.040	-0.125	-0.262	-0.066	-0.094	0.441	3.71%
	(-2.10)	(2.05)	(-0.57)	(0.85)	(-6.84)	(-1.00)	(-0.30)	(-0.13)	(1.57)	
4	-0.018	0.571	-0.401	0.037	-0.124	-0.258	-0.090	-0.052	0.454	4.06%
	(-2.20)	(2.15)	(-0.72)	(0.79)	(-7.21)	(-0.98)	(-0.41)	(-0.07)	(1.64)	
5	-0.016	0.422	-0.297	0.029	-0.110	-0.271	-0.096	0.103	0.397	3.34%
	(-1.92)	(1.77)	(-0.62)	(0.63)	(-6.50)	(-1.05)	(-0.45)	(0.14)	(1.42)	
6	-0.015	0.371	-0.266	0.029	-0.104	-0.274	-0.105	0.187	0.377	3.15%
	(-1.86)	(1.68)	(-0.62)	(0.62)	(-6.34)	(-2.05)	(-0.49)	(0.24)	(1.33)	

Panel B. Relative Performance of VaR and MIDAS

k	$Constant$	VaR_t^{ew}	$MIDAS_t$	R_t	$dummy$	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	-0.012		2.812	0.035	-0.126	-0.218	-0.025	-0.040	0.371	2.95%
	(-1.70)		(2.68)	(0.72)	(-5.58)	(-0.84)	(-0.12)	(-0.05)	(1.29)	
2	-0.018	0.466		0.044	-0.134	-0.230	-0.041	-0.111	0.430	3.68%
	(-2.31)	(3.47)		(0.92)	(-6.65)	(-0.88)	(-0.18)	(-0.15)	(1.51)	
3	-0.020	0.726	-2.325	0.048	-0.123	-0.242	-0.065	0.003	0.413	3.87%
	(-2.34)	(2.03)	(-0.84)	(0.98)	(-5.39)	(-0.92)	(-0.29)	(0.01)	(1.48)	

Table VII. Conditional VaR

This table presents results from the time-series regressions of the value-weighted excess market return on the conditional value at risk $E_t(VaR_{t+1})$. In Panel A, conditional VaR is computed by using conditional mean and conditional standard deviation along with the skewness and tail-thickness parameters of the skewed t density. In Panel B, conditional VaR is defined as the portion of one-month ahead VaR explained by the lagged VaR and macroeconomic variables. R_t, dummy, and macroeconomic variables are defined in Table II. In each regression, the dependent variable is the one-month-ahead excess market return, R_{t+1}. In each regression, the first row gives the estimated coefficients. The second row gives the Newey-West adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. Conditional VaR from Skewed-t GARCH

$Constant$	$E_t(VaR_{t+1})$	R_t	$dummy$	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.001	0.205							0.65%
(-0.18)	(1.98)							
-0.003	0.332	0.032	-0.122					2.19%
(-1.13)	(3.51)	(0.74)	(-7.27)					
-0.016	0.368	0.033	-0.121	-0.245	-0.066	-0.112	0.432	3.56%
(-2.10)	(3.22)	(0.70)	(-6.55)	(-0.93)	(-0.29)	(0.16)	(1.53)	

Panel B. Conditional VaR from 2-Stage Regression

$Months$	$Constant$	$E_t(VaR_{t+1})$	R_t	$dummy$	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	-0.055	2.529	0.075	-0.143	-0.234	0.077	-1.616	1.095	2.73%
	(-2.26)	(2.06)	(1.33)	(-4.10)	(-0.92)	(0.35)	(-1.25)	(2.07)	
2	-0.028	0.927	0.058	-0.142	-0.264	-0.024	-0.503	0.625	3.76%
	(-3.23)	(4.04)	(1.17)	(-7.19)	(1.02)	(-0.11)	(-0.67)	(2.11)	
3	-0.021	0.576	0.039	-0.128	-0.274	-0.058	-0.295	0.518	3.65%
	(-2.56)	(3.63)	(0.83)	(-6.88)	(-1.04)	(-0.26)	(-0.41)	(1.80)	
4	-0.020	0.509	0.035	-0.126	-0.261	-0.065	-0.266	0.509	3.91%
	(-2.50)	(3.74)	(0.75)	(-6.89)	(-0.99)	(-0.29)	(-0.37)	(1.78)	
5	-0.017	0.352	0.028	-0.110	-0.254	-0.075	-0.067	0.430	3.23%
	(-2.05)	(2.69)	(0.61)	(-6.22)	(-0.97)	(-0.34)	(-0.09)	(1.49)	
6	-0.016	0.289	0.029	-0.103	-0.253	-0.077	-0.004	0.408	3.03%
	(-1.94)	(2.40)	(0.61)	(-6.06)	(-0.97)	(-0.35)	(0.01)	(1.4)	

Table VIII. Alternative Measures of Downside Risk

This table presents results from the time-series regressions of the value-weighted excess market return on the 2.5% and 5% Tail Risk ($TR_t^{2.5\%}$ and $TR_t^{5\%}$) and the 2.5% and 5% Expected Shortfall ($ES_t^{2.5\%}$ and $ES_t^{5\%}$) measures. 2.5% (5%) Tail Risk is computed as the sum of squared deviations of the lowest 2.5 percentile (5 percentile) of the NYSE/AMEX/Nasdaq index returns from the mean of index returns during the latest 100 days. 2.5% (5%) Expected Shortfall is computed as the average of the lowest 2.5 percentile (5 percentile) of NYSE/AMEX/Nasdaq index returns observed during the last 100 days as of the end of month t. R_t, dummy, and macroeconomic variables are defined in Table II. The second row gives the Newey-West adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. 2.5% Tail Risk

Constant	$TR_t^{2.5\%}$	R_t	dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.008	1.163	0.024	-0.110	-0.278	-0.119	0.194	0.301	2.77%
(-1.09)	(2.30)	(0.52)	(-5.61)	(-1.07)	(-0.53)	(0.28)	(1.11)	

Panel B. 5% Tail Risk

Constant	$TR_t^{5\%}$	R_t	dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.008	1.053	0.024	-0.107	-0.276	-0.111	0.163	0.311	2.75%
(-1.14)	(2.29)	(0.52)	(-5.71)	(-1.06)	(-0.50)	(0.23)	(1.14)	

Panel C. 2.5% Expected Shortfall

Constant	$ES_t^{2.5\%}$	R_t	dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.016	0.453	0.032	-0.104	-0.249	-0.050	-0.087	0.397	3.03%
(-2.04)	(2.56)	(0.68)	(-6.23)	(-0.95)	(-0.23)	(-0.12)	(1.38)	

Panel D. 5% Expected Shortfall

Constant	$ES_t^{5\%}$	R_t	dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.016	0.524	0.032	-0.096	-0.253	-0.043	-0.102	0.393	2.83%
(-2.05)	(2.36)	(0.69)	(-6.30)	(-0.97)	(-0.20)	(0.13)	(1.35)	

APPENDIX

We provide evidence that VaR is an important measure of market risk in predicting returns on the value-weighted NYSE/AMEX/Nasdaq index. However, the positive relation between VaR and expected returns on the NYSE/AMEX/Nasdaq index may well be due to the Nasdaq stocks which tend to be smaller stocks with return distributions that significantly deviate from a normal distribution. Therefore, here we consider other stock market indices. Panels A to D report results for the value-weighted NYSE/AMEX, NYSE, Nasdaq, and S&P 500 index portfolios, respectively. For all market indices, there is a positive and highly significant relation between VaR and excess market return; the t-statistics vary from 2.33 for S&P 500 to 2.64 for Nasdaq. Moreover, the R^2 values vary between 3.11% for NYSE and 4.20% for Nasdaq. The results indicate that the strong positive relation between VaR and expected return is robust across different stock market indices.

Panel A. NYSE/AMEX

Constant	VaR_t	R_t	dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.011	0.263	0.009	-0.117	-0.230	-0.103	0.175	0.284	3.12%
(-1.69)	(2.59)	(0.18)	(-6.47)	(-0.89)	(-0.49)	(0.25)	(1.15)	

Panel B. NYSE

Constant	VaR_t	R_t	dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.011	0.256	0.002	-0.118	-0.228	-0.095	0.195	0.273	3.11%
(-1.66)	(2.53)	(0.04)	(-6.49)	(-0.90)	(-0.46)	(0.29)	(1.11)	

Panel C. Nasdaq

Constant	VaR_t	R_t	dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.039	0.548	0.112	-0.071	-0.193	0.158	-0.447	0.841	4.20%
(-2.23)	(2.64)	(2.15)	(-3.06)	(-0.52)	(0.46)	(-0.40)	(1.74)	

Panel D. S&P 500

Constant	VaR_t	R_t	dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
-0.011	0.248	-0.025	-0.113	-0.226	-0.070	0.147	0.276	3.15%
(-1.47)	(2.33)	(-0.52)	(-6.67)	(-0.88)	(-0.32)	(0.21)	(1.00)	