Variance Trading and Market Price of Variance Risk*

Oleg Bondarenko†

December 2007 (first version: June 2003)

Abstract

This paper develops a new approach for the variance trading. The approach departs from the existing literature by focusing on the empirically relevant realized variance, as opposed to the unobservable integrated variance. We show that the discretely-sampled realized variance can be robustly replicated under very general conditions. The replication strategy specifies the optimal timing for rebalancing in the underlying. The deviations from the optimal schedule can lead to surprisingly large hedging errors. The results have important implications for hedging actual instruments, such as OTC variance swaps and CBOE variance futures.

In the empirical application, we synthesize the prices of the variance contract on S&P 500 index over the 17-year period from 01/1990 to 12/2006. We find that the market variance risk is priced, its risk premium is negative and economically very large. The variance risk premium cannot be explained by the known risk factors and option returns. The variance contract appears to be even more "expensive" than the already puzzling index puts. Several examples demonstrate the important role of the variance risk in the economy.

JEL Classification: G12, G13, G23

Keywords: Variance Risk, Option Valuation, Risk-Neutral Density, Stochastic Volatility

*Some results of this paper first appeared under title "Market Price of Variance Risk and Performance of Hedge Funds" (2003). I would like to thank for helpful comments Torben Andersen, Peter Carr, Gib Bassett, Bill Fung, Mila Getmansky, Will Goetzmann, Campbell Harvey, David Hsieh, Jens Jackwerth (ESSFM 2003 discussant), Andrew Lo (AFA 2006 discussant), Eugene Kandel, Shmuel Kandel (Gutmann Center Symposium 2004 discussant), Per Mykland, Lubos Pastor, and seminar participants at Northwestern University, Stockholm School of Economics, University of Chicago, University of Florida, University of Illinois at Chicago, Washington University in St. Louis, Morningstar, CEPR European Summer Symposium in Financial Markets 2003, Gutmann Center Symposium on Hedge funds 2004, IXIS/NYU Conference on Hedge Funds 2005, American Finance Association meeting 2006, and Chicago Quantitative Alliance meeting 2006. I am grateful to Torben Andersen and David Hsieh for the help with the data. Any remaining errors are mine.

†Department of Finance (MC 168), University of Illinois at Chicago, 601 S. Morgan St., Chicago, IL 60607. Phone: (312) 996-2362. Email: olegb@uic.edu

1 Introduction

It is well-known that volatility of equity returns tends to change unpredictably over time. In particular, many empirical studies have documented that broad market indexes often experience large shifts in volatility. However, it is less-understood whether investors require compensation for the volatility risk and, if yes, to what extent. More generally, what role does the volatility risk play in the economy? The answers to these questions are important for a number of applications, including option pricing, optimal portfolio construction, risk management, performance evaluation, and others.

Recently, there has been a growing demand for new instruments to hedge shifts in volatility. Several innovative derivative instruments have been introduced in over-the-counter market. One important example is variance swaps on various assets. A variance swap has a payoff, which is equal to the difference between the realized variance over a given period and the contract variance. Thus, a variance swap allows traders to bet directly on the level of future variance.[1] Trading in variance swaps has grown dramatically in the aftermath of the recent financial turmoils, such as the summer of 1997, the fall of 1998, and the fall of 2001. In May 2004, CBOE launched a closely related instrument, the futures on the S&P 500 realized variance.

In principle, giving a history of prices for the variance swaps or the CBOE variance futures, it would be relatively straightforward to study the variance risk. Unfortunately, this approach is not practical because both markets are still rather illiquid and have relatively short histories. Moreover, transaction prices for OTC variance swaps are not readily available to researches. In this paper, we implement a different approach, where market prices of the variance contract are recovered from observed prices of exchange-listed options. The approach is model-free, because no parametric structure is imposed on the underlying price process.

To explain our approach, suppose first that an underlying asset follows a continuous price process F_t. Let $IV_T = IV(0,T)$ denote the continuously-sampled integrated variance from time 0 to time T. Consider the variance contract, which at maturity T has a payoff of IV_T. The valuation approach for this contract has been developed by path-breaking work of Dupire (1993), Neuberger (1994), Carr and Madan (1998), Demeterfi, Derman, Kamal, and Zou (1999), and Britten-Jones and Neuberger (2000). The key insight in these papers is that the payoff IV_T can be robustly replicated by a static portfolio of standard options and dynamic trading in the underlying asset.

The existing approach is very elegant, but it has several important limitations. First, it requires the price process F_t to be continuous. This can be restrictive in view of ample empirical evidence of jumps in the S&P 500, as well as prices of other assets. Second, the approach requires continuous trading in the underlying asset, which is impossible in practice. Finally and most importantly, the existing approach focuses on valuing the *integrated* variance IV_T. The integrated variance is a theoretical concept, which makes the analysis tractable, but it is unobservable in practice. The actually traded instruments, such as OTC variance swaps or the CBOE futures, are all based on the discretely-sampled *realized* variance, typically computed from daily returns.

Specifically, let $RV_T = RV(\mathcal{T})$ denote the realized variance computed using the sampling

[1] Both *variance* and *volatility* swaps are available over-the-counter, with volatility being defined as the square root of variance. Although the two types of swaps are closely related, variance swaps appear to be more popular. This is mainly due to the fact that variance swaps offer useful theoretical advantages and as such are easier to hedge.

partition $\mathcal{T} = \{t_0, t_1, \ldots, t_n\}$ with equidistant trading dates $0 = t_0 < t_1 < \ldots < t_n = T$, and let $\Delta t = t_i - t_{i-1}$. When the sampling interval Δt tends to zero, the realized variance RV_T converges to the integrated variance IV_T, however, the difference between the two can be very large in typical situations. For example, consider a trader who needs to replicate the payoff RV_T with $T = 1$ month and $\Delta t = 1$ day. If the trader follows the existing replication strategy for the payoff IV_T, then the average hedging error is close to 40%. In other words, the existing approach exposes the trader to huge risks and it would be unacceptable in practice.

The first objective of our paper is to develop a new approach for valuing the payoffs based on the *empirically relevant* realized variance, as opposed to the *theoretical* integrated variance. Our main theoretical result demonstrates that the perfect replication of the discretely-sampled realized variance RV_T is possible under very general conditions.[2] The market does not have to be complete. The underlying price process can be completely general and, in particular, can have jumps. The partition $\mathcal{T}$ can be arbitrary: Δt does not have to be small and the partition does not have to be equidistant. Our approach nests the existing approach as a special case when (1) $\Delta t \to 0$ and (2) there are no jumps.

The replication strategy for the realized variance RV_T consists of two components: (i) a static portfolio of standard puts and calls with the same time to maturity T and different strikes, and (ii) a dynamic trading strategy in the underlying asset on dates $t_0 < t_1 < \ldots < t_n$. (The existing approach consists of the same components (i) and (ii), except the underlying must be rebalanced continuously.)

Suppose that the realized variance is computed from daily closing prices, as typically the case with actually traded instruments. To replicate the variance contract, the trader must at time 0 form a specific option portfolio and then trade in the underlying asset once a day, at the market close. That is, the replication strategy specifies the *exact* timing for rebalancing. Moreover, we demonstrate that if the trader deviates from the optimal timing, the hedging errors could be surprisingly high.

For example, if the trader rebalances too frequently (twice a day) or not frequently enough (every two days), then the average error of the replication strategy is around 26-29%. Rebalancing even more (less) frequently would expose the trader to even larger hedging errors. When the trader rebalances once a day, but not necessarily at the market close, then the hedging errors could again be surprisingly large. For example, if the trader rebalances only 5 minutes before the market close, the average hedging error exceeds 7%. This finding has important implications. As trading in the variance products becomes sufficiently active, there might be a predictable price impact at the market close due to the fact that traders, trying to minimize their hedging risks, will execute their orders at essentially the same time. For related reasons, one might also be concerned about possible price manipulation around the market close.

The second objective of our paper is to document the economic properties of the variance risk. Our theoretical result implies that, even when the variance contract is not traded, its unique arbitrage-free price can be inferred from prices of standard put and call options in a model-free fashion. In the empirical part of the paper, we use CME options on the S&P 500 index futures to synthesize prices of the variance contract and generate a series of monthly variance return r_v. Our sample is from 01/1990 to 12/2006 and allows us to study histori-

[2]In this paper, we study three variations of the realized variance. The three versions coincide in the continuous-time limit with no jumps. Although only one version allows for robust replication, we show that the differences between the three versions are negligible for practical purposes. In contrast, the difference between the integrated variance and any versions of the realized variance is typically very large.

cal performance of the variance contract over a very long period. The main results can be summarized as follows:

- We find that the variance risk is priced, its risk premium is negative and very large in economic terms. In particular, the unconditional variance risk premium is estimated as -31.24% per month, which is highly statistically significant. Over the same period, the risk premium on the market return r_m is 0.55% per month.

  Intuitively, because the variance contract pays off when the realized variance is high, investors are willing to pay high prices for the hedge against sharp increases in the market variance.

- The distribution of the variance return is highly non-normal. In particular, the variance return exhibits substantial positive skewness.

- Consistent with the previous literature, we find that the variance return r_v and the market return r_m are negatively correlated. This raises an important question as to whether the variance risk is priced *beyond* its negative correlation with the market return. We show that the negative correlation accounts for only a small portion of the variance risk premium (-2.72% per month). The remaining portion (-28.52% per month) can be interpreted as the compensation for the "pure" variance risk.

- Similarly, the strong negative variance risk premium cannot be explained by other known risk factors. In particular, when the variance return is regressed onto the market return, the size factor and the value factor of Fama and French (1993), the momentum factor of Jegadeesh and Titman (1993), as well as returns on ATM and OTM puts and calls, the intercept remains highly negative and statistically significant.

- We argue that selling the variance contract would have resulted in very high profits over the studied 17-year period. In particular, the Sharpe ratio for selling the variance contract is about 5 times as large as the Sharpe ratio for the market itself and almost 3 times as large as the Sharpe ratio for selling at-the-money put option.

- We find that the economic value to investors of being able to trade the variance contract could be very high. For example, consider an investor with the CRRA preferences who has $1 mln of the investable wealth. If the investor's risk aversion coefficient is 3 (10), then he is willing to pay up to $52,000 ($18,000) per month for being able to sell the variance contract. The economic value of introducing the variance contract is large even when the investor can also trade options on the market.

Overall, we conclude that the variance represents a new risk factor, which the market prices very heavily. This has important implications for explaining returns on various assets. If assets prices are sensitive to changes in the market variance, then assets expected returns must reflect the exposure to the variance risk. Even a slight exposure to the variance risk could have a non-trivial impact, because the variance risk premium is so substantial. The variance risk might be important to various institutions. In particular, most hedge fund categories exhibit negative exposure to the variance return. This explains the fact that hedge funds generally do better in quite periods than in volatile ones, even after accounting for the exposure to the market return. There exists some evidence that the variance return is correlated with proxies for the credit, liquidity, and correlation risks. Intuitively, when the variance suddenly rises, credit spreads often become wider, liquidity dries out, and correlations between different assets increase.

The remainder of the paper is organized as follows. This section concludes with the review of the related literature. Section 2 presents the new methodology for valuing the variance contract. Section describes the dataset of S&P 500 futures options. Section 4 documents the statistical properties of the variance return. Section 5 and Section 6 provide the discussion of the empirical results and various applications. Section 7 concludes.

1.1 Related Literature

The original approach for replicating the variance payoff has been developed in Dupire (1993), Neuberger (1994), Carr and Madan (1998), Demeterfi, Derman, Kamal, and Zou (1999), Britten-Jones and Neuberger (2000). More recently, Carr and Lee (2005) develop valuation of arbitrary functions of the variance, although their approach requires continuous trading in options. Jiang and Tian (2005) and Carr and Wu (2007) investigate the effect of jumps on variance trading. All these papers focus on the integrated variance and assume that continuous monitoring and rebalancing are possible. Branger and Schlag (2005) examine various economic reasons for variance trading.

The issue whether the variance risk is priced has been investigated in a number papers. Bakshi, Cao, and Chen (2000) and Buraschi and Jackwerth (2001) present evidence that equity index options are non-redundant securities, suggesting that other risks (in addition to the underlying's price) are factored in option prices. One interpretation for their results is that the variance risk may be priced. These approaches do not allow one to determine the sign or magnitude of the variance risk premium.

Coval and Shumway (2001) study daily returns on the ATM zero-beta straddles and argue that the variance risk premium is negative. In their approach, some (weak) parametric assumptions on the option pricing model are needed (to compute option betas), and the sensitivity of the ATM zero-beta straddle to the variance risk is model-specific. Furthermore, the approach requires frequent trading in options, which would be prohibitively costly in practice.

Bakshi and Kapadia (2003a, 2003b) investigate statistical properties of delta-hedged gains of equity call options, where a long option position is dynamically hedged with the underlying. In the stochastic volatility framework, they show that the sign of delta-hedged option portfolios corresponds to the sign of the variance risk premium. Empirically, they find that the variance risk is priced and is negative. In their approach, some (weak) parametric assumptions are required (to compute delta-hedges), and the magnitude of the variance risk premium cannot be determined.

There is also an extensive strand of the literature, in which structural models, such as Hull and White (1987), Heston (1993), Bates (2000), are calibrated to option data. Recent examples include Bakshi, Cao, and Chen (1997), Guo (1998), Bates (2000), Chernov and Ghysels (2000), Benzoni (2002), and Pan (2002), Broadie, Chernov, and Johannes (2005), among others. These approaches make the strongest assumptions and are subject to model misspecification.

More recently, Ang, Hodrick, Xing, and Zhang (2006) investigate performance of stock portfolios ranked by their sensitivity to changes in the VIX index. They find evidence of negative volatility risk premium in the cross-section of stock returns. Bollerslev, Gibson, and Zhou (2007), Wu (2005), and Todorov (2007) use high-frequency returns and the CBOE VIX index to study the dynamics of the variance risk premium. Egloff, Leippold, and Wu (2007) estimate a two-factor affine model to explain the term structure variation of OTC variance swap rates.

2 Theoretical Methodology

Throughout this section we fix the current time at $t = 0$ and consider only contracts which pay out at a fixed future date T. Let F_t denote the value of the S&P 500 futures contract expiring at date $T' \geq T$, and let $P_t(K) = P(K, T; F_t, t)$ and $C_t(K) = C(K, T; F_t, t)$ be the prices of European put and call with strike K and maturity T. To simplify exposition, we assume that the risk-free rate is zero.[3]

The option prices can be computed using the risk-neutral density (RND):

$$C_t(K) = \int_0^\infty (F_T - K)^+ \, h_t(F_T) \, dF_T, \qquad P_t(K) = \int_0^\infty (K - F_T)^+ \, h_t(F_T) \, dF_T,$$

where $h_t(F_T) = h(F_T, T; F_t, t)$ is RND. RND satisfies the relationship first discovered in Ross (1976), Breeden and Litzenberger (1978), Banz and Miller (1978):

$$h_t(F_T) = \left. \frac{\partial^2 C_t(K)}{\partial K^2} \right|_{K=F_T} = \left. \frac{\partial^2 P_t(K)}{\partial K^2} \right|_{K=F_T}. \tag{1}$$

For future needs, it is convenient to denote as $M_t(K)$ the minimum of the put and call:

$$M_t(K) = \min(P_t(K), C_t(K)).$$

In other words, of the two plain vanilla options with strike K, $M_t(K)$ equals the price of the one that is currently out-of-the-money.

2.1 The Integrated Variance Contract

To present the main idea of our approach, we initially assume that F_t follows a continuous process. Continuity of F_t and the absence of arbitrage opportunities imply that there is a risk-neutral measure under which F_t evolves as

$$\frac{dF_t}{F_t} = \sqrt{v_t} dB_t^*, \tag{2}$$

where B_t^* is a standard Brownian motion and v_t is a strictly positive adapted process. The instantaneous variance v_t can be a very general stochastic process. In particular, it can have jumps. It is only assumed that the following integral exists:

$$IV(t_1, t_2) := \int_{t_1}^{t_2} \left(\frac{dF_t}{F_t} \right)^2 = \int_{t_1}^{t_2} v_t dt,$$

where $IV(t_1, t_2)$ denotes the integrated variance over the period $[t_1, t_2]$.

[3] In reality, the risk-free rate is nonzero. However, in empirical tests, we convert *spot* prices of options into *forward* prices (for delivery at time T). To obtain forward prices, spot prices are multiplied by $e^{r_f(T-t)}$, where r_f is the risk-free rate over $[t, T]$. For example, the forward put price is $P_t(K) = e^{r_f(T-t)} P_t^s(K)$, where $P_t^s(K)$ is the spot put price. A similar approach has been used in, for example, Dumas, Fleming, and Whaley (1998).

Suppose that the holding period $[0, T]$ is fixed and consider the *variance contract*, which at maturity T has the payoff of $IV_T = IV(0, T)$. We are interested in the time-0 value of this derivative security, $E_0^*[IV_T]$, where $E^*[\cdot]$ denotes the expectation under the risk-neutral measure. (Although the extension to arbitrary t is straightforward, we will not need it.) A comment on the terminology might be in order. In practice, the variance is most commonly traded using variance swaps. The variance swap has a value of zero at inception and at maturity it pays the difference between IV_T and the fixed leg $E_0^*[IV_T]$. The variance contract can be obtained by entering the variance swap and investing the fixed leg of the swap at the risk-free rate. Focusing on the variance contract, as opposed to the variance swap, allows for an easier comparison with other securities, such as put and call options, and for a natural definition of the return on the variance.

The valuation approach for the variance contract has been developed in path-breaking work of Dupire (1993) and Neuberger (1994).[4] The key role in their approach is played by the following portfolio of put and call options:

$$OP_T \quad := \quad 2\left(\int_0^{F_0} \frac{P_T(K)}{K^2}dK + \int_{F_0}^{\infty} \frac{C_T(K)}{K^2}dK\right) = 2\int_0^{F_0} \frac{M_T(K)}{K^2}dK. \tag{3}$$

In other words, OP_T represents the payoff of the portfolio of European puts and calls with same maturity T and a continuum of strikes K from 0 to ∞, where portfolio weights are equal to $2/K^2$. The portfolio consists of puts for low strikes ($K \leq F_0$) and calls for high strikes ($K \geq F_0$), which reflects the fact that in practice out-of-the-money (OTM) options are more liquid than in-the-money (ITM) options.

The time-0 value of this portfolio is denoted as

$$U_0 := E_0^*\left[OP_T\right] = 2\left(\int_0^{F_0} \frac{P_0(K)}{K^2}dK + \int_{F_0}^{\infty} \frac{C_0(K)}{K^2}dK\right) = 2\int_0^{F_0} \frac{M_0(K)}{K^2}dK. \tag{4}$$

Proposition 1 establishes the replication strategy for the variance contract and shows that the value of the variance contract is equal to U_0. For completeness, its proof is provided in Appendix.

Proposition 1 *Assume that the process F_t is continuous. Then the payoff IV_T can be perfectly replicated by*

(i) a time-T payoff equal to OP_T;

(ii) a dynamic trading strategy, which at any time t holds $2\left(\frac{1}{F_t} - \frac{1}{F_0}\right)$ shares of the underlying.

Therefore,

$$E_0^*\left[IV_T\right] = U_0.$$

Proof: See Appendix A.

[4]See also Carr and Madan (1998), Demeterfi, Derman, Kamal, and Zou (1999), and Britten-Jones and Neuberger (2000).

Proposition 1 states that the two payoffs, IV_T and OP_T, have the same market values at time 0. This does *not* mean that the two payoffs are the same at maturity T. In fact, the two payoffs are quite different, because IV_T depends on the whole path of F_t over the period $[0,T]$, while OP_T only depends on the final value of the underlying at maturity T. However, the difference between the path-*dependent* payoff IV_T and path-*independent* payoff OP_T can be perfectly replicated by continuous trading in the underlying, with the risk-neutral expectation of the dynamic trading strategy (ii) being zero.

In view of Proposition 1, one can estimate the market price of the variance contract U_0 in the following way. Suppose one observes market prices of standard calls and puts with various strikes and the same maturity T. One can interpolate option prices with available strikes to obtain the call and put pricing functions $C_0(K)$ and $P_0(K)$ for a continuum of strikes. The two integrals in (4) are then evaluated by integrating $C_0(K)$ and $P_0(K)$ over K with the weight function $1/K^2$.

2.2 The Realized Variance Contract

The approach described in the previous subsection is very elegant, but it has three important limitations. First, although the approach imposes little structure on the variance process v_t, it does require the price process F_t to be continuous. This assumption can be restrictive: by now the empirical literature has documented strong evidence of jumps in the S&P 500 index, as well as prices of other financial assets.

Second, even under the restrictive assumption that the price path is continuous, computation of the integrated variance IV_T would be problematic because it requires continuous path monitoring. In practice, of course, the variance must be computed from discretely-sampled data. Importantly, the discretely-sampled variance can differs considerably from its continuously-sampled counterpart.

Finally, the approach assumes that the dynamic replication of the payoff IV_T in (26) is done *continuously*, which is, again, impossible in practice. Even though the underlying futures can be traded with low transaction costs, the overall costs of continuous rebalancing would be infinite. On the other hand, if the replicating portfolio is rebalanced only at a finite number of dates, then perfect replication of the payoff IV_T cannot be achieved.

We now propose a novel way around these limitations. From now on, we no longer require that the process F_t is continuous. Let $\mathcal{T} = \{t_0, t_1, \ldots, t_n\}$ denote a sampling partition with trading dates $0 = t_0 < t_1 < \ldots < t_n = T$ and let

$$\Delta t := \max_i (t_i - t_{i-1}).$$

Often, the trading dates are set at equal intervals, in which case $t_i - t_{i-1} = \Delta t = T/n$ for all i, but, in general, this does not have to be the case.

Let $x_i = F_{t_i}/F_{t_{i-1}} - 1$ denote the return over $[t_{i-1}, t_i]$. The realized variance at the partition $\mathcal{T}$ can be defined in several ways, including

$$RV_T^{(1)} = RV^{(1)}(\mathcal{T}) := \sum_{i=1}^{n} \left[\ln(1+x_i)\right]^2, \quad (5)$$

$$RV_T^{(2)} = RV^{(2)}(\mathcal{T}) := \sum_{i=1}^{n} x_i^2. \quad (6)$$

Over-the-counter variance swaps all use the daily sampling. Most of them are based on definition (5), although some are based on definition (6). The recently introduced CBOE futures on the realized variance also use daily sampling and are based on definition (5).[5]

The realized variance may be viewed as an empirical estimate of the integrated variance. It is constructed from discrete returns x_i and is specific to the partition. We also introduce another specification for the realized variance:

$$RV_T^{(3)} = RV^{(3)}(\mathcal{T}) := 2\sum_{i=1}^{n}\left(x_i - \ln(1+x_i)\right). \tag{7}$$

Although the specification in (7) may not "look" like variance, it has all the intuitive properties of variance. In particular,

- $RV_T^{(3)}$ is always nonnegative;
- In the special case when the price process F_t is continuous, all three measures $RV_T^{(1)}$, $RV_T^{(2)}$, and $RV_T^{(3)}$ converge to IV_T as Δt approaches zero;
- In typical applications, the three measures $RV_T^{(1)}$, $RV_T^{(2)}$, and $RV_T^{(3)}$ are very close to each other.

Figure 1 compares the functions

$$\begin{aligned} f^{(1)}(x) &= [\ln(1+x)]^2, \\ f^{(2)}(x) &= x^2, \\ f^{(3)}(x) &= 2\left(x - \ln(1+x)\right), \end{aligned}$$

used in the definitions of $RV_T^{(1)}$, $RV_T^{(2)}$, and $RV_T^{(3)}$. The figure illustrates several points, which are straightforward to verify analytically. First, the function $f^{(3)}(x)$ is positive for all returns $|x| > 0$. Second, the three functions are very close to each other for the range of typical daily returns. In particular, the differences between them are virtually zero when $|x| \leq 5\%$ and are still very small when $|x| \leq 10\%$.[6] Third, the function $f^{(3)}(x)$ always lies between $f^{(1)}(x)$ and $f^{(2)}(x)$. Finally, the function $f^{(3)}(x)$ is closer to $f^{(1)}(x)$ than to $f^{(2)}(x)$.[7]

What is so special about the new definition in (7)? The answer is provided by the following proposition.

Proposition 2 *For any partition $\mathcal{T}$, the payoff $RV_T^{(3)}$ can be perfectly replicated by*

(i) a time-T payoff equal to OP_T;

(ii) a dynamic trading strategy, which is rebalanced on dates $t_i \in \mathcal{T}$ to maintain $2\left(\frac{1}{F_{t_i}} - \frac{1}{F_0}\right)$ shares of the underlying.

[5]In principle, the realized variance can also be adjusted for the drift in the underlying. In this case, one subtracts from the expressions in (5)-(6) the adjustment term $\frac{1}{n}\left(\ln\frac{F_T}{F_0}\right)^2$. However, for typical variance swaps, such an adjustment is very small and is rarely used in practice.

[6]It is easy to check that the three functions only differ in terms of order $O(x^3)$.

[7]Specifically, the linear combination $\frac{2}{3}f^{(1)}(x) + \frac{1}{3}f^{(2)}(x)$ approximates the function $f^{(3)}(x)$ up to terms of $O(x^4)$.

Therefore,

$$E_0^*\left[RV_T^{(3)}\right] = U_0.$$

Proof: Definition (7) can be re-arranged as

$$\begin{aligned} RV_T^{(3)} &= -2\ln\frac{F_T}{F_0} + 2\sum_{i=1}^{n}\left(\frac{F_{t_i}-F_{t_{i-1}}}{F_{t_{i-1}}}\right) \\ &= 2\left(\frac{F_T-F_0}{F_0} - \ln\frac{F_T}{F_0}\right) + 2\sum_{i=1}^{n}\left(\frac{1}{F_{t_i}} - \frac{1}{F_0}\right)(F_{t_i}-F_{t_{i-1}}) \\ &= OP_T + 2\sum_{i=1}^{n}\left(\frac{1}{F_{t_i}} - \frac{1}{F_0}\right)(F_{t_i}-F_{t_{i-1}}), \end{aligned} \tag{8}$$

where the last equality follows from the proof of Proposition 1. Equation (8) identifies the two components of the replication strategy for the payoff $RV_T^{(3)}$: (i) the path-independent payoff OP_T, and (ii) discrete rebalancing in the underlying performed on dates $t_0, t_1, \ldots, t_n$.

Since F_t is a martingale under the risk-neutral measure, equation (8) implies that the market price of the realized variance $RV_T^{(3)}$ is equal to U_0.

□

Proposition 2 establishes the replication strategy for the realized variance $RV_T^{(3)}$. The replication strategy parallels that of Proposition 1. It consists of (i) the same path-independent payoff OP_T, and (ii) the modified dynamic trading strategy, where the underlying is rebalanced not continuously but on discrete dates. Proposition 1 may be viewed as a special case of Proposition 2 when there are no jumps and $\Delta t \to 0$.

In view of Proposition 2, the payoff $RV_T^{(3)}$ can be decomposed as:

$$RV_T^{(3)} = OP_T + DS_T, \tag{9}$$

where

$$DS_T := 2\sum_{i=1}^{n}\left(\frac{1}{F_{t_{i-1}}} - \frac{1}{F_0}\right)(F_{t_i}-F_{t_{i-1}})$$

is the profit/loss of the dynamic strategy, which rebalances on dates $t_i \in \mathcal{T}$ to maintain $2\left(\frac{1}{F_{t_i}} - \frac{1}{F_0}\right)$ shares of the underlying. In this dynamic trading strategy, the initial position in the underlying is zero. At time t_i, the position is negative if $F_{t_i} > F_0$, and is positive if $F_{t_i} > F_0$. The dynamic strategy can be viewed as contrarian: the position in the underlying is reduced (increased) when the price goes up (down). Intuitively, the strategy realizes profits when the underlying frequently bounces up and down while generally remaining in a narrow range of F_0. On the other hand, the strategy incurs losses when the price trends in either direction. In contrast, the payoff of the portfolio of options OP_T is the higher the farther the underlying moves in either direction from F_0, with the payoff being zero when $F_T = F_0$. Empirically, we will find that the two components OP_T and DS_T are negatively correlated.

While its proof is very simple, Proposition 2 has important implications. First, the result holds for completely general price processes, not only continuous ones. The only maintained assumption is the absence of arbitrage. Second, the result holds for any sampling partition

$\mathcal{T}$. In particular, $\mathcal{T}$ does not have to be equidistant and Δt does not have to be small. Third, the replication strategy for $RV_T^{(3)}$ requires *discrete* rebalancing in the underlying and the proposition identifies the exact timing for rebalancing.

Below we discuss the implications of Proposition 2 in more detail.

2.2.1 Allowing for jumps

As mentioned earlier, an important limitation of the existing approach is its reliance on the assumption of continuous price process. The empirical literature have presented strong evidence that prices of many financial assets can jump. Recent studies for the S&P 500 index include Bakshi, Cao, and Chen (1997), Andersen, Benzoni, and Lund (2002), Pan (2002), Eraker, Johannes, and Polson (2003), Aït-Sahalia (2002), Carr and Wu (2003), among others.

Allowing for jumps is important for many derivatives, but this is especially the case with the variance contract, whose payoff could be very sensitive to extreme (but, perhaps, rare) returns. To see this intuitively, suppose that, on a typical day, the volatility of the underlying return is 1% (which corresponds to about 16% annual volatility). Consider also a day when the underlying jumps by 10%. The contribution of this day to the variance payoff will be equivalent to 100 days with no jumps (or, about 5 months). Similarly, a single day when the underlying jumps by 30% will be equivalent to 900 "normal" days (or, more than 3.5 years).

In recent studies, Jiang and Tian (2005) and Carr and Wu (2007) investigate the effect of jumps on the variance contract. When the underlying can jump, the existing approach can no longer replicate the variance exactly. The two papers derive the theoretical approximation error due to jumps, which is model specific. Both papers focus on the integrated, not realized, variance and assume that continuous monitoring and rebalancing are possible.

2.2.2 It is realized variance that matters

The existing literature focuses exclusively on valuing payoffs based on the integrated variance, which is an unobservable, theoretical quantity. On the other hand, actually traded variance instruments are based on *realized* variance, typically computed from daily returns.[8] The implicit presumption in the literature is that the unobservable integrated variance and the realized variance are close substitutes and that the difference between two can be safely ignored. As we argue in this subsection, however, this presumptions is unjustified for actually traded instruments.

While it is true that the realized variance converges to the integrated variance when Δt tends to zero, the difference between the two can be very large in typical applications. Stated differently, the existing literature develops the replication strategy for the "wrong" payoff.

To clarify this point, suppose that $T = 1$ month, and let $\mathcal{T}$ be the equidistant sampling partition with $\Delta t = 1$ day. Consider a trader who has sold the variance contract with the payoff based on the realized variance RV_T (which could be computed according to any of the three definitions $RV_T^{(1)}$, $RV_T^{(2)}$, or $RV_T^{(3)}$). The trader wishes to dynamically replicate the payoff by following the replication strategy of Proposition 1. In the absence of jumps and transaction costs, the trader is able to perfectly replicate the payoff IV_T. The problem, however, is that the target payoff RV_T will typically differ from the replicated payoff IV_T considerably.

[8] In practice, a derivative contract can only be viable if its payoff is easily verifiable by all counterparties. This requirement effectively precludes using high-frequency data in the definition of the derivative's payoff.

To gauge the accuracy of the replicating strategy, we consider two canonical models: Black-Scholes (1973) and Heston (1993). For the first model (BS), the price process under the objective measure is

$$\frac{dF_t}{F_t} = \mu dt + \sigma dB_t,$$

where B_t is a standard Brownian motion. For the Heston square-root stochastic volatility model (SV), the price process under the objective measure is

$$\frac{dF_t}{F_t} = \mu dt + \sqrt{v_t} dB_t, \tag{10}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma_v \sqrt{v_t}\, dW_t, \tag{11}$$

where W_t is another standard Brownian motion, correlated with B_t such that $E[dB_t dW_t] = \rho dt$. We set the parameters for the two models as

$$BS: \quad \mu = 0.11, \quad \sigma = 0.15, \tag{12}$$

$$SV: \quad \mu = 0.11, \quad \kappa = 5.8, \quad \theta = 0.023, \quad \sigma_v = 0.35, \quad \rho = -0.4. \tag{13}$$

The parameters for the SV model correspond closely to those estimated in Eraker, Johannes, and Polson (2003). For each model, we conduct a Monte Carlo experiment with 5000 repetitions and compute the hedging error (HE), which is defined as the RMSE of the relative error:

$$\text{HE} = \left(E\left[\left(\frac{IV_T}{RV_T} - 1 \right)^2 \right] \right)^{1/2}.$$

The following table reports HE for the two models and the three definitions of the realized variance:

	$RV_T^{(1)}$	$RV_T^{(2)}$	$RV_T^{(3)}$
BS	0.387	0.387	0.387
SV	0.393	0.392	0.392

In all cases, HE is around 39%, which means that the considered replication strategy exposes the trader to huge risks and would be completely unacceptable in practice. The results are very similar across the two models and they are virtually indistinguishable across the three definitions of the realized variance $RV_T^{(1)}$, $RV_T^{(2)}$.

The important conclusion of this subsection is that the replication strategy in Proposition 1, which is designed to replicate the *theoretical* continuously-sampled integrated variance, is simply inappropriate for replicating the *empirically relevant* payoffs sampled at the daily frequency.

The contrast between the integrated and realized variances is the easiest to see for the case of the Black-Scholes model. For this model, the integrated variance is deterministic, $IV_T = \sigma^2 T$, and Proposition 1 essentially describes a complicated way to replicate the constant payoff (i.e., the payoff of a risk-less bond). On the other hand, the *discretely-sampled* payoffs $RV_T^{(1)}$, $RV_T^{(2)}$, and $RV_T^{(3)}$ are random variables. These payoffs are the more uncertain, the larger the sampling interval Δt. However, Proposition 2 allows one to perfectly replicate the uncertain payoff $RV_T^{(3)}$ for any Δt. One way to interpret Proposition 2 is that, although the difference

$RV_T^{(3)} - IV_T$ can become large as Δt increases, the *market price* of that difference is always zero.

2.2.3 Optimal timing for rebalancing

Proposition 2 establishes an interesting fact that, in order to perfectly replicate the payoff $RV_T^{(3)}$, one must rebalance the position in the underlying on a finite number of fixed dates. If the realized variance is computed using the sampling partition $\mathcal{T} = \{t_0, t_1, \ldots, t_n\}$, then rebalancing must take place on the same dates $t_0, t_1, \ldots, t_n$.

On a surface, this fact might appear unremarkable. After all, the path-dependent payoff $RV_T^{(3)}$ only depends on values of the underlying on the sampling dates $t_0, t_1, \ldots, t_n$ and not on what happens to the underlying between them. Is it then surprising that the sampling dates play a special role in the replication strategy?

However, this property is truly unique in option pricing and, *a priori*, it is far from obvious. One can recall that European put and call options, as well as other path-independent options, have payoffs which only depend on the value of the underlying at maturity, F_T. Yet, to replicate these payoffs, continuous trading in the underlying is necessary. Similarly, exotic path-dependent options (such as lookback or Asian options) are often defined using a finite number of fixing dates, but they also require continuous rebalancing. (Furthermore, the above examples require that the market is complete and the true model is known.)

In this subsection, we report on two experiments which illustrate that, if the trader deviates from the optimal rebalancing strategy even slightly, the hedging errors could be surprisingly high. Let $\mathcal{T} = \{t_0, t_1, \ldots, t_n\}$ be an equidistant sampling partition with $\Delta t = 1$ day. This partition is used to compute the realized variance $RV_T = RV(\mathcal{T})$. To replicate this payoff, the trader follows the strategy described in Proposition 2, except the underlying is rebalanced using a different partition $\mathcal{T}' = \{t'_0, t'_1, \ldots, t'_m\}$, where for simplicity we assume that $t'_0 = t_0$ and $t'_m = t_n = T$. This replication strategy generates payoff RV'_T, which is generally different from the target payoff RV_T. As before, we consider the Black-Scholes model and the Heston model with parameters as in (12)-(13). The relative hedging error is now computed as

$$\text{HE} = \left(E\left[\left(\frac{RV'_T}{RV_T} - 1 \right)^2 \right] \right)^{1/2}.$$

In the first experiment, we assume that the partition $\mathcal{T}'$ is equidistant and its sampling interval $\Delta t' = 1$ minute, 10 minutes, 1 hour, 1/2 day, 1 day, 2 days, and 5 days. Table 1 shows that rebalancing too frequently or too infrequently might lead to considerable hedging errors. For example, when $\Delta t' = 1/2$ day (2 days), the average error of the replication strategy is about 26% (29%) for both the Black-Scholes and Heston models.[9] Table 1 also reports related results for the S&P 500 futures based on real data. The hedging errors are now computed using 5-minute returns over the period from 1990 to 2006. For the S&P 500, the average error of the replication strategy is about 26% (27%) when $\Delta t' = 1/2$ day (2 days).

[9]The examples in this and previous subsections illustrate the common property. For a *fixed* partition $\mathcal{T}$, the difference between the three versions of the realized variance $RV_T^{(1)}$, $RV_T^{(2)}$, are $RV_T^{(3)}$ are very small under realistic settings. However, the realized variances $RV_T^{(j)}$ computed for two *different* partitions $\mathcal{T}$ and $\mathcal{T}'$ could differ considerably.

In the second experiment, we assume that the trader rebalances once a day, but not necessarily at the exact times specified by the partition $\mathcal{T}$. Specifically, although actually traded instruments typically use daily closing prices, suppose instead that the trader rebalances at times $\mathcal{T}' = \{t_0, t'_1, \ldots, t'_{n-1}, t_n\}$, where

$$t'_i = t_i - \delta t, \qquad i = 1, \ldots, n-1.$$

Table 2 reports the results for $\delta t = 1$ minute, 5 minutes, 15 minutes, and 1 hour.

The striking finding in the table is that the hedging errors are very sensitive to even slightest non-synchronicity. If the trader executes orders only 1 (5) minutes before the market close, the average error of the replication strategy is non-negligible: 3.1% (7.1%) for the Black-Scholes model and 3.2% (7.1%) for the Heston model. Similarly, using actual 5-minute return for the S&P 500, we find that the average hedging error is 7.4% if the trader rebalances 5 minutes prior to the market close. In other words, deviating even slightly from the optimal rebalancing schedule would expose the trader to considerable risk.

This observation has important implications. As trading in the variance products becomes sufficiently active, there might be a predictable price impact at the market close due to the fact that traders, trying to minimize their hedging risks, will execute their orders at essentially the same time. For related reasons, one might also be concerned about possible price manipulation around the market close. In particular, if the closing price of the underlying is pushed up (down) on days with large positive (negative) returns, then the payoff of the variance contract can be increased by a sizeable amount. This issue might become particularly important for the variance contracts on less liquid individual stocks.

2.2.4 Irregular sampling and high-frequency data

As emphasized in Section 2.2.2, the daily-sampled realized variance can be a very poor estimator of the integrated variance. This insight has been well understood in the literature (see, for example, Andersen and Bollerslev (1998)), and several papers have argued that more accurate estimates of the integrated variance can be obtained by using high-frequency data (see Andersen, Bollerslev, Diebold, and Labys (2003), Barndorff-Nielsen and Shephard (2002), among others). Recent work has focused on constructing optimal high-frequency estimators in the presence of market microstructure noise (Bandi and Russel (2003) and Zhang, Mykland, and Aït-Sahalia (2005)).

One unresolved issue remains in the high-frequency literature: How to estimate the integrated variance when high-frequency returns are only available for a part of each day? In a typical situation, an econometrician can observe high-frequency returns over the regular trading hours, but not over the period when the exchange is closed. For example, the econometrician may collect 5-minute returns on S&P 500 futures over the regular exchange hours (from 9:30 to 15:15 CT) and one additional close-to-open return over the remaining $17\frac{1}{4}$ hours.

The realized variance for the full day can be represented as

$$RV = RVN + RVD,$$

where RVN is the realized variance for the over-night (close-to-open) period and RVD is the realized variance for the intra-day (open-to-close) period. The component RVD can be computed using high-frequency data, however, there is only one over-night return available to compute the component RVN.

What are relative contributions of the two components for the S&P 500 futures? Using high-frequency data over the period from 01/1990 to 12/2006, we find that RV_{oc} and RV_{co} on average represent 78.8% and 21.2% of the total variance RV.[10] Clearly, the component RVN is quite large and it should not be ignored. Intuitively, the amount of the over-night variance is equivalent to 1 hours 49 minutes of the active period. (Per *unit of time*, the over-night variance is equal to about 11% of the variance for the active period.)

In these circumstances, how should one construct the estimator of the total variance? Two approaches have been used in the literature:

(1) discard the over-night return and rescale the high-frequency component RVD by an appropriate constant;

(2) treat the over-night return the same way as high-frequency returns.

Both approaches, however, are not completely satisfactory.[11] The first approach effectively treats over-night returns as pure noise, despite the fact that they clearly contain some useful information about events and news which happen during the inactive hours.[12] On the other hand, the second approach almost defeats the very purpose of using high-frequency data. Even if one uses ultra high-frequency returns and is able to estimate very accurately the integrated variance over the active period, it is the over-night return that determines the (relatively poor) quality of the overall estimator. The asymptotic theory arguments break down when there is one low-frequency return.

The main point of the above discussion is that the unobservability of the integrated variance remains a difficult problem, even when one is armed with quality high-frequency data and assuming that other critical issues (market microstructure noise, presence of jumps, etc.) can be adequately resolved. This underscores the advantage of our approach, which works directly with the easily measurable realized variance, as opposed to its unobservable theoretical counterpart.

In this context, two properties of Proposition 2 are particularly useful. First, our result does not rely on the asymptotic theory and it holds for any sampling frequency, whether high or not. Second, the result allows for *irregular* sampling partitions $\mathcal{T}$, where high-frequency and low-frequency returns can be mixed. The latter is important because it is the case that financial data often arrive sampled at irregular intervals. Among other reasons, this happens because

- open-to-close periods are sampled more frequently than close-to-open periods;
- some close-to-open periods are longer than others, because of weekends, exchange holidays, and other special circumstances;
- transactions data are observed at random time intervals.[13]

[10] Hansen and Lunde (2005) report similar results for the 30 stocks of the Dow Jones Industrial Average, for which they find that about 80% of daily variance occurs during the active period.

[11] Hansen and Lunde (2005) propose an estimator which is a linear combination of approaches (1) and (2).

[12] Perhaps the most extreme case of the "over-night" return is provided by the days following the 9/11 attack. The exchange did not open on Tuesday, 09/11, and the trading resumed only on Monday, 09/17. In this case, the close-to-open period covers almost one week, and high-frequency returns contain very little information on one of the most volatile and dramatic events in the recent history.

[13] See Aït-Sahalia and Mykland (2003) for inference under these circumstances.

2.3 Continuous-Time Limit

To gain some additional intuition, we can look at the three definitions of the realized variance in the continuous-time limit. Suppose that the price process under the objective measure is given by the following general jump-diffusion process:

$$\frac{dF_t}{F_{t-}} = \mu_t dt + \sqrt{v_t} dB_t + \xi_t dN_t(\lambda_t), \tag{14}$$

where $N_t(\lambda_t)$ is a Poisson process, λ_t is jump arrival intensity, ξ_t is a random jump with a possibly time-varying distribution.

When Δt tends to zero, one obtains the three versions of the integrated variance:

$$\begin{aligned}
IV_T^{(1)} &:= \lim_{\Delta t \to 0} RV_T^{(1)} = \int_0^T (d\ln F_t)^2 = \int_0^T v_t dt + \sum_{j=1}^{N_T} \left(\ln(1+\xi_j)\right)^2, \\
IV_T^{(2)} &:= \lim_{\Delta t \to 0} RV_T^{(2)} = \int_0^T \left(\frac{dF_t}{F_t}\right)^2 = \int_0^T v_t dt + \sum_{j=1}^{N_T} \xi_j^2, \\
IV_T^{(3)} &:= \lim_{\Delta t \to 0} RV_T^{(3)} = \int_0^T 2\left(\frac{dF_t}{F_t} - d\ln F_t\right) = \int_0^T v_t dt + \sum_{j=1}^{N_T} 2\left(\xi_j - \ln(1+\xi_j)\right),
\end{aligned}$$

where $\xi_1, \ldots, \xi_{N_T}$ denote the jumps that arrive over the period $[0, T]$.

The integrated variances $IV_T^{(1)}$, $IV_T^{(2)}$, and $IV_T^{(3)}$ differ only with regard to how they treat jumps. When there are no jumps, the three definitions coincide with IV_T. This confirms the fact that Proposition 1 is a special case of Proposition 2 when $\Delta t \to 0$ and when there are no jumps.

To find the market prices of the three integrated variances, one needs to evaluate the risk-neutral expectation $E_0^*[\cdot]$ of the respective payoffs. In general, this expectation is equal to U_0 only for the integrated variance $IV_T^{(3)}$, but not for $IV_T^{(1)}$ and $IV_T^{(2)}$. Moreover, if under the risk-neutral measure negative jumps on average dominate positive jumps (as common in many applications), then

$$E_0^*\left[IV_T^{(1)}\right] > U_0, \qquad \text{and} \qquad E_0^*\left[IV_T^{(2)}\right] < U_0.$$

That is, U_0 understates the market price of the integrated variance $IV_T^{(1)}$ and overstates the market price of $IV_T^{(2)}$.

2.4 The Generalized Variance Contracts

We conclude the theory section with a discussion the *generalized* variance contracts. Let $g(F)$ be a general function, whose second derivative $g''(F)$ is continuous almost everywhere. Suppose initially that F_t follows the continuous process as in (2). Consider a contract which at maturity T pays the *generalized* integrated variance defined as

$$gIV_T = \int_0^T g''(F_t)(dF_t)^2 = \int_0^T g''(F_t)F_t^2\, v_t dt.$$

In the above payoff, the variance is integrated with the weight function

$$\omega(F) := g''(F)F^2,$$

which depends on the price level. One interesting choice for the weight function is

$$\omega(F) = F^a,$$

in which case the payoff is computed as the power-price-weighted integrated variance (PPWIV)

$$gIV_T = \int_0^T F_t^a \, v_t dt.$$

The standard variance contract corresponds to the case of $a = 0$. When $a > 0$, the PPWIV contract's instantaneous exposure to the variance increases with the price level F_t, the property that might be appealing to some traders. In particular, since for the S&P 500 index as well as many other assets, the price and the variance are negatively correlated, the payoff of the PPWIV contract ($a > 0$) is less variable than that of the standard variance contract ($a = 0$). Another potential advantage of the PPWIV contract with $a > 0$ is that its payoff is defined even when the underlying price can drop to zero, unlike the standard variance contract.[14]

Another interesting choice of the weight function is

$$\omega(F) = I(F; B_1, B_2),$$

where $I(F; B_1, B_2)$ denotes the indicator function,

$$I(F; B_1, B_2) := 1[B_1 \leq F \leq B_2].$$

The contract now pays the *corridor* integrated variance (CIV), or the variance calculated only when the price F_t is between the barriers B_1 and B_2:

$$gIV_T = \int_0^T I(F_t; B_1, B_2) \, v_t dt.$$

Carr and Madan (1998) show how to price the CIV contract when the process F_t is continuous and under continuous trading.[15] Andersen and Bondarenko (2007) study several corridor implied volatility measures constructed for the S&P 500. Interestingly enough, they find that narrow corridor measures are more useful predictors of future volatility than broad corridor measures. The special cases of the barriers (B_1, B_2) equal to $(0, F_0)$ and (F_0, ∞) correspond to the down- and up-variance payoffs, respectively.

Many other choices for the weight function can be considered, yielding even more complex, custom-tailored payoffs and allowing investors to bet on even more specific aspects of the variance. For example, one can imaging a contract which pays the corridor power-price-weighted integrated variance. Although the second derivative $g''(F)$ is typically non-negative, in general, it does not have to be. One can also incorporate explicit dependence on time, i.e., $g(F, t)$, but for simplicity we will not consider such an extension.

[14] For the case of $a = 1$, PPWIV contract is sometimes called the gamma contract. In the gamma contract, the exposure to the variance is proportional to the price level F_t. If the market increases by x%, then the variance exposure also increases by x%. The name "gamma" refers to the fact that the gamma contract has a linear cash-gamma, whereas the standard variance contract has a constant cash-gamma. Several banks, including BNP Paribas and Deutsche Bank, have recently started to offer gamma swaps.

[15] In a recent paper, Carr and Lewis (2007) propose an approximate replication strategy for discretely-sampled corridor variance swap.

How can one value the generalized variance contracts? As before, we are interested in the approach which works under empirically relevant discrete-sampling and for general price processes, including those with jumps. For given sampling partition $\mathcal{T} = \{t_0, t_1, \ldots, t_n\}$, we consider three definitions for the generalized realized variance:

$$gRV_T^{(1)} = gRV^{(1)}(\mathcal{T}) := \sum_{i=1}^{n} g''(F_{i-1})F_{i-1}^2 \left[\ln \frac{F_i}{F_{i-1}}\right]^2, \tag{15}$$

$$gRV_T^{(2)} = gRV^{(2)}(\mathcal{T}) := \sum_{i=1}^{n} g''(F_{i-1})(F_i - F_{i-1})^2, \tag{16}$$

$$gRV_T^{(3)} = gRV^{(3)}(\mathcal{T}) := 2\sum_{i=1}^{n} \left[g(F_i) - g(F_{i-1}) - g'(F_{i-1})(F_i - F_{i-1})\right]. \tag{17}$$

The three measures $gRV_T^{(1)}$, $gRV_T^{(2)}$, and $gRV_T^{(3)}$ only differ in terms of order $O(\Delta F_i^3)$ and, in typical applications, are very close to each other. However, only the definition $gRV_T^{(3)}$ admits model-independent replication.

To state the theoretical result, it is convenient to introduce the following function of two arguments:

$$G(x_1, x_2) := 2\left[g(x_2) - g(x_1) - g'(x_1)(x_2 - x_1)\right].$$

With this notation, the definition of $gRV_T^{(3)}$ simplifies to

$$gRV_T^{(3)} = \sum_{i=1}^{n} G(F_{i-1}, F_i).$$

Furthermore, we define the following payoff at maturity T:

$$gOP_T := G(F_0, F_T).$$

As follows from (25), this path-independent payoff can be staticly replicated by a portfolio of standard options:

$$gOP_T = 2\int_0^\infty g''(K)M_T(K)dK.$$

The right hand side represents the payoff of the portfolio of European puts and calls with same maturity T and a continuum of strikes K from 0 to ∞, where portfolio weights are equal to $2g''(K)$.

Proposition 3 *For any partition $\mathcal{T}$, the payoff $gRV_T^{(3)}$ can be perfectly replicated by*

(i) a time-T payoff equal to gOP_T;

(ii) a dynamic trading strategy, which is rebalanced on dates $t_i \in \mathcal{T}$ to maintain $2(g'(F_0) - g'(F_{t_i}))$ shares of the underlying.

Therefore,

$$E_0^*\left[gRV_T^{(3)}\right] = E_0^*\left[gOP_T\right] = 2\int_0^\infty g''(K)M_0(K)dK.$$

Proof: Definition (17) can be re-arranged as

$$gRV_T^{(3)} = 2(g(F_T) - g(F_0)) - 2\sum_{i=1}^{n} g'(F_{i-1})(F_i - F_{i-1})$$

$$= 2(g(F_T) - g(F_0) - g'(F_0)(F_T - F_0)) + 2\sum_{i=1}^{n}(g'(F_0) - g'(F_{i-1}))(F_i - F_{i-1})$$

$$= gOP_T + 2\sum_{i=1}^{n}(g'(F_0) - g'(F_{i-1}))(F_i - F_{i-1}).$$

Since F_t is a martingale under the risk-neutral measure, $E_0^*\left[gRV_T^{(3)}\right] = E_0^*\left[gOP_T\right]$. □

The significance of Proposition 3 should not be diminished by simplicity of its proof. The proposition establishes the model-free replication strategy for a broad class of payoffs $gRV_T^{(3)}$, for completely general price processes, and for arbitrary sampling partitions $\mathcal{T}$. Clearly, Proposition 2 is a special case of Proposition 3 corresponding to $g(F) = -\ln F$.

For a continuous price process F_t, the three measures $gRV_T^{(1)}$, $gRV_T^{(2)}$, and $gRV_T^{(3)}$ converge to gIV_T as $\Delta \to 0$. Specifically, if the price process is given by the jump-diffusion process in (14), then one obtains the three versions of the generalized integrated variance:

$$\begin{aligned}
gIV_T^{(1)} &:= \lim_{\Delta t \to 0} gRV_T^{(1)} = \int_0^T g''(F_t)F_t^2 \,(d\ln F_t)^2 = gIV_T + \sum_{j=1}^{N_T} g''(F_j)F_j^2 \,(\ln(1+\xi_j))^2, \\
gIV_T^{(2)} &:= \lim_{\Delta t \to 0} gRV_T^{(2)} = \int_0^T g''(F_t)F_t^2 \left(\frac{dF_t}{F_t}\right)^2 = gIV_T + \sum_{j=1}^{N_T} g''(F_j)F_j^2 \,\xi_j^2, \\
gIV_T^{(3)} &:= \lim_{\Delta t \to 0} gRV_T^{(3)} = \int_0^T 2\left(dg(F_t) - g'(F_t)dF_t\right) = gIV_T + \sum_{j=1}^{N_T} G(F_j, F_j(1+\xi_j)),
\end{aligned}$$

where $\xi_1, \ldots, \xi_{N_T}$ denote the jumps that arrive over the period $[0, T]$ and F_j denotes the price before jump ξ_j. As before, the three definitions $gIV_T^{(1)}$, $gIV_T^{(2)}$, and $gIV_T^{(3)}$ differ only with regard to how they treat jumps. When there are no jumps, they all coincide with gIV_T.

3 Implementation Issues and Data

3.1 Variance Return

The basic idea of our empirical study is to view the variance as an alternative asset class. Let

$$r_v = \frac{RV_T}{U_0} - 1 \tag{18}$$

denote the simple net return on the variance contract where the underlying is the S&P 500 futures. (Recall that, by construction, r_v is already defined in *excess* of the risk-free rate.)

In the empirical application, we construct monthly variance returns over the period from 01/1990 to 12/2006. Armed with such a time-series, we then can investigate the economic properties of the variance return. In particular, we can compute the market price of the variance risk as the unconditional expectation $\pi_v = E[r_v]$ and use the variance return as a new risk factor for various assets.

To compute the variance return, two quantities are needed. The numerator is the realized variance, which is easy to compute from daily closing prices of the S&P 500 futures. In our application, it makes little difference whether the realized variance is measured as $RV_T^{(1)}$, $RV_T^{(2)}$, or $RV_T^{(3)}$.[16] For the empirical results reported in Section 4, we measure the realized variance according to the definition $RV_T^{(3)}$, in which case the market price of the variance is precisely equal to U_0. To simplify notation, in what follows we write $RV_T = RV_T^{(3)}$.

The denominator in (18) is the market price of the variance contract U_0. In view of Proposition 2, we can infer U_0 from prices of standard S&P 500 options. The intuition here is that, even though the price of the variance contract might not be directly observable, there is only one arbitrage-free price for the variance contract which is consistent with observed prices of S&P 500 options.[17]

An analogy may help to clarify our empirical approach. Suppose for a moment that put options were not traded on the exchange. Instead, an econometrician could only observe market prices of European calls with various strikes and maturities. The econometrician could still study the economic properties of put returns, because synthetic put positions can be constructed from put-call parity. This approach is model-free, as the put-call parity relationship only requires the absence of arbitrage and a frictionless market. Similarly, although the replication strategy for the variance contract is more involved than that for a put option, the price of the variance contract can still be derived in a model-independent fashion from observed prices of traded options. In other words, we can view the variance contract as an (essentially) tradeable security.

To estimate the market value of the variance contract U_0, we use a new method developed in Bondarenko (2003a). The method is termed *Positive Convolution Approximation* (PCA) and it allows one to infer the conditional RND $\hat{h}_0(F_T)$ through the relationship in (1). The method directly addresses the important limitations of option data that (a) options are only traded for a discrete set of strikes, as opposed to a continuum of strikes, (b) very low and very high strikes are usually unavailable, and (c) option prices are recorded with substantial measurement errors, which arise from nonsynchronous trading, price discreteness, and the bid-ask bounce. The PCA method is fully nonparametric, always produces arbitrage-free estimators, and controls against overfitting while allowing for small samples. In addition to the estimate for RND, the method also produces the estimates for the call and put pricing functions $\hat{C}_0(K)$ and $\hat{P}_0(K)$ for a continuum of strikes. Thus, we obtain an estimate $\hat{U}_0$ from (4) by integrating $\hat{C}_0(K)$ and $\hat{P}_0(K)$ over the strike K with the weight function $1/K^2$.

3.2 CME Options

Our data consist of daily prices of options on the S&P 500 futures traded on the Chicago Mercantile Exchange (CME) and the S&P 500 futures themselves. The data are obtained

[16] With daily sampling, we find that the correlations between any two definitions are always above 0.99.

[17] As discussed in Introduction, the price of the variance contract could, in principle, be derived from prices of OTC variance swaps. Unfortunately, this approach is not practical because the OTC market is still rather illiquid, it has a short history, and, most importantly, its transaction prices are not readily available to researches. Similarly, the CBOE futures on the realized variance have an even shorter history and, at the time of writing, their trading volume is almost non-existing.

from the Futures Industry Institute and CME directly.[18] The sample period is from January 1990 through June 2005. The earlier version of the paper used data only up to December 2000. By extending the sample period for additional 6 years, we are able to confirm our finding out-of-sample.

The S&P 500 futures have four different maturity months from the March quarterly cycle. The contract size is $250 times S&P 500 futures price (before November 1997, the contract size was $500 times S&P 500 futures price). On any trading day, the CME futures options are available for six maturity months: four months from the March quarterly cycle and two additional nearby months ("serial" options). The CME options expire on a third Friday of a contract month. The quarterly options expire at the market open, while the serial options expire at the market close. For the serial options, we measure time to maturity as the number of calendar days between the trade date and the expiration date. For the quarterly options, we use the number of calendar days remaining less one. The option contract size is one S&P 500 futures. The minimum price movement is 0.05. The strikes are multiples of 5 for near-term months and multiples of 25 for far months. If at any time the S&P 500 futures contract trades through the highest or lowest strike available, additional strikes are usually introduced.

The CME options on the S&P 500 futures and options on the S&P 500 Index itself, traded on the Chicago Board Option Exchange (CBOE), have been a focus of many empirical studies. For short maturities, prices of the CME and CBOE options are virtually indistinguishable. Nevertheless, there are a number of practical advantages in using the CME options:

- As well known, there is a 15-minute difference between the close of the CBOE markets and the NYSE, AMEX, and NASDAQ markets, where the S&P 500 components are traded. This difference leads to non-synchronicity biases between the recorded closing prices of the options and the level of the Index. In contrast, the CME options and futures close at the same time (3:15 pm CT).

- It is easier to hedge options using highly liquid futures as opposed to trading the 500 individual stocks. On the CME, futures and futures options are traded in pits side by side. This arrangement facilitates hedging, arbitrage, and speculation. It also makes the market more efficient. In fact, even traders of the CBOE options usually hedge their positions with the CME futures.

- Another complication is that the S&P 500 Index pays dividends. Because of this, to estimate the risk-neutral densities from the CBOE options, one needs to make some assumptions about the Index dividend stream. No such assumptions are needed in the case of the CME futures options.

A disadvantage of the CME options is their American-style feature. However, our empirical analysis is conducted in such a way that the effect of the early exercise is minimal.

We estimate $U_0 = U(0,T)$ over calendar months. For example, for month of June, dates 0 and T in the holding period $[0,T]$ correspond to the last business days in May and June, respectively. However, the S&P 500 options expire not at the end, but roughly in the middle of calendar months. (More precisely, the exchange sets the option maturity as the Saturday immediately following the third Friday of the contract month.) Therefore, we obtain $U_0 = U(0,T)$ by linear interpolation between the two neighboring option maturities T_1 and T_2 which

[18]We are very grateful to David Hsieh for help with the data.

straddle the maturity of the variance contract T. Formally, for $T_1 \leq T \leq T_2$, we compute

$$\frac{U(0,T)}{T} = \frac{T_2 - T}{T_2 - T_1} \cdot \frac{U(0,T_1)}{T_1} + \frac{T - T_1}{T_2 - T_1} \cdot \frac{U(0,T_2)}{T_2}, \tag{19}$$

where $U(0,T_1)$ and $U(0,T_2)$ can be directly estimated from traded options via the relationship in (4).[19] Options with maturities T_1 and T_2, which correspond to approximately 15 and 45 calendar days, are among most actively traded contracts.

Overall, there are $N = 204$ one-month periods in the sample period. Specific steps for calculating the variance return r_v are explained in Appendix B. Briefly, those steps include 1) filtering out unreliable option data, 2) checking that option prices satisfy the theoretical no-arbitrage restrictions, 3) inferring forward prices of European puts and calls, 4) estimating RND, the put and call pricing functions for a continuum of strikes, and 5) estimating the market price of the variance contract using the relationships in (4) and (19).

3.3 Monte-Carlo Experiment

To assess the empirical performance of the estimator $\hat{U}_0$, we conduct a Monte-Carlo experiment in conditions that closely approximate real data. The specific details are provided in Appendix C. In the experiment, we start with a flexible parametric model for the price process F_t, whose parameters are chosen to calibrate a typical cross-section of the S&P 500 options. Theoretical call and put prices are generated for same strikes which were actually available on the selected day. Then, we create simulated option prices by adding random noise to the theoretical prices. The purpose of adding random noise is to model observational errors that arise from various market imperfections. We consider three cases: No noise, Moderate noise, and High noise. In these cases, the level of noise is varied in relation to typical bid-ask spreads of traded options.

Simulated option prices are generated 5000 times. Each time, the estimator $\hat{U}_0$ is obtained via the PCA method. Its accuracy is measured according to the normalized RMSE criterion:

$$\text{RMSE}(\hat{U}_0) = \frac{1}{U_0} \left(E\left[\left(\hat{U}_0 - U_0 \right)^2 \right] \right)^{1/2}, \tag{20}$$

where the true value of U_0 is known in simulations.

The results of the Monte-Carlo experiment can be summarized as follows:

	No noise	Moderate noise	High noise
RMSE($\hat{U}_0$):	0.008	0.014	0.021

Even when there is no noise in option prices, RMSE is strictly positive. In this case, the estimation errors are introduced because options are unavailable for very high or low strikes and because of interpolation in (19). Naturally, the accuracy of the estimator declines when the level of noise increases. Nevertheless, we conclude that the procedure for estimating the market value of the variance contract is quite accurate, with RMSE being just 1.4% (2.1%) for Moderate (High) noise.

[19] When calculating its new VIX, CBOE relies on similar interpolation. In the previous version, we computed the variance return from one option maturity to the next one, so that no interpolation was needed. However, many financial series are only available for calendar months. One important example is returns of hedge funds. Therefore, it is convenient to measure the variance return in the same way. The change in the empirical design has little effect on the results reported in Section 4.

4 Statistical Properties of the Variance Return

4.1 The Market Price of the Variance Risk

Table 3 reports various statistics for monthly returns on the variance contract, the S&P 500 futures, and two additional portfolios which will be discussed shortly. We denote as

$$r_m = \frac{F_T}{F_0} - 1$$

the excess return on the S&P 500 futures contract, which is interpreted in this paper as the return on the market portfolio. The average excess return (AER) for the variance contract is -31.4% per month, indicating that the variance risk is priced and its risk premium is negative. Over the same period, AER for the market is about 0.55% per month.

The negative risk premium π_v on the variance contract return is partly due to negative correlation between r_v and r_m. To separate the effect of negative correlation, Panel B of Table 3 reports the Jensen's alpha and beta coefficients with respect to the market return. Specifically, for any return r_i, we compute the alpha and beta coefficients as

$$\alpha_i = E[r_i] - \beta_i E[r_m], \qquad \beta_i = \frac{Cov(r_i, r_m)}{Var(r_m)}.$$

The beta coefficient for r_v is -4.99, which accounts for about -2.72% of the risk-premium π_v; the alpha coefficient is -28.52% per month. High leverage of the variance contract complicates interpretation of its alpha coefficient. Therefore, Panel B of Table 3 reports several risk-adjusted measures that are unaffected by leverage:

- the Sharpe ratio, $SR := \frac{E[r_i]}{\sqrt{Var(r_i)}}$,
- the Treynor's measure, $TM := \frac{\alpha_i}{\beta_i}$, and
- M-squared of Modigliani and Modigliani (1997), $M^2 := \frac{E[r_i]}{\sqrt{Var(r_i)}} \sqrt{Var(r_m)}$.

In particular, the table demonstrates that selling the variance contract produces a very high Sharpe ratio of 0.70, which is about 5 times as large as the Sharpe ratio for the market over the same period. The M^2 statistics for selling the variance contract is 2.81% per month. Intuitively, this statistics shows the return that an investor would earn if the variance contract is de-leveraged with the risk-free asset to match the standard deviation on the market.

Additional insights are provided by Figures 2-4. Figure 2 shows the S&P 500 level and the daily return on the underlying S&P 500 futures. In Figure 3, the top panel shows the distribution of the variance return over time. It also indicates months with largest positive variance returns: August 1990, October 1997, August 1998, September 2001, and July 2002. The bottom panel of Figure 3 contrasts time-series for the market price of the variance contract and for the realized variance. For most months, the market price of the variance contract substantially exceeds the realized variance and the variance return is negative. Occasionally, however, the realized variance spikes up and the variance contract realizes a high positive return. In Figure 4, the top left panel is the histogram of the variance return r_v. The histogram demonstrates that the distribution of the variance return is highly non-normal. In particular, the variance return exhibits substantial positive skewness. The top right panel of Figure 4 is a scatter plot of the variance return r_v versus the market return r_m, which confirms the fact that returns r_v and r_m are negatively correlated (the correlation coefficient is -0.45).

Overall, the reported results suggest that selling the variance contract would have resulted in very high profits over the studied 17-year period. The profits are high on the risk-adjusted basis, where the Sharpe ratio or related risk measures are used. Table 3 analyzes monthly returns on two special portfolios, the *market-neutral portfolio* (MNP) and the *mean-variance portfolio* (MVP). In MNP, the short position in the variance contract is hedged with the S&P 500 futures in such a way that the sample correlation between the portfolio's return and the market return is zero. That is, the MNP's return is given by

$$r_{mn} = -r_v + \beta_v r_m = -r_v - 4.99_m.$$

MVP is also formed from the variance contract and the S&P 500 futures. Now, however, the portfolio weights are chosen by the mean-variance optimization to achieve the highest possible Sharpe ratio. Its return is

$$r_{mv} = -0.24 r_v - 0.76 r_m.$$

For MVP, the Sharpe ratio is 0.73 and the M^2 measure is 2.92. Figure 5 compares cumulative returns of the four strategies that invest in 1) the S&P 500 futures, 2) the short variance contract, 3) MNP, and 4) MVP. To make the comparison meaningful, the last three strategies are de-leveraged with the risk-free asset so that the standard deviations of their monthly returns over the sample period are the same as for the first strategy. Figure 5 shows that the cumulative returns for strategies 2)-4) are much higher than that for strategy 1).

We also find it reassuring that, after extending the original sample for additional 6 years, our results change very little. For example, the mean variance return is -31.31% for the original period 01/1990–12/2000 and -31.12% for 01/2001–12/2006.

In concurrent work, Carr and Wu (2007) implement a related methodology to estimate the variance risk premium. In particular, they too estimate the risk premium for the 1-month S&P 500 realized variance and find it highly negative. However, their risk premium estimate differs substantially from ours and it might be instructive to reconcile these findings.

Carr and Wu use a much shorter 7-year period (01/1996 to 02/2003) and there are some differences in the empirical design and estimation of the market price of the variance contract. However, the main reason for the differing estimates is due to the different definitions of the risk premium. Carr and Wu measure the risk premium in two ways.

In the first approach, they employ the log transformation for the variance return (i.e., $\log(1 + r_v)$) and find that the average log variance return (termed the *log variance premium*) is -66% per month. This figure is much lower than our estimate of -31.24% for the average simple return. However, because of a very high variability of the variance return and the effect of the Jensen's inequality, the average log return *should* be much lower than the average simple return. Carr and Wu's primary motivation for the log transformation is to make the distribution closer to normal, but the downside is that, for highly leveraged securities, such as the variance contract and options, the mean log return can no longer be interpreted as a risk premium, even approximately. In those circumstances, the log transformation induces a very strong downward bias. Put differently, even if the variance risk is not priced, the mean log variance return will be not zero, but highly negative.[20]

In the second approach, Carr and Wu compute the variance risk premium based on the variance swap payoff

$$VS_T := RV_T - E_0^*[RV_T] = RV_T - U_0.$$

They find that the sample average of VS_T is -0.0274 (expressed in annualized variance units). When interpreting this estimate one should keep in mind that, unlike the variance return r_v, the variance swap payoff $VS_T = r_v U_0$ is not normalized by the overall level of the variance. Most parametric models predict that VS_T is related to the level of variance[21] and we find strong empirical support for this fact: the variance swap payoff VS_T is on average much larger (in absolute term) for high volatility periods then for low volatility ones. In contrast, the variance return r_v is much more homoscedastic and less dependent on U_0. The results of the following two regressions support this claim (t-statistics are shown in the second row):

$$\begin{array}{ccccccc} VS_T & = & 0.0008 & + & (-0.32)\ U_0, & & \bar{R}^2 = 16.3. \\ & & (0.34) & & (-6.37) & & \\ & & & & & & \\ r_v & = & -34.13 & + & 75.5\ U_0, & & \bar{R}^2 = -0.25. \\ & & (-6.64) & & (0.71) & & \end{array}$$

Intuitively, our definition of the variance risk premium, $E[r_v]$, corresponds to the average excess return of the strategy which every month has the same *dollar* exposure in the variance contract. In contrast, $E[VS_T]$ corresponds to the average excess return of the strategy which holds the same *number of shares* of the variance contract. The latter strategy has exposure to the variance risk which varies from month to month with the price of variance contract U_0. Importantly, the price U_0 changes dramatically over our sample period. For example, U_0 was equal to 0.0093 in February 1994 and 0.245 in September 1998, corresponding to an increase of more than 26 times (in standard deviation units, 0.097 and 0.495, respectively). This fact has two implications. First, the sample average for the variance swap payoff VS_T is very sensitive to the chosen period. Second, the sample average for VS_T is dominated by months with high volatility with the contribution of low volatility months being very small.

[20]The effect of the log transformation is most extreme for options, for which simple returns could take value of -100% and the average log return is $-\infty$.

[21]For example, consider the classical Heston stochastic volatility model. The variance v_t evolves under the objective measure as in (11) and under the risk-neutral measure as

$$dv_t = \kappa^*(\theta^* - v_t) + \sigma_v \sqrt{v_t} dW_t^*,$$

where $\kappa^*\theta^* = \kappa\theta$, $\kappa^* = \kappa + \eta_v \sigma_v^2$, and η_v is the market price of risk due to shocks in the Brownian motion. With continuous sampling, one can compute the conditional risk premium $E_0[VS_T] = E_0\left[IV_T\right] - E_0^*\left[IV_T\right]$ as

$$E_0\left[VS_T\right] = T\left(\left(1 - \frac{1-e^{-\kappa T}}{\kappa T}\right)\theta + \frac{1-e^{-\kappa T}}{\kappa T} v_0 - \left(1 - \frac{1-e^{-\kappa^* T}}{\kappa^* T}\right)\theta^* - \frac{1-e^{-\kappa^* T}}{\kappa^* T} v_0\right).$$

For small κT and $\kappa^* T$, the above expression can be approximated as

$$E_0\left[VS_T\right] \approx T\left(\frac{\kappa T}{2}\theta + \left(1 - \frac{\kappa T}{2}\right) v_0 - \frac{\kappa^* T}{2}\theta^* - \left(1 - \frac{\kappa^* T}{2}\right) v_0\right) = \frac{T^2}{2}\eta_v \sigma_v^2 v_0,$$

which indicates that the magnitude of $E_0\left[VS_T\right]$ increases with the instantaneous variance v_0. For small maturities T, the relationship is approximately linear.

4.2 Comparison to Put and Call Options

One might expect a high correlation between returns on the variance contract and put options. Similar to the variance contract, puts earn considerable negative risk premiums and exhibit positive skewness. See, for example, Bondarenko (2003b). Still, there are substantial differences in returns on the variance contract and put options.

Table 4 reports various statistics for monthly excess returns of several options on the S&P 500 futures. Specifically, we consider puts and calls with different moneyness as well as the ATM straddle, which is often viewed as a bet on future volatility. We compute returns on standardized puts and calls, which at the beginning of each period have constant moneyness $k_0 = K/F_0$ equal to 0.96, 0.98, 1.00, 1.02, and 1.04, and which mature on the same date as the corresponding variance contract. Prices of the options with fixed moneyness and one-month maturity are obtained via interpolation.[22]

Compared to the variance contract, puts generally exhibit a higher degree of leverage and generate more extreme returns. In particular, put returns have much larger standard deviations. As k decreases, put returns become more skewed and have more negative means. The Sharpe ratio, M^2, and the Treynor's measure for selling puts are the higher, the lower k. Still, selling the variance contract results in higher Sharpe ratio, M^2, and the Treynor's measure than for any put option. In particular, the Sharpe ratio for selling the variance is almost 3 times as large as that for selling the at-the-money put (ATM) with k=1.00.

As expected, correlations between the variance and put returns are positive, but they are not too high (ranging from 0.50 to 0.57). The largest correlation (0.57) is between the variance contract and the ATM put and the 2% OTM put. It worth pointing out that the correlations of puts with the market return are always higher (in absolute value) than the correlations with the variance return. Call options are negatively correlated with the variance contract. The correlations are generally smaller (in absolute value) than those for puts. Calls too have stronger correlations with the market return than with the variance return. The correlation between the variance contract and the ATM straddle is 0.35. The bottom panels in Figure 4 show two scatter plots: the variance return r_v versus the return of the ATM put, and the variance return r_v versus the return of the ATM straddle.

All in all, the variance return seems to contain quite different information from that contained in option returns. To see this more clearly, consider any option that matures at time T. The option payoff at maturity is *path-independent*, because the payoff is a function of the final value of the underlying F_T only. In contrast, the payoff of the variance contract is *path-dependent*. In other words, while F_T completely determines payoffs of all puts and calls that mature at time T, one needs the whole history of F_t prior to T in order to determine the payoff of the variance contract. This also means that even a collection of infinitely many options with a continuum of strikes does not span the variance return.

To gain additional insights, we use equation (9) to decompose the variance return as

$$r_v = \left(\frac{OP_T}{U_0} - 1\right) + \frac{DS_T}{U_0} = r_{op} + r_{ds},$$

[22] In other words, the approach for constructing option returns is consistent with that for the variance return. To compute option returns over the calendar month $[0, T]$, we estimate RNDs at time t for the two closest available maturities T_1 and T_2 ($T_1 < T < T_2$), compute Black-Scholes implied volatilities for the seven fixed levels of moneyness, obtain the corresponding implied volatilities for T, and then convert them into option prices.

where r_{op} is the return on the option portfolio (i), and r_{ds} is the normalized profit/loss of the zero-cost dynamic strategy (ii). The two components are compared in the following table:

	Mean	SD	Skew.	Kurt.	α	β	Correlations r_v	r_{op}	r_{ds}	r_m
r_v	-31.24	44.71	2.19	9.04	-28.52	-4.99	1.00	0.44	0.18	-0.45
r_{op}	-46.44	75.30	3.23	17.79	-44.49	-3.58	0.44	1.00	-0.80	-0.19
r_{ds}	15.94	67.65	-1.45	9.04	16.71	-1.41	0.18	-0.80	1.00	-0.08

The results can be summarized as follows:

(a) The average excess return for the variance contract is decomposed into -44.49% from the option portfolio and 16.71% from the dynamic strategy, indicating that losses from the option portfolio are partially offset by profits from the dynamic strategy.

(b) The standard deviations for r_{op} and r_{ds} are much larger than that for r_v.

(c) Both r_{op} and r_{ds} have negative correlations with the market return r_m. However, their correlations are lower (in absolute value) than the correlation between r_v and r_m.

(d) The correlation between the variance return r_v and the return on the static option portfolio r_{op} is not particularly high (0.44). (For comparison, this correlation is lower than correlations of r_v with ATM or OTM puts.)

(e) The two components r_{op} and r_{ds} have a strong negative correlation (-0.80), the fact that helps to explain findings (b)-(d).

4.3 Relation to Other Risk Factors

Table 5 investigates whether the variance risk premium continues to be significantly negative in the presence of other known risk factors. The table reports the results of several OLS specifications where r_v is regressed onto the market return, Fama-French size and value factors, the momentum factor, as well as returns of four options: ATM and OTM puts ($k = 1.00$ and $k = 0.96$) and ATM and OTM calls ($k = 1.00$ and $k = 1.04$).

In all regressions, the intercept, or alpha coefficient, remains highly negative and statistically significant. In particular, in the specification with r_m, SMB, HML, and UMD, all coefficients come out negative. The intuition for this result is as follows. When the market variance unexpectedly increases, the overall market performs poorly in general, however, small stocks, value stocks, and (to a lesser degree) recent winners tend to do even worse. After accounting for the standard risk factors, the alpha coefficient only slightly declines in absolute value (-24.98%). This suggests that these factors can account for only small portion of the variance risk premium π_v.

Similarly, option returns also cannot explain the variance risk premium. In the specification with r_m, SMB, HML, UMD, and the four option returns, the alpha coefficient still remains highly negative (-18.70%) and statistically significant. This supports the conclusion of the previous subsection that options returns do not span the variance return very well.

4.4 Portfolio Choice Implications

To put in perspective the economic value of the variance as a new asset class, we consider an investor who maximizes the expected value of a Constant Relative Risk Aversion (CRRA) utility function

$$u(W_T) = \begin{cases} \frac{1}{1-\gamma} W_T^{1-\gamma} & \text{if } \gamma \neq 1 \\ \log W_T & \text{if } \gamma = 1, \end{cases}$$

where W_T is the final wealth and $\gamma > 0$ is the coefficient of risk aversion. Suppose that the investor can only invest in the market portfolio and the variance contract and let w_m and w_v denote the corresponding portfolio weights. Then, the final wealth can be represented as

$$W_T = (1 + r_f + w_m r_m + w_v r_v) W_0,$$

where r_f is the risk-free rate and W_0 is the initial wealth. (Recall that returns r_m and r_v are defined in excess of the risk-free rate.) The investor chooses the portfolio weights w_m and w_v to maximize $E[u(W_T)]$. Because of homotheticity of CRRA preferences, this is equivalent to maximizing $E[u(1 + r_f + w_m r_m + w_v r_v)]$.

Panel A of Table 6 reports the optimal portfolio weights computed using the historical distribution of returns. The investment horizon is one month and the coefficient of risk aversion γ is set to 1, 2, 3, 5, 10, 20, and 50. The portfolio weights are reported for two cases when the weight w_v is 1) constrained to zero and 2) unconstrained. By contrasting the two cases, we can quantify the economic value to the investor of being able to trade the variance contract. Specifically, we compute the certainty equivalent rate (CER), which measures how much the investor is willing to pay for the opportunity to trade the variance contract. The table reports CER as a percentage of the investor's initial wealth.[23]

For the constrained case, the investor with CRRA preferences is always long the market. However, for the unconstrained case, the investor is short both the market and the variance contract. Shorting the market now serves as a hedge for shorting the variance contract. As risk aversion increases, the portfolio weights (in absolute value) monotonically decline. The economic value of introducing the variance contract is very substantial. For example, an investor with \$1 mln of the investable wealth and the risk aversion coefficient of 3 (10) is willing to pay up to about \$52,000 (\$18,000) per month for being able to sell the variance contract.

Next, we verify that the economic value of introducing the variance contract remains large even when the investor can also trade options on the market. In Panel B of Table 6, we repeat the previous analysis but now assume that the investor can also trade the ATM put option. The final wealth is now represented as

$$W_T = (1 + r_f + w_m r_m + w_p r_p + w_v r_v) W_0,$$

where w_p and r_p are the portfolio weight and return on the ATM put. In this case, the investor is short the market, the ATM put, and the variance contract. When the variance contract is available, the negative position in the ATM put generally becomes much smaller, while the negative position in the market remains about the same and sometimes decreases slightly. Interestingly, the weights w_v are very similar to those reported in Panel A. That is, although

[23]Here we follow Driessen and Maenhout (2003), who have used a similar approach to measure the economic value of introducing various option strategies to the investment opportunity set.

the presence of the ATM put return has a large effect on the market weights w_m, there is almost no effect on the variance weights w_v.

The economic value of introducing the variance contract is still considerable. An investor with \$1 mln of the investable wealth and the risk aversion coefficient of 3 (10) is now willing to pay about \$43,000 (\$15,000) per month to be able to sell the variance contract.[24]

4.5 Power-Price-Weighted Variance Contracts

As our final exercise, we study performance of several generalized variance contracts. Table 7 reports various statistics for monthly excess returns of the power-price-weighted variance contracts (PPWVC), for which the payoff $RV_T^{(3)}$ is defined in (17) with the weight function $\omega(F) = F^a$. The market price of PPWVC is computed using Proposition 3 for different values of the power a.

As the power a increases from -3 to 5, the return on PPWVC gradually becomes less variable and its distribution becomes less non-normal (lower skewness and kurtosis). Moreover, the correlation with the market return r_m and beta both decrease (in absolute value). These results are to be expected, since the level of the S&P 500 index and the variance are negatively correlated. As a increases, selling the PPWVC becomes slightly less profitable. In particular, the sample average return of PPWVC and its Sharpe ratio become less negative. These findings can be partially explained by the fact that, as a increases, the option portfolio gOP_T used in the replication strategy for the payoff $RV_T^{(3)}$ puts less weight on more "expensive" OTM puts and more weight on "cheaper" OTM calls.

5 Discussion

In the next section, we argue that the variance return is a new risk factor, which helps to explain behavior of various assets. Before that, however, it may be useful to address some common questions regarding the variance contract and its return.

5.1 Alternative Measures of Variance and Volatility

The variance return r_v introduced in this paper is the return on the contract which pays the future realized variance. However, many alternative measures of variance or volatility are available. Those include the VIX index, ATM implied volatility, local volatility, range-based estimators of volatility, GARCH variance, realized semi-variance, and others. To various degrees, these quantities are all related to each other, and it is natural to expect that they are all priced by the market.

The distinguishing feature of the realized variance $RV_T^{(3)}$ is that it admits a completely general, model-free pricing. Proposition 2 allows one to obtain the price of the variance contract, even when the variance contract is not traded. In contrast, to value payoffs based on the alternative measures, one would have to write down a formal model and thus to rely on strong parametric assumptions. Moreover, because those are very complex, path-dependent payoffs, the valuation might be particularly sensitive to model misspecification. This paper focuses on

[24]When we repeat the same analysis assuming that the investor can, in addition to the ATM put, also trade the 4% OTM put, the results for CER are quantitatively similar.

the realized variance $RV_T^{(3)}$ not because it is superior to the alternatives, but because it is the only measure for which accurate estimation of its return is possible.[25]

5.2 Comparison to VIX

One of the most popular benchmarks of the equity market volatility is the CBOE VIX index. The index is widely followed by market participants and is often referred to as the "investor fear gauge". This index has been used in numerous academic studies.[26]

How different is behavior of the variance return compared to that of VIX? Or, perhaps, the correlation between the two series is close to 100%? If so, then what would be the empirical advantage of using the variance return?

Table 8 reports correlations between r_v, r_m, and several other measures of interest. In particular, it shows that the correlation between r_v and VIX is relatively low (0.41). The correlation is higher between r_v and the first difference ΔVIX (0.62), although it is still far from perfect. We also note that the correlation between ΔVIX and r_v is even smaller (in absolute value) than the correlation between ΔVIX and r_m (-0.67).

Below we summarize some differences between the variance return r_v and the VIX index:

- The important advantage of r_v is that it is a return on a tradeable strategy. This means that one can use the sample average of r_v as an estimate of the unconditional variance risk-premium. Moreover, in a regression analysis with tradeable factors, the intercept can be interpreted as the abnormal performance (or alpha).

  In contrast, the VIX index does not correspond to a tradeable strategy. VIX is a forward-looking, risk-neutral measure, which reflects both investors' beliefs about the future variance and investors' risk preferences. Changes in VIX can be due to changes in the perceived level of the future variance and/or changes in risk attitudes.[27]

- To compute the variance return two quantities are needed. The denominator in (18) is the market price of the variance contract U_0 at the beginning of the trading period. Both U_0 and VIX_0 are derived from market prices of options and are closely related.[28] However, the numerator in (18) is the realized variance RV_T, which can be quite different

[25] In principle, one can envision even more complex and exotic contracts, whose payoffs are based on higher moments (skewedness, kurtosis), cross-moments (covariance, coskewedness), and others. It would be interesting to know if and to what extent those quantities are priced by the market. However, unless such contracts are actually traded and their prices are available over a reasonably long history, these questions would be difficult to answer in a general, model-free setting.

[26] Currently, there are two versions of the index. In September 2003, CBOE made several changes to the original methodology of the volatility index. The new version retained the original name of VIX, while the old version was renamed to VXO. CBOE has calculated the new VIX back to 1990 from historical option prices. The two versions are highly correlated, with a correlation coefficient of 0.985. In all our illustrations, we use the new VIX.

[27] A close alternative to the approach of our paper would be to replace the variance return with the return on the VIX *futures*, which CBOE began trading in Spring 2004. Unfortunately, this market also has a very short history and, at the time of writing, its liquidity is very low. Furthermore, the return on the VIX futures might be harder to interpret as it incorporates two layers of risk-premiums: one in the VIX index itself and the other in the futures price.

[28] The methodology that CBOE uses to compute the new VIX is based on the result of Dupire (1993) and Neuberger (1994) and, conceptually, VIX_0 is equivalent to $\sqrt{U_0}$. However, Jiang and Tian (2007) point out that the CBOE's implementation has important flaws, which lead to systematic biases in the VIX. They argue that studies based on the VIX may misestimate the variance risk premium.

from VIX_T. This is the primary reason why the two series, r_v and ΔVIX, contain rather different information and exhibit rather different behavior.

- In the empirical applications, such as performance evaluation for hedge funds, we find that if r_v is replaced with, say, ΔVIX, many results become weaker or disappear completely. (Table 8 also demonstrates that the variance return r_v has stronger correlations with proxies for the credit, liquidity, and correlation risks.)

5.3 Alternative Volatility Instruments

A unique feature of the variance contract is that it allows investors to obtain *pure* exposure to the future variance, which is independent of changes in the underlying price and other factors. Recently, several other instruments have been proposed with the objective of hedging the variance or volatility risk, including

- at-the-money option straddle – Coval and Shumway (2001),
- at-the-money-forward straddle – Brenner, Ou, and Zhang (2006),
- options on a volatility index, such as the CBOE Volatility Index (VIX) – Grunbichler and Longstaff (1996).

Common shortcomings of these alternative instruments are that 1) the sensitivity with respect to the variance (or volatility) is nonlinear, 2) sensitivities to changes in other factors (the underlying price, volatility of volatility, etc.) are typically nonzero, and 3) sensitivities are model-specific and/or time-varying. As an example, consider the at-the-money straddle. When the straddle is first created, it is approximately delta-neutral and its price is primarily determined by the expected future variance. However, as the underlying drifts away from its initial level, the straddle is no longer delta-neutral and is now affected by directional movements of the underlying. (Table 4 reports that the correlation between the variance return and the return on ATM straddle is relatively low 0.36.)

In contrast, the variance contract allows one to isolate the variance risk from other risks. In particular, at any time the variance contract has zero sensitivity to changes in the level of the underlying F_t, i.e., it is always delta-neutral. This property is important in applications, because changes in the level and the variance of the S&P 500 index exhibit a non-trivial correlation.

6 Applications

The variance risk plays an important role in the economy for several reasons:

- The variance risk is likely to be related to the credit, liquidity, and correlation risks. Typically, when the variance increases abruptly, credit spreads become wider, liquidity dries out, and correlations between different assets go up.
- The variance often "spills over" across different regions and asset classes. The most severe market turmoils are associated with the so called "flights to quality".

- Various financial institutions might be sensitive to the variance risk. In particular, many institutions do well in quite markets, but lose money in volatile ones. This pattern indicates a negative exposure to the variance risk.

Below we discuss different implications of our analysis and various uses for the variance return.

6.1 Credit Risk

Table 8 documents that the correlation between the variance return and the change in the credit spread ΔCS is 0.27. The credit spread is defined as the difference between the Moody's BAA yield and the 10-year Treasury constant maturity yield, or $CS = BAA - T10Y$.[29] (It appears that the correlation between r_v and ΔCS has increased recently. Over the subperiod from 01/2000-12/2006, the correlation is 0.52.)

Explaining the historical behavior of credit spreads on investment grade bonds has always been a difficult problem. As Chen, Collin-Dufresne, and Goldstein (2006) argue, standard structural models for credit spreads face two main challenges:

(i) *The level puzzle* – the magnitude of observed credit spreads is very hard to rationalize in the context of standard structural models. When standard models are calibrated to match historical default probabilities (which are relatively low) and recovery rates (which are relatively high), the models all produce about the same (and too narrow) spreads. See Huang and Huang (2003).

(ii) *The time-variation puzzle* – the standard models predict counterfactually low time-series variation in spreads. Intuitively, in these one-factor models, the spread is a function of the firm's leverage only, which does not vary over time that much or that fast.

Stochastic variance can help in explaining puzzles (i) and (ii). Intuitively, consider a model which predicts that spreads are positively related to the level of the market variance. If spreads are sufficiently sensitive to changes in the variance, they will exhibit substantial time-variation. Moreover, spreads will reflect both (1) the probability of default *and* (2) the risk premium due to the correlation with the variance. Even when defaults are relatively rare, the contribution to spreads of effect (2) can be still substantial. This is due to the fact that, empirically, the variance risk premium is negative (i.e., of the required sign) and very large. The finding that the variance return is correlated with ΔCS is consistent with this intuition. We leave development of a formal model for future research.

6.2 Liquidity Risk

The variance risk is likely to be related to the liquidity risk. In volatile market conditions, liquidity often dries up. During most severe market turmoils, bid-ask spreads for some securities may become extremely large, and reliable quotes may not exist at all. For instance, this is how *The Economist* described the October 1997 liquidity crisis in Brazil:

[29]Historical yields are downloaded from the website of the Board of Governors of the Federal Reserve System, http://www.federalreserve.gov.

> Debt prices swung wildly, from Latin America to Russia and Eastern Europe... Even worse, those markets were so illiquid that it become impossible to gauge the going price. One London investment bank, asked to quote yields for a few of the more heavily traded bonds, did not care.
>
> *The Economist,* "A week on the wild side," November 1-7, 1997, p. 78.

The variance return in October 1997 was 174% (while the market return r_m was -3.2%), which provides at least anecdotal support for the link between the variance and liquidity risks.

Liquidity, however, is not easy to quantify. It is a broad concept, for which there exist a number of definitions. One measure of liquidity is proposed in Pastor and Stambaugh (2002). For the US equity market, they construct the liquidity factor LIQ based on the idea that order flow induces a temporary price impact. Table 8 reports that there is a negative correlation (-0.11) between the variance return and LIQ. (After 01/2000, the correlation is -0.35).[30] In other words, the level of liquidity generally tends to decline in those months for which the variance return is high.

6.3 Correlation and Covariance Risk

A number of recent empirical studies have documented that correlations among asset returns are time-varying. Generally, correlations tend to increase with the market variance. During the periods of market turmoil, the variance "spills" over across different regions, industries, and asset classes. Asset returns become more correlated and the benefits of diversification are reduced precisely when they are needed the most. This suggests that investors might require compensation for bearing the correlation risk and that the correlation risk might be closely related to the variance risk.

To illustrate this point, we construct a measure of average correlation between stocks using daily returns of the 30 industry portfolios.[31] Within each month, we compute all pairwise time-series correlations and denote their mean as $\bar{\rho}$. Over the sample period 01/1990-12/2006, the average correlation $\bar{\rho}$ has the mean of 0.48 and the standard deviation of 0.14. Interestingly, the change in $\bar{\rho}$ has a rather high correlation of 0.60 with the variance return, see Table 8. That is, correlations generally tend to increase in those months for which the variance return is high. This observation also suggests that the correlation risk should be priced and that it should have a negative risk premium. The recent empirical studies of Driessen, Maenhout, and Vilkov (2006) and Krishnan, Petkova, and Ritchken (2006) confirm this conclusion.

To formalize the link between the variance and correlation risks, consider the following simple model. Suppose that under the objective measure the price for stock-i evolves according to a generalized market model

$$\frac{dS_{i,t}}{S_{i,t}} = \alpha_i + \beta_i \frac{dM_t}{M_t} + \gamma_i dB_{i,t}, \tag{21}$$

where M_t denote the market portfolio and $B_{i,t}$ is the standard Brownian motion which models the firm-specific risk. The process $B_{i,t}$ is independent of M_t and other stocks. The market

[30] It should be pointed out that, because the factor of Pastor and Stambaugh is intended to measure liquidity of the US stocks, the factor may be less informative with respect to global liquidity crises similar to the one discussed above. For example, in October 1997, LIQ was about 0.5 standard deviations below its mean, indicating only a relatively modest drop in liquidity.

[31] The industry returns are downloaded from Ken French's web site.

portfolio can follow a general dynamic, which is left unspecified. In particular, the market can have time-varying, stochastic variance v_t. To simplify exposition, however, we assume the absence of jumps and continuous sampling.

The integrated covariance between stocks i and j is given by

$$ICOV_{ij}(0,T) = \int_0^T \frac{dS_{i,t}}{S_{i,t}}\frac{dS_{i,t}}{S_{i,t}} = \beta_i\beta_j \int_0^T \left(\frac{dM_t}{M_t}\right)^2 = \beta_i\beta_j IV_M(0,T), \tag{22}$$

where IV_M is the integrated variance for the market:

$$IV_M(0,T) = \int_0^T v_t dt.$$

It is clear from (22) that the risk premium for the integrated covariance is equal to the risk premium for the market integrated variance times a constant. If the market variance has a negative risk premium, the covariance between the two stocks with positive betas will also have a negative risk premium.

The situation is similar with the correlation risk. In particular, the model in (21) implies that the instantaneous correlation between two stocks is stochastic:

$$\rho_{ij,t} = corr\left(\frac{dS_{i,t}}{S_{i,t}}, \frac{dS_{j,t}}{S_{j,t}}\right) = \frac{\beta_i\beta_j v_t}{\sqrt{\beta_i^2 v_t + \gamma_i^2}\sqrt{\beta_j^2 v_t + \gamma_j^2}}.$$

The correlation increases when the variance v_t increases and vice versa. This means that bets on the realized correlation are closely related to bets on the market variance, the fact that has been recognized by many traders of recently developed correlation swaps. Moreover, the sign and the magnitude of the correlation risk premium are linked to the sign and magnitude of the variance risk premium.

6.4 Performance of Hedge Funds

Many financial institutions might have a negative exposure to the variance risk, which is reflected in the fact that they generally do better in quite markets than in volatile ones, even after accounting for the market exposure.

In a companion paper, Bondarenko (2007) investigates how performance of hedge funds is related to the variance return. The study presents the first direct evidence that most hedge funds exhibit negative exposure to the variance return. For example, among Fund Research (HFR) categories, the largest “sellers” of the variance risk include Distressed Securities, Emerging Markets, Equity Non-Hedge, Event-Driven, Fixed Income: High Yield, Macro, Composite, and Fund of Funds Composite. The overall hedge fund industry earns about 7% annually by shorting the variance risk.[32] The variance return remains an important factor even in the presence of option returns as in Agarwal and Naik (2004).

These findings have important implications for hedge fund performance evaluation. Previous studies often conclude that hedge funds, as a group, deliver superior risk-adjusted returns (see, for example, Goetzmann, Ibbotson, and Brown (1999)). That is, linear regressions of

[32]Hedge funds can attain the negative exposure to the variance risk both *directly* (via trades in variance swaps and other derivatives) and *indirectly* (for example, by pursuing relative-value and event-driven arbitrage strategies, which often perform poorly during volatile periods.

fund returns on various market factors produce positive and statistically significant alpha coefficients. However, after correcting for the variance risk exposure, the performance of most categories becomes less impressive, with positive alphas often becoming negative or statistically insignificant. As a group, hedge funds no longer seem to "add value".

To be fair, however, we note that hedge funds could still be very useful to individual investors. As follows from Subsection 4.4, a typical CRRA investor would be willing to pay considerable amounts for being able to sell the variance contract. In the past, investors did not have a cheap and practical way to trade the variance contract or related securities. Under those circumstances, hedge funds are valuable because they offer investors an indirect way to attain negative exposure to the variance risk. Even though hedge funds charge very high fees, investing through them can greatly improve investors' utility. The situation, however, could change fundamentally in the future: if trading in the CBOE realized variance futures becomes more active, there would be an easy and direct way to trade the variance.

The hedge fund industry keeps growing at a very fast pace and its role in financial markets keeps expanding. This raises an interesting question: If hedge funds continue to sell the variance risk, will they eventually drive its risk premium down?

6.5 Cross-Section of Stock Returns

In a recent paper, Ang, Hodrick, Xing, and Zhang (2006) (hereafter AHXZ) investigate whether the systematic volatility is priced in the cross-section of stock returns. They find that stocks with high sensitivities to daily changes in VIX have lower average returns. This implies that aggregate volatility is priced in the cross-section of stocks and that it has a negative risk premium.

AHXZ infer the risk premium indirectly, by constructing the factor mimicking portfolio $FVIX$. Over the period from 01/1986 to 12/2000, they estimate the volatility risk premium as -0.08% per month. Although the estimate is statistically significant, economically, it appears to be rather small.[33] It would be interesting to compliment AHXZ's analysis by using the variance return, which has the important advantage of being an observable and tradeable factor. In contrast, the true factor mimicking portfolio $FVIX$ is unobservable and the construction of its proxy depends critically on various elements of the empirical design: the universe of stocks, the choice of the base assets, the choice of the correct factor model, and others.

7 Conclusion

This paper develops a new approach for valuing payoffs tied to the variance. The paper deviates from the existing literature by focusing on the empirically relevant realized variance, as opposed to the unobservable integrated variance. We argue that this distinction is critical for hedging actual variance instruments, such as OTC variance swaps and the CBOE variance futures. In realistic situations, the discretely-sampled realized variance differs considerably from the continuously-sampled integrated variance and the existing approach produces unacceptably large hedging errors.

Our approach is completely model-free. It allows for jumps in the underlying price process and applies to any sampling partition. The approach is useful for several reasons. It identifies

[33] The estimate in AHXZ is not directly comparable to our estimate of the variance risk premium, because the factor mimicking portfolio $FVIX$ and the variance contract can differ considerably in terms of leverage.

the exact replication strategy for the realized variance and shows that hedging errors are very sensitive to even slightest deviations from the optimal timing. The approach allows one to recreate historical prices of the variance contract and to study its risk characteristics. In particular, CBOE can rely on this information to determine the proper margin requirement for its futures, while traders and regulators can use it to assess the extent of the counterparty risk present in OTC variance swaps. We also show how to extend our approach to value the generalized variance contracts, for which the variance is accumulated according to a general weight function.

In the empirical application, we study the economic properties of the variance contract on S&P 500 over the long period from 01/1990 to 12/2006. We find that the variance risk is priced, its risk-premium is negative and very large in economic terms. The variance risk premium cannot be explained by known risk factors and option returns. We argue that ability to trade the variance can considerably enhance the utility of a typical investor. Trading strategies which involve selling the variance contract would have resulted in very high profits over the studied period. The variance contract appears to be even more "expensive" than the already puzzling index put options. Overall, the empirical findings present a difficult challenge for the theoretical literature to develop an equilibrium model which could justify the variance risk premium observed empirically.

One important implication of this paper is that the variance return should be considered a new risk factor for explaining a wide range of assets, both traditional and alternative. In particular, we conjecture that this factor might come out significant for small stocks, corporate bonds, emerging market stocks, bonds, and currencies, as well as credit default swaps, correlation swaps, and others innovative products.

There is some evidence that the variance risk is related to the proxies of the credit, liquidity, and correlation risks. This is reflected in the fact that unexpected increases in the variance are often associated with wider spreads, lower liquidity, and higher correlations between different assets. Empirically, it might be difficult to separate the interrelated risks, but the important advantage of our approach is that it allows one to measure the variance risk directly, while the credit, liquidity, and correlation risks might be harder to quantify precisely.

The variance return goes a long way toward explaining the returns of hedge funds, but it might also be important for other institutions, such as market makers and banks. Portfolio managers should pay attention to the variance risk for two reasons: (1) assets expected returns could be directly affected by the exposure to the variance risk, (2) correlations across different assets increase when the variance increases, reducing the effect of diversification in volatile markets.

Appendix

A Proof of Proposition 1

We first establish the fact that

$$OP_T = 2\left(\frac{F_T - F_0}{F_0} - \ln\frac{F_T}{F_0}\right). \tag{23}$$

The proof of (23) follows Carr and Madan (1998) and Bakshi and Madan (2000). Let $g(F)$ be a general function, whose second derivative $g''(F)$ is continuous almost everywhere. Then the following representation holds

$$g(F_T) = g(x) + g'(x)(F_T - x) + \int_0^x g''(K)(K - F_T)^+ dK + \int_x^\infty g''(K)(F_T - K)^+ dK. \tag{24}$$

for any $x \geq 0$.[34] In particular, when $x = F_0$, the above equation becomes

$$\begin{aligned} g(F_T) &= g(F_0) + g'(F_0)(F_T - F_0) + \int_0^{F_0} g''(K)(K - F_T)^+ dK + \int_{F_0}^\infty g''(K)(F_T - K)^+ dK \\ &= g(F_0) + g'(F_0)(F_T - F_0) + \int_0^\infty g''(K) M_T(K) dK. \end{aligned} \tag{25}$$

Applying (25) to the function $g(F) = \ln F$, one obtains

$$\ln F_T = \ln F_0 + \frac{1}{F_0}(F_T - F_0) - \int_0^\infty \frac{M_T(K)}{K^2} dK,$$

which proves (23).

Next, equation (2) and Ito's lemma imply that

$$d(\ln F_t) = \frac{dF_t}{F_t} - \frac{1}{2} v_t dt.$$

Therefore,

$$\begin{aligned} IV_T = \int_0^T v_t dt = -2\ln\frac{F_T}{F_0} + 2\int_0^T \frac{dF_t}{F_t} &= 2\left(\frac{F_T - F_0}{F_0} - \ln\frac{F_T}{F_0}\right) + 2\int_0^T \left(\frac{1}{F_t} - \frac{1}{F_0}\right) dF_t \\ &= OP_T + 2\int_0^T \left(\frac{1}{F_t} - \frac{1}{F_0}\right) dF_t. \end{aligned} \tag{26}$$

Equation (26) identifies the replication strategy for the variance contract. It consists of

(i) the static position with the payoff OP_T, and

(ii) the dynamic strategy, which at each moment holds $2\left(\frac{1}{F_t} - \frac{1}{F_0}\right)$ shares of the futures contract.

Since the process F_t is a martingale under the risk-neutral measure, the risk-neutral expectation of the dynamic strategy (ii) is zero. This means that $E_0^*[IV_T] = E_0^*[OP_T] = U_0$.

□

[34] One way to derive (24) is to first write the payoff $g(F_T)$ as

$$g(F_T) = \int_0^x g(K)\delta(K - F_T)dK + \int_x^\infty g(K)\delta(F_T - K)dK,$$

where $\delta(\cdot)$ is the Dirac delta function, and then integrate each integral by parts two times.

B Construction of Dataset

To construct our final dataset, we follow several steps:

1. For both options and futures we use settlement prices. Settlement prices (as opposed to closing prices) do not suffer from nonsynchronous/stale trading of options and the bid-ask spreads. CME calculates settlement prices simultaneously for all options, based on their last bid and ask prices. Since these prices are used to determine daily margin requirements, they are carefully scrutinized by the exchange and closely watched by traders. As a result, settlement prices are less likely to suffer from recording errors and they rarely violate basic no-arbitrage restrictions. In contrast, closing prices are generally less reliable and less complete.

2. In the dataset, we match all puts and calls by trading date t, maturity T, and strike. For each pair (t, T), we drop very low (high) strikes for which put (call) price is less than 0.1. To convert spot prices to forward prices, we approximate the risk-free rate r_f over $[t, T]$ by the rate of Treasury bills.

3. Because the CME options are American type, their prices $P_t^A(K)$ and $C_t^A(K)$ could be slightly higher than prices of the corresponding European options $P_t(K)$ and $C_t(K)$. The difference, however, is very small for short maturities that we focus on. This is particularly true for OTM an ATM options.[35]

To infer prices of European options $P_t(K)$ and $C_t(K)$, we proceed as follows. First, we discard all ITM options. That is, we use put prices for $K/F_t \leq 1.00$ and call prices for $K/F_t \geq 1.00$. Prices of OTM and ATM options are both more reliable and less affected by the early exercise feature. Second, we correct American option prices $P_t^A(K)$ and $C_t^A(K)$ for the value of the early exercise feature by using Barone-Adesi and Whaley (1987) approximation.[36] Third, we compute prices of ITM options through the put-call parity relationship

$$P_t(K) + F_t = C_t(K) + K.$$

4. We check option prices for violations of the no-arbitrage restrictions. To preclude arbitrage opportunities, call and put prices must be monotonic and convex functions of the strike. In particular, the call pricing function $C_t(K)$ must satisfy

$$(a) \quad C_t(K) \geq (F_t - K)^+, \qquad (b) \quad -1 \leq C_t'(K) \leq 0, \qquad (c) \quad C_t''(K) \geq 0.$$

The corresponding conditions for the put pricing function $P_t(K)$ can be obtained from put-call parity. When restrictions (a)-(c) are violated, we enforce them by running the so-called *Constrained Convex Regression* (CCR). This procedure has been implemented in Bondarenko (2000). Intuitively, CCR searches for the smallest (in the sense of least squares) perturbation of option prices that restores the no-arbitrage restrictions. The procedure also allows one to identify possible recording errors or typos.

5. For each pair (t, T), we estimate RND using the *Positive Convolution Approximation* (PCA) procedure of Bondarenko (2000, 2003a). Armed with RND, we obtain the put and call pricing functions. The market price of the variance contract is then computed using the relationships in (4) and (19). The market price of PPWVC is computed similarly, using Proposition 3.

C Monte-Carlo Experiment

In the Monte-Carlo experiment, we assume that the actual RND $h_0(F_T)$ is given by a flexible parametric specification. For that, we use the stochastic volatility model with contemporaneous jumps in price and volatility (SVCJ), which is shown in Broadie, Chernov, and Johannes (2005) to provide an excellent

[35]As shown in Whaley (1986), the early exercise premium increases with the level of the risk-free rate, volatility, time to maturity, and degree to which an option is in-the-money.

[36]It is important to point out that this correction is always substantially smaller than typical bid-ask spreads. In particular, the correction generally does not exceed 0.2% of an option price.

fit to the S&P 500 futures options. Given the model parameters estimated by Broadie, Chernov, and Johannes, we set the instantaneous variance v_t to fit a typical cross-section of the S&P 500 futures options in our sample. In particular, we assume that

- the initial date is June 30, 1997, when F_0=890.25;
- the maturity date for the variance contract is July 31, 1997, so that $T = 31$ days;
- the closest option maturity is July 18, 1997, so that $T_1 = 18$ days; for that maturity there are 50 strikes available ranging from 725 to 970;
- the next option maturity is August 15, 1997, so that $T_2 = 46$ days; for that maturity there are 48 strikes available ranging from 680 to 970;

With the parametric model, we can compute the theoretical price of the variance contract $U_0 = U(0, T)$. For each option maturity T_j $(j = 1, 2)$, we also compute theoretical call and put prices $\{C_{ij}, P_{ij}\}$ for those strikes K_i that are available for given maturity T_j.

To create simulated option prices, we add to the theoretical call and put prices random noise, which models observational errors that arise from non-synchronicity, bid-ask spread, and other market imperfections. Specifically, for a given theoretical option price q, we introduce measurement errors $\varepsilon(q)$, which are independent and for which $E[\varepsilon(q)] = 0$. The details for specification $\varepsilon(q)$ are explained below.

We generate 500 simulated cross-sections of options. For each cross-section, an estimate $\hat{U}_0$ is obtained in the same way as with real data. Given 500 estimates $\hat{U}_0$, the RMSE measure is computed as in (20).

Error specification

To model the measurement errors $\varepsilon(q)$, we have used several approaches and have found that the relative performance of the estimator $\hat{U}_0$ is qualitatively very similar. Since CME calculates option settlement prices by using the last bid and ask prices, we assume that the introduced measurement error $\varepsilon(q)$ is uniformly distributed on $[-0.5s(q), 0.5s(q)]$, where $s(q)$ is the "effective spread" corresponding to the price q. Simple approaches (not reported) would be to assume that (i) the level of noise is the same across options in *absolute* terms, so that $s(q) = c$ for some constant c, or (ii) the level of noise is the same in *relative* terms, so that $s(q) = cq$ for some constant c. However, both approaches are not completely satisfactory.

The alternative approach is more realistic, although it is also a bit more complex. We observe that options exchanges typically set upper limits for bid-ask spreads. For example, the CBOE rules limit the maximum bid-ask spread for the S&P 500 options as:

	$q < 2$	$2 \le q < 5$	$5 \le q < 10$	$10 \le q < 20$	$20 \le q$
Maximum spread	1/4	3/8	1/2	3/4	1

To approximate these rules, we construct function $M(q)$, which represents the maximum spread for a given option price q. Specifically, let

$$M(0) = \frac{1}{8}, \quad M(2) = \frac{1}{4}, \quad M(5) = \frac{3}{8}, \quad M(10) = \frac{1}{2}, \quad M(20) = \frac{3}{4}, \quad M(q) = 1, \quad q \ge 50,$$

and $M(q)$ is linearly interpolated for all other $q \in [0, 50]$. In particular, $M(q)$ is about 19% (4%) of the option price when $q = \$1$ ($q = \$20$). Then we assume that $s(q)$ is proportional to the maximum spread, or $s(q) = cM(q)$. By varying the constant c, the level of noise can be increased or decreased across all options. We report the results for three cases: 1) No noise, $c = 0$, 2) Moderate noise, $c = 0.5$ (the effective spreads are half of the exchange allowed maximum), and 3) High noise, $c = 1$ (the effective spreads are equal to the allowed maximum). The advantages of the proposed specification for $\varepsilon(q)$ are that (a) noise is smaller in the *absolute* terms for OTM options, (b) noise is larger in the *relative* terms for OTM options, and (c) simulated option prices are always nonnegative.

References

Aït-Sahalia, Y., 2002, "Telling from discrete data whether the underlying continuous-time model is a diffusion," *Journal of Finance* 96, 2075-2112.

Aït-Sahalia, Y., and P. Mykland, 2003, "The Effects of Random and Discrete Sampling When Estimating Continuous-Time Diffusions," *Econometrica* 71, 483-549.

Agarwal, V., and N. Naik, 2004, "Risks and Portfolio Decisions Involving Hedge Funds," *Review of Financial Studies* 17, 63-98.

Andersen, T., L. Benzoni, and J. Lund, 2002, "An Empirical Investigation of Continuous-Time Equity Returns Models," *Journal of Finance* 57, 1239-1284.

Andersen, T., and T. Bollerslev, 1998, "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review* 39, 885-905.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys, 2003, "Modeling and forecasting realized volatility," *Econometrica* 71, 579-625.

Andersen, T., O. Bondarenko, 2007, "Construction and Interpretation of Model-Free Implied Volatility," *Volatility as an Asset Class*, London: Risk Books, Ed. by I. Nelken, 141-181.

Ang, A., B. Hodrick, Y. Xing, and X. Zhang, 2006, "The Cross-Section of Volatility and Expected Returns," *Journal of Finance*, 61, 259-299.

Bakshi, G., C. Cao, and Z. Chen, 1997, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance* 52, 2003-2049.

Bakshi, G., C. Cao, and Z. Chen, 2000, "Do Call Prices and the Underlying Stock Always Move in the Same Direction?" *Review of Financial Studies* 13, 549-584.

Bakshi, G., and N. Kapadia, 2003a, "Delta-Hedged Gains and the Negative Market Volatility Premium," *Review of Financial Studies* 16, 527-566.

Bakshi, G., and N. Kapadia, 2003b, "Volatility Risk Premium Embedded in Individual Equity Options: Some New Insights," working paper.

Bakshi, G., and D. Madan, 2000, "Spanning and Derivative-Security Valuation," *Journal of Financial Economics* 55, 205-238.

Bandi, F., and Russel, J., 2003, "Microstructure Noise, Realized Volatility, and Optimal Sampling," working paper, University of Chicago.

Banz, R., and M. Miller, 1978, "Prices for State-Contingent Claims: Some Estimates and Applications," *Journal of Business* 51, 653-672.

Barndorff-Nielsen, O., and N. Shephard, 2002, "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B* 64, 252-280.

Barone-Adesi, G., and Whaley, R., 1987, "Efficient Analytical Approximation of American Option Values," *Journal of Finance* 41, 301-320.

Bates, D., 2000, "Post-'87 Crash Fears in the S&P 500 Futures Option Market," *Journal of Econometrics* 94, 181-238.

Benzoni, L., 2002, "Pricing Options under Stochastic Volatility: An Empirical Investigation," working paper, University of Minnesota.

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81, 637-654.

Bollerslev, T., M. Gibson, and H. Zhou, 2007, "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities," working paper.

Bondarenko, O., 2000, "Recovering Risk-Neutral Densities: A New Nonparametric Approach," UIC working paper.

Bondarenko, O., 2003a, "Estimation of Risk-Neutral Densities Using Positive Convolution Approximation," *Journal of Econometrics* 116, 85-112.

Bondarenko, O., 2003b, "Why are Put Options So Expensive?" UIC working paper.

Bondarenko, O., 2007, "Do Hedge Funds Sell Short Variance?" UIC working paper.

Branger, N., and C. Schlag, 2005, "An Economic Motivation for Variance Contracts," working paper.

Breeden, D., and R. Litzenberger, 1978, "Prices of State Contingent Claims Implicit in Options Prices," *Journal of Business* 51, 621-652.

Brenner, M., E. Ou, and J. Zhang, 2006, "Hedging Volatility Risk" *Journal of Banking and Finance* 30, 811-821.

Britten-Jones, M. and A. Neuberger, 2000, "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance* 55, 839-866.

Broadie, M., M. Chernov, and M. Johannes, 2005, "Model Specification and Risk Premia: Evidence from Futures Options," *Journal of Finance* forthcoming.

Buraschi, A., and J. Jackwerth, 2001, "The Price of a Smile: Hedging and Spanning in Option Markets," *Review of Financial Studies* 14, 495-527.

Carr, P., and R. Lee, 2005, "Robust Replication of Volatility Derivatives," working paper.

Carr, P., and K. Lewis, 2007, "Corridor Variance Swaps," *Volatility as an Asset Class*, London: Risk Books, Ed. by I. Nelken, 287-305.

Carr, P., and D. Madan, 1998, "Towards a Theory of Volatility Trading," *Volatility*, Risk Publications, R. Jarrow, Ed., 417-427.

Carr, P., and L. Wu, 2003, "What Type of Process Underlies Options? A Simple Robust Test," *Journal of Finance* 58, 2581-2610.

Carr, P., and L. Wu, 2007, "Variance Risk Premia," *Review of Financial Studies*, forthcoming.

Chen, L., P. Collin-Dufresne, R. Goldstein, 2006, "On the Relation Between the Credit Spread Puzzle and the Equity Premium Puzzle," working paper.

Chernov, M., and E. Ghysels, 2000, "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics* 56, 407-458.

Coval, J., and T. Shumway, 2001, "Expected Option Returns," *Journal of Finance* 56, 983-1009.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou, 1999, "More Than You Ever Wanted to Know About Volatility Swaps," working paper, Goldman Sachs, http://www.gs.com/qs/doc/volswaps.pdf.

Driessen, J., P. Maenhout, 2004, "A Portfolio Perspective on Option Pricing Anomalies," working paper.

Driessen, J., P. Maenhout, and G. Vilkov, 2006, "Option-Implied Correlations and the Price of Correlation Risk," working paper.

Dumas, B., J. Fleming, and R. Whaley, 1998, "Implied Volatility Functions: Empirical Tests," *Journal of Finance* 53, 2059-2106.

Dupire, B., 1993, "Model Art," *Risk* (September) 118-120.

Egloff, D., M. Leippold, and L. Wu, 2007, "Variance Risk Dynamics, Variance Risk Premia, and Optimal Variance Swap Investments," working paper.

Eraker, B., M. Johannes, and N. Polson, 2003, "The Impact of Jumps on Volatility and Returns," *Journal of Finance* 58, 1269-1300.

Fama, E., and K. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics* 33, 3-56.

Goetzmann, W., R. Ibbotson, and S. Brown, 1999, "Offshore Hedge Funds: Survival and Performance 1989-1995," *Journal of Business* 72, 91-117.

Grunbichler, A., and F. Longstaff, 1996, "Valuing Futures and Options on Volatility," *Journal of Banking and Finance* 20, 985-1001.

Guo, D.,1998, "The Risk Premium of Volatility Implicit in Currency Options," *Journal of Business and Economic Statistics* 16, 498-507.

Hansen, P., and A. Lunde, 2005, "A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data," *Journal of Financial Econometrics* 3, 525-554.

Heston, S., 1993, "A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies* 6, 327-343.

Hull, J., and A. White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance* 42,

Jegadeesh, N., and S. Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications For Stock Market Efficiency," *Journal of Finance* 48, 65-92.

Jiang, G., and Y. Tian, 2005, "The Model-Free Implied Volatility and Its Information Content," *Review of Financial Studies* 18, 1305-1342.

Jiang, G., and Y. Tian, 2007, "Extracting Model-Free Volatility from Option Prices: An Examination of the VIX Index," *Journal of Derivatives*, Spring, 1-26.

Krishnan, C., R. Petkova, and P. Ritchken, 2006, "The Price of Correlation Risk," working paper.

Modigliani, F., and L. Modigliani, 1997, "Risk-Adjusted Performance," *Journal of Portfolio Management* 23 (2), 45-54.

Neuberger, A., 1994, "The Log Contract: A New Instrument to Hedge Volatility," *Journal of Portfolio Management*, Winter, 74-80.

Pastor, L., and R. Stambaugh, 2002, "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy* 111, 642-685.

Ross, S., 1976, "Options and Efficiency," *Quarterly Journal of Economics* 90, 75-89.

Todorov, V., 2007, "Variance Risk Premium Dynamics", working paper.

Wu, L., 2005, "Variance Dynamics: Joint Evidence from Options and High-Frequency Returns," working paper.

Zhang, L., Mykland, P., and Y. Aït-Sahalia, 2005, "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High Frequency Data," *Journal of the American Statistical Association*, forthcoming.

Table 1: **Hedging Errors for Different $\Delta t'$**

Panel A: BS Model

	1 min	10 min	1 hour	1/2 day	1 day	2 days	5 days
$RV_T^{(1)}$	0.387	0.383	0.353	0.255	0.0024	0.285	0.578
$RV_T^{(2)}$	0.387	0.383	0.353	0.256	0.0047	0.286	0.584
$RV_T^{(3)}$	0.387	0.383	0.353	0.255	0	0.285	0.579

Panel B: SV Model

	1 min	10 min	1 hour	1/2 day	1 day	2 days	5 days
$RV_T^{(1)}$	0.392	0.388	0.361	0.261	0.0025	0.291	0.594
$RV_T^{(2)}$	0.392	0.388	0.361	0.261	0.0050	0.293	0.602
$RV_T^{(3)}$	0.392	0.388	0.361	0.261	0	0.292	0.596

Panel C: S&P 500

	1 min	10 min	1 hour	1/2 day	1 day	2 days	5 days
$RV_T^{(1)}$	n/a	0.329	0.285	0.262	0.0030	0.271	0.468
$RV_T^{(2)}$	n/a	0.329	0.285	0.261	0.0060	0.274	0.469
$RV_T^{(3)}$	n/a	0.329	0.285	0.262	0	0.272	0.468

Notes: The target payoff RV_T has maturity $T = 1$ month and is sampled over equidistant partition with $\Delta t = 1$ day. The relative hedging errors are reported for the three definitions of the realized variance and for different $\Delta t'$, which measures how frequently the underlying is actually rebalanced. For the Black-Scholes and Heston models, parameters are set as in (12)-(13) and the hedging errors are computed using Monte-Carlo simulations with 5000 repetitions, where we assume that one year consists of 252 trading days and one day is equal to 7 trading hours. For the S&P 500 futures, the hedging errors are computed using 5-minute returns over the period from 01/1990 to 12/2006.

Table 2: **Hedging Errors for Different δt**

Panel A: BS Model

	0	1 min	5 min	15 min	1 hour
$RV_T^{(1)}$	0.0024	0.031	0.071	0.121	0.236
$RV_T^{(2)}$	0.0047	0.032	0.071	0.122	0.236
$RV_T^{(3)}$	0	0.031	0.071	0.122	0.236

Panel B: SV Model

	0	1 min	5 min	15 min	1 hour
$RV_T^{(1)}$	0.0025	0.032	0.071	0.124	0.241
$RV_T^{(2)}$	0.0050	0.032	0.072	0.124	0.241
$RV_T^{(3)}$	0	0.032	0.071	0.124	0.241

Panel C: S&P 500

	0	1 min	5 min	15 min	1 hour
$RV_T^{(1)}$	0.0030	n/a	0.074	0.113	0.251
$RV_T^{(2)}$	0.0060	n/a	0.075	0.114	0.252
$RV_T^{(3)}$	0	n/a	0.074	0.113	0.251

Notes: The target payoff RV_T has maturity $T = 1$ month and is sampled over equidistant partition with $\Delta t = 1$ day. The relative hedging errors are reported for the three definitions of the realized variance and for different δt, which measures deviations from the optimal times. For the Black-Scholes and Heston models, parameters are set as in (12)-(13) and the hedging errors are computed using Monte-Carlo simulations with 5000 repetitions, where we assume that one year consists of 252 trading days and one day is equal to 7 trading hours. For the S&P 500 futures, the hedging errors are computed using 5-minute returns over the period from 01/1990 to 12/2006.

Table 3: **Monthly Returns for Variance Contract, S&P 500, Market-Neutral Portfolio, and Mean-Variance Portfolio**

Panel A: Return Distribution

	Min.	1%	5%	10%	Med.	90%	95%	99%	Max.
VC	-83.08	-81.60	-74.35	-70.23	-40.08	13.97	66.19	162.12	174.48
SP500	-15.05	-10.52	-6.67	-4.53	0.90	5.63	6.57	8.96	10.75
MNP	-158.52	-117.42	-55.03	-20.83	39.48	65.10	75.78	87.11	95.02
MVP	-40.12	-32.79	-15.16	-4.24	9.85	16.41	17.63	20.34	20.46

Panel B: Risk Characteristics

	Mean	SD	Skew.	Kurt.	α	β	SR	TM	M^2
VC	-31.24	44.71	2.19	9.04	-28.52	-4.99	-0.70	5.71	-2.81
SP500	0.55	4.02	-0.51	3.92	0.00	1.00	0.14	0.00	0.55
MNP	28.52	39.96	-1.60	6.53	28.52	0.00	0.71	n/a	2.87
MVP	7.20	9.92	-1.92	7.79	6.95	0.46	0.73	15.07	2.92

Notes: Statistics are reported for excess returns of the variance contract (VC), the S&P 500 futures, the market-neutral portfolio (MNP), and the mean-variance portfolio (MVP). Returns are expressed in monthly percentage term. MNP is short the variance contract and hedged with the S&P 500 futures so that the sample correlation with the market return r_m is zero; its return $r_{mn} = -r_v - 4.99r_m$. MVP is the mean-variance optimal portfolio constructed using the variance contract and the S&P 500 futures; its return $r_{mv} = -0.24r_v - 0.76r_m$. Panel A shows minimum, median, maximum, and 1%, 5%, 10%, 90%, 95%, and 99% percentiles. Panel B shows mean, standard deviation (SD), skewness, kurtosis, α and β coefficients (with respect to S&P 500), Sharpe ratio (SR), Treynor's measure (TM), M-squared measure (M^2). The sample period is 01/1990–12/2006.

Table 4: **Risk Characteristics for Monthly Returns of S&P 500 Options**

Panel A: Puts

k	Mean	SD	Skew.	Kurt.	α	β	SR	TM	M^2	ρ_m	ρ_v
0.96	-64.12	127.04	4.47	25.44	-52.97	-20.43	-0.50	2.59	-2.03	-0.65	0.53
0.98	-47.53	130.54	2.99	12.43	-34.39	-24.07	-0.36	1.43	-1.46	-0.74	0.57
1.00	-31.64	119.68	2.02	6.95	-18.06	-24.89	-0.26	0.73	-1.06	-0.84	0.57
1.02	-21.45	97.30	1.44	4.94	-9.34	-22.19	-0.22	0.42	-0.89	-0.92	0.54
1.04	-14.12	76.17	1.04	4.16	-4.11	-18.35	-0.19	0.22	-0.75	-0.97	0.50

Panel B: Calls

k	Mean	SD	Skew.	Kurt.	α	β	SR	TM	M^2	ρ_m	ρ_v
0.96	1.90	66.71	0.16	2.45	-6.65	15.66	0.03	-0.42	0.11	0.94	-0.37
0.98	0.50	86.47	0.51	2.65	-9.96	19.17	0.01	-0.52	0.02	0.89	-0.31
1.00	-1.32	116.56	1.13	3.89	-13.96	23.16	-0.01	-0.60	-0.05	0.80	-0.22
1.02	-10.68	163.18	2.27	9.05	-25.27	26.74	-0.07	-0.94	-0.26	0.66	-0.14
1.04	-41.26	207.60	5.98	46.62	-55.03	25.22	-0.20	-2.18	-0.80	0.49	-0.04

Panel C: ATM Straddle

k	Mean	SD	Skew.	Kurt.	α	β	SR	TM	M^2	ρ_m	ρ_v
1.00	-16.48	59.90	0.94	4.08	-16.01	-0.86	-0.28	18.60	-1.11	-0.06	0.35

Notes: Statistics are reported for excess returns of 1-month options on the S&P 500 futures: puts and calls with different moneyness k and ATM straddles. Statistics include mean, standard deviation (SD), skewness, kurtosis, α and β coefficients (with respect to S&P 500), Sharpe ratio (SR), Treynor's measure (TM), M-squared measure (M^2), and correlations with the market and variance contract. Returns are expressed in monthly percentage term. The sample period is 01/1990–12/2006.

Table 5: **OLS Regressions for Variance Return**

	Const	SP500	SMB	HML	UMD	ATMP	OTMP	ATMC	OTMC	R^2
Reg 1	-28.52	-4.99								0.20
	(-10.08)	(-7.14)								
Reg 2	-26.08	-5.84	-4.71	-2.73						0.31
	(-9.61)	(-8.01)	(-5.77)	(-2.79)						
Reg 3	-24.98	-6.19	-4.62	-2.94	-1.01					0.31
	(-9.02)	(-8.23)	(-5.68)	(-3.00)	(-1.76)					
Reg 4	-21.15	0.99				0.17	0.08	-0.02	0.01	0.33
	(-6.98)	(0.45)				(3.07)	(2.07)	(-0.39)	(0.55)	
Reg 5	-18.70	-1.81	-3.74	-2.51	-0.99	0.11	0.08	0.00	0.01	0.40
	(-6.19)	(-0.80)	(-4.76)	(-2.69)	(-1.74)	(1.96)	(2.38)	(-0.04)	(0.41)	

Notes: This table reports results of OLS regressions of the monthly variance return. The independent variables include the return on S&P 500 futures, the Fama-French size and value factors (SMB and HML), the momentum factor UMD, returns of at-the-money ($k = 1.00$) and out-of-the-money ($k = 0.96$) put options, and returns of at-the-money ($k = 1.00$) and out-of-the-money ($k = 1.04$) call options. All returns are in excess of risk-free rate and expressed in percentage term; t-statistics are reported in parentheses. The adjusted R^2 is provided in the last column. The sample period is 01/1990–12/2006.

Table 6: **Portfolio Weights and Certainty Equivalent Rate for CRRA Preferences**

Panel A: Market Portfolio and Variance Contract

	γ=1	γ=2	γ=3	γ=5	γ=10	γ=20	γ=50
w_m	3.089	1.604	1.080	0.652	0.328	0.164	0.066
w_m	-2.967	-2.191	-1.594	-1.004	-0.515	-0.261	-0.107
w_v	-0.601	-0.447	-0.333	-0.215	-0.113	-0.058	-0.024
CER	5.075	7.098	5.214	3.351	1.760	0.906	0.375

Panel B: Market Portfolio, ATM Put, and Variance Contract

	γ=1	γ=2	γ=3	γ=5	γ=10	γ=20	γ=50
w_m	-5.777	-3.085	-2.096	-1.276	-0.647	-0.329	-0.137
w_p	-0.290	-0.162	-0.112	-0.069	-0.035	-0.018	-0.007
w_m	-6.077	-3.253	-2.150	-1.269	-0.626	-0.314	-0.131
w_p	-0.105	-0.042	-0.023	-0.011	-0.005	-0.002	-0.001
w_v	-0.624	-0.441	-0.326	-0.211	-0.111	-0.057	-0.023
CER	7.643	5.905	4.316	2.770	1.453	0.746	0.308

Notes: This table reports optimal portfolio weights for the investor with one month horizon and CRRA preferences (γ is the risk aversion coefficient). The investor can trade the market portfolio and the variance contract (Panel A), or the market portfolio, the ATM put, and the variance contract (Panel B). The top portions of both panels show portfolio weights when trading in the variance contract is not allowed. The certainty equivalent rate (CER) is the percent of wealth that investor is willing to pay each month to be able to trade the variance contract. The sample period is 01/1990–12/2006.

Table 7: **Risk Characteristics for Monthly Returns of Power-Price-Weighted Variance Contracts**

a	Mean	SD	Skew.	Kurt.	α	β	SR	TM	M^2	ρ_m	ρ_v
-3	-36.27	49.31	2.78	12.75	-32.83	-6.32	-0.74	5.20	-2.96	-0.51	0.99
-2	-34.37	47.79	2.58	11.32	-31.14	-5.91	-0.72	5.27	-2.89	-0.50	0.99
-1	-32.70	46.24	2.38	10.10	-29.72	-5.47	-0.71	5.44	-2.84	-0.47	1.00
0	-31.24	44.71	2.19	9.04	-28.52	-4.99	-0.70	5.71	-2.81	-0.45	1.00
1	-30.05	43.23	2.01	8.14	-27.59	-4.50	-0.70	6.13	-2.79	-0.42	1.00
2	-28.85	41.90	1.83	7.32	-26.67	-4.00	-0.69	6.66	-2.77	-0.38	0.99
3	-27.89	40.66	1.68	6.63	-25.98	-3.49	-0.69	7.44	-2.76	-0.35	0.98
5	-26.33	38.62	1.41	5.58	-24.98	-2.47	-0.68	10.11	-2.74	-0.26	0.95

Notes: Statistics are reported for excess returns of 1-month power-price-weighted variance contracts with different values of power a. Statistics include mean, standard deviation (SD), skewness, kurtosis, α and β coefficients (with respect to S&P 500), Sharpe ratio (SR), Treynor's measure (TM), M-squared measure (M^2), and correlations with the market and variance contract. Returns are expressed in monthly percentage term. The sample period is 01/1990–12/2006.

Table 8: **Correlations**

	r_v	r_m	VIX	ΔVIX	SMB	HML	UMD	LIQ	ΔCS	$\Delta\bar{\rho}$
r_v	1.00	-0.45	0.41	0.62	-0.31	0.19	-0.01	-0.11	0.27	0.60
r_m	-0.45	1.00	-0.29	-0.67	0.02	-0.41	-0.23	-0.08	-0.03	-0.36
VIX	0.41	-0.29	1.00	0.28	-0.08	0.00	0.03	-0.03	0.22	0.14
ΔVIX	0.62	-0.67	0.28	1.00	-0.16	0.29	0.18	0.04	0.08	0.44

Notes: Time-series correlations are reported for the variance return r_v, the return on S&P 500 futures r_m, VIX, the change in VIX, the size factor, the value factor, the momentum factor, the value-weighted liquidity factor of Pastor and Stambaugh (2002), the change in the credit spread (where CS is the difference between the Moody's BAA yield and the 10-year Treasury constant maturity yield), the change in the average correlation $\bar{\rho}$. The sample period is 01/1990–12/2006 for all measures, except LIQ, for which the period is 01/1990–12/2004.

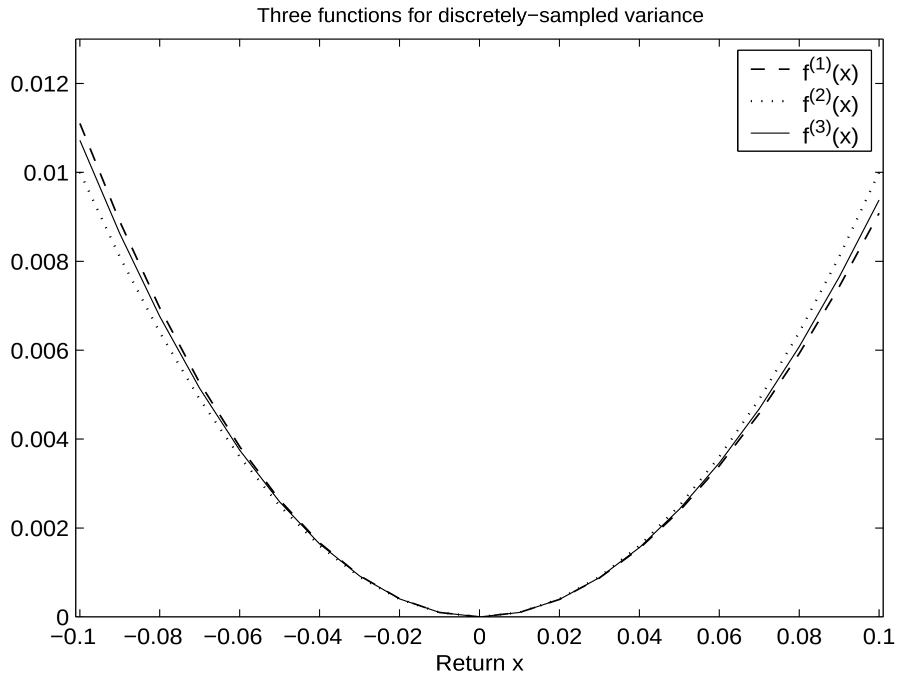

Figure 1: This figure shows behavior of functions $f^{(1)}(x) = [\ln(1+x)]^2$, $f^{(2)}(x) = x^2$, and $f^{(3)}(x) = 2\,(x - \ln(1+x))$ used in the definitions of the discretely-sampled variance $RV_T^{(1)}$, $RV_T^{(2)}$, and $RV_T^{(3)}$.

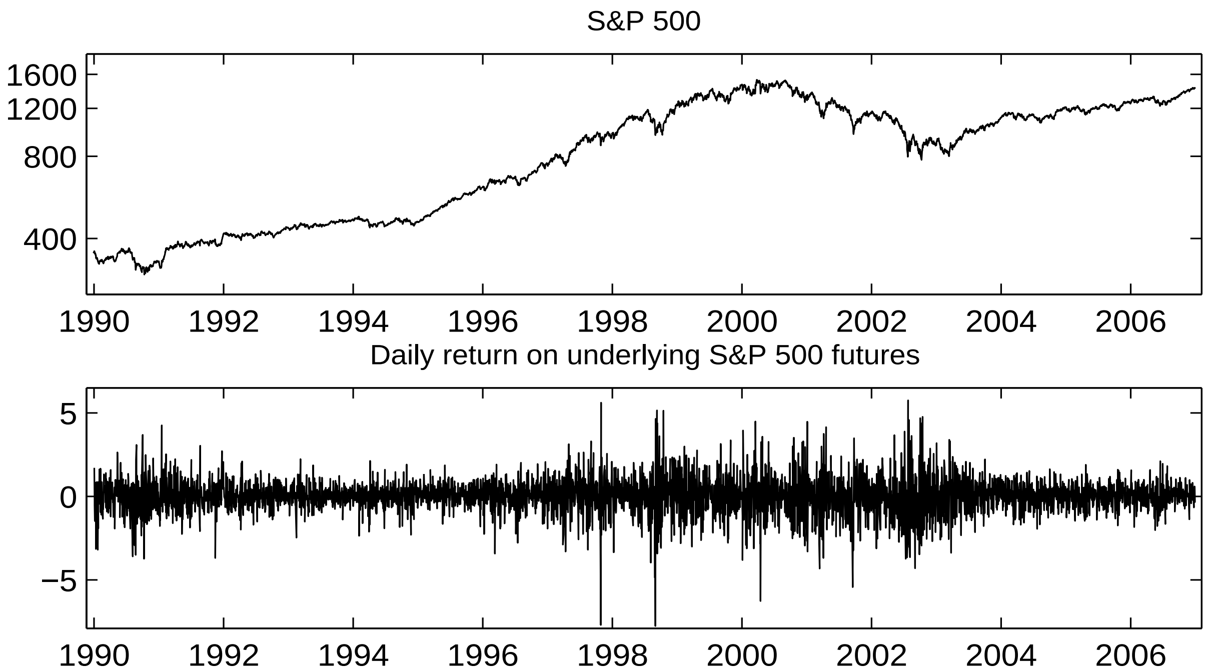

Figure 2: This figure plots the level of the S&P 500 Index and daily return of the S&P 500 futures.

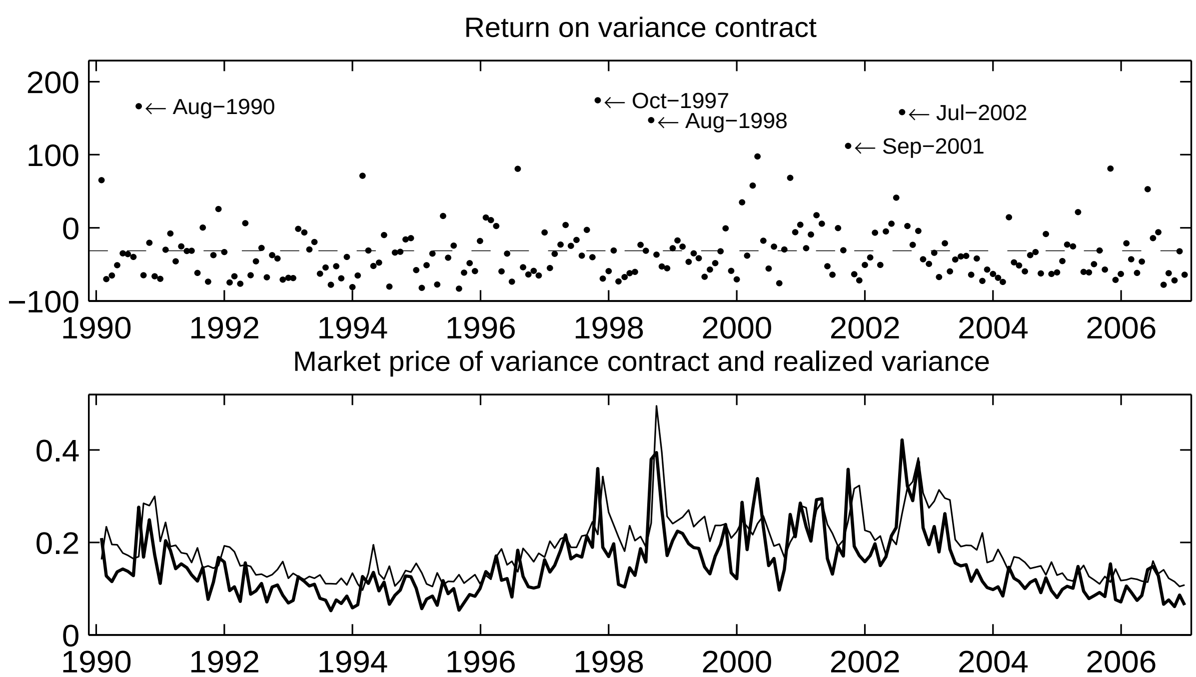

Figure 3: The top panel plots the monthly variance return from 01/1990 to 12/2006. Its sample mean is shown with the dashed line. The bottom panel plots the square root of the market price of the variance contract (the thick line) and the realized variance of S&P 500 futures (the thin line). The holding period is one month. Both series are annualized.

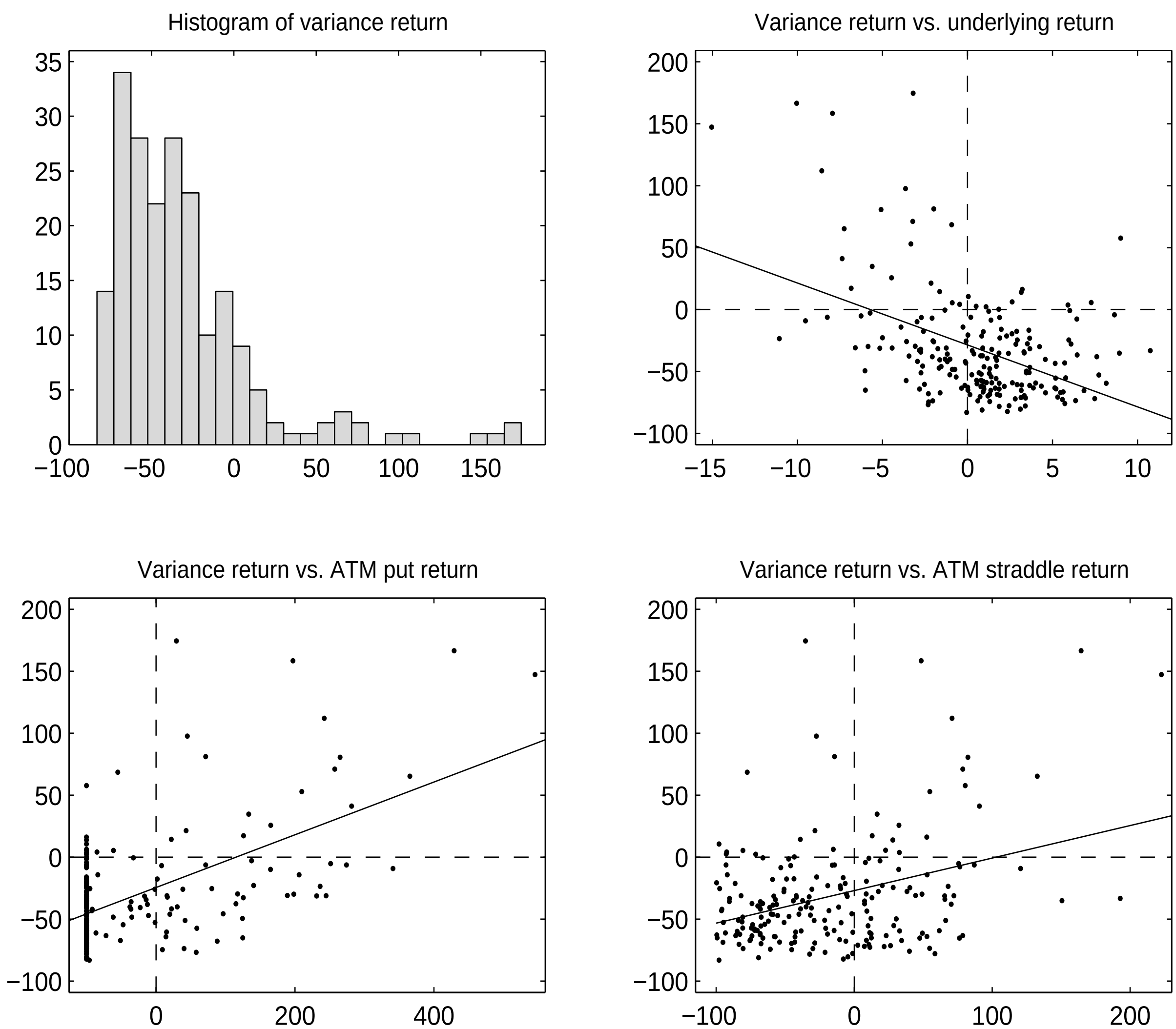

Figure 4: The top left panel shows the histogram of the variance return. The three other panels are scatter plots of the variance return versus the return on the underlying S&P 500 futures, the return on the ATM put, and the return on the ATM straddle. Also shown are corresponding OLS regression lines. All returns are monthly. The sample period is 01/1990–12/2006.

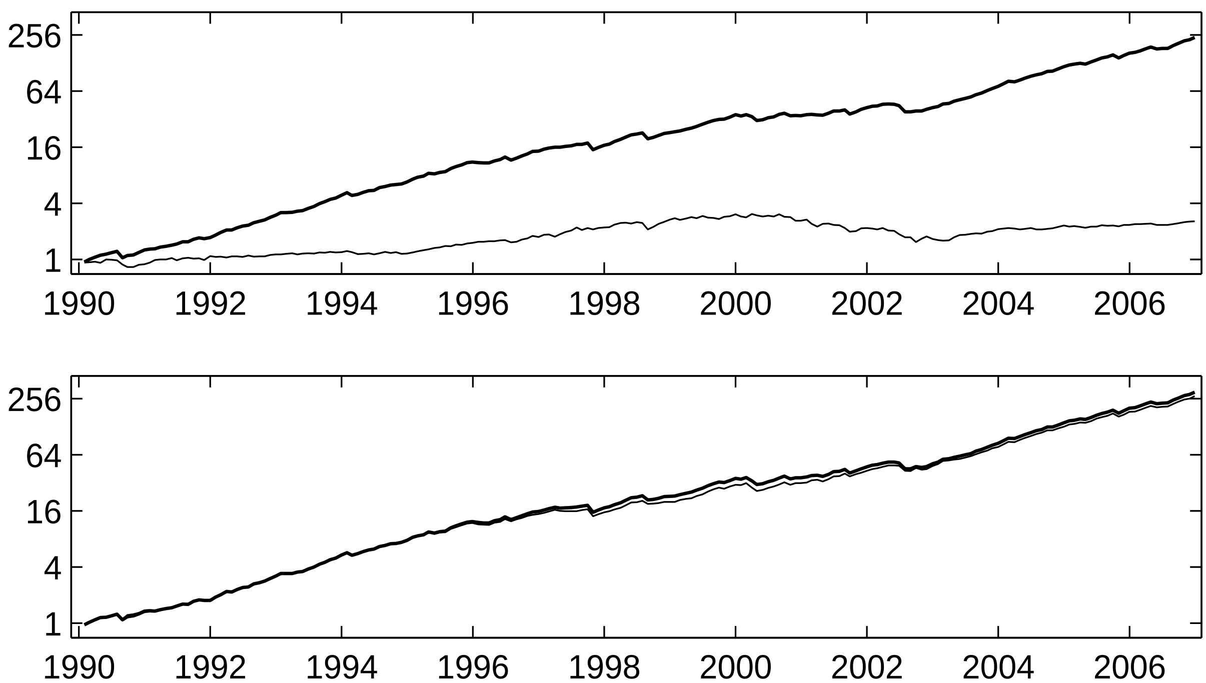

Figure 5: This figure shows the cumulative returns for four strategies that invest in 1) the S&P 500 futures, 2) the short variance contract, 3) the market-neutral portfolio MNP, and 4) the mean-variance portfolio MVP. Strategies 2)-4) are de-leveraged with the risk-free asset so that the standard deviations of their monthly returns over the sample period are the same as that for strategy 1). The top panel shows strategy 1) (the thin line) and strategy 2) (the thick line). The bottom panel shows strategy 3) (the thin line) and strategy 4) (the thick line).