What does the volatility risk premium say about liquidity provision and demand for hedging tail risk?[*†]

Jianqing Fan

jqfan@princeton.edu

Michael B. Imerman[‡]

mbi212@lehigh.edu

Wei Dai

weidai@princeton.edu

December 31, 2014

JEL Classification: C01, C58, G12

Keywords: volatility risk premium, integrated volatility, ultra-high-frequency data, microstructure noise, Fourier transform, tail risk, Big Data risk analytics

---

[*]Jianqing Fan is the Frederick L. Moore Professor of Finance, Bendheim Center for Finance, Princeton University, Princeton, NJ 08544; Michael B. Imerman is an Assistant Professor, Perella Department of Finance, Lehigh University, Bethlehem, PA 18015; and Wei Dai is a Ph.D. Student, Department of Operations Research & Financial Engineering, Princeton University, Princeton, NJ 08544.

[†]We would like to thank participants at the 20^{th} Annual Conference on Pacific Basin Finance, Economics, Accounting, and Management at Rutgers University and the Second Measuring Risk Conference at Princeton University as well as the Finance Department Seminar Series at Rutgers Business School. We are especially grateful for comments from Yacine Aït-Sahalia and Peter Carr. Also, feedback from the Derivatives Team at Morgan Stanley provided useful insight into the events of Fall 2008.

[‡]Contact Author Email: mbi212@lehigh.edu; Phone: (610) 758-6380.

What does the volatility risk premium say about liquidity provision and demand for hedging tail risk?

ABSTRACT

This paper provides a data-driven analysis of the volatility risk premium, using tools from high-frequency finance and Big Data analytics. We argue that the volatility risk premium, loosely defined as the difference between realized and implied volatilities, can best be understood when viewed as a systematically priced bias. We first use ultra-high-frequency transaction data on SPDRs and a novel approach for estimating integrated volatility on the frequency domain to compute realized volatility. From that we subtract the daily *VIX*, our measure of implied volatility, to construct a time series of the volatility risk premium. To identify the driving factors behind the volatility risk premium as a priced bias we decompose it into magnitude and direction. We find compelling evidence that the magnitude of the deviation of the realized volatility from implied volatility represents supply and demand imbalances in the market for hedging tail risk. It is difficult to conclusively accept the hypothesis that the direction or sign of the volatility risk premium reflects expectations about future levels of volatility. However, strong evidence supports the hypothesis that the sign of the volatility risk premium is indicative of gains or losses on a delta-hedged portfolio.

1 Introduction

Whether realized volatility is greater than or less than implied volatility is an empirical question, and one that has been studied over time (see Mixon, 2009). However, the literature contains conflicting evidence about implied volatility as an unbiased estimator of future realized volatility (e.g. Canina and Figlewski, 1993; Christensen and Prabhala, 1998). Theory suggests that implied volatility should be a biased estimate of future realized volatility since implied volatility includes the market price of risk; that is, implied volatility is the expected "actual" (or statistical) volatility plus a risk premium. In mathematical finance this is formalized in terms of a change of measure. The *volatility risk premium* is defined as the difference between the expected future volatility under the physical measure (ex-ante forecast of realized volatility) and the expected future volatility under the risk-neutral measure (implied volatility from option prices). Therefore, the existence of a non-zero volatility risk premium indicates that not only is implied volatility a biased estimator of future realized volatility but, furthermore, that the bias is systematically priced.

The volatility risk premium has been an active area of research in financial economics for some time now. Whereas existing studies typically begin with an asset pricing model or some framework of stochastic or time-varying volatility, we take a step back from the theoretical foundation of volatility in financial markets and perform a purely data-driven analysis of the volatility risk premium, leveraging the insights of *Big Data analytics*. That is, we start with a massive data-set of transaction level prices: our sample includes over half a billion trades in SPDRs, the ETF that tracks the S&P 500 index, from 2006 to 2011. We use this data to estimate the realized volatility of the market using a robust methodology with minimal parametric assumptions. Then, we compare this realized volatility to a measure of model-free implied volatility, daily over the same five year period. Our proxy for the model-free implied volatility is the VIX volatility index, which is commonly used in other studies of the volatility risk premium. Since our objective is to quantify the volatility risk premium in a model-free, nonparametric manner, we compare the computed ex-post realized volatility to the contemporaneous level of VIX.[1] As a result, we essentially return to the fundamental idea of the volatility risk premium as a *bias* – not a random bias, but a systematically priced bias. Through our data-driven analysis we seek to better understand the economic determinants of this bias.

Our study of the volatility risk premium represents a shift in the existing paradigm in financial research and risk analysis. In any quantitative field, there are two approaches to conducting research: model-based and data-driven. This dichotomy is perhaps more pronounced now than ever in the quantitative fields of financial economics – i.e. asset pricing, derivatives, and risk management. Traditionally, the researcher would construct a model based on theory and then use data to empirically verify or validate the model and justify the economic intuition it conveys. However, with the abundance of financial data being generated every day and the increasing popularity of "Big Data" along with data mining and machine learning techniques making their way into the financial engineer's toolbox, a new data-driven approach to research in these areas is gaining popularity. We view our study as an extension of this research philosophy to better understand how the market prices volatility.

As mentioned above, the model-based approach is found in most of the existing studies of the volatility risk premium; that is, they begin with an asset pricing model with time-varying volatility, then use data to test hypotheses related to the insights obtained from the model. From such studies we can identify several stylized facts about the volatility risk premium, upon which our data-driven analysis will build. First, traditional risk factors have poor explanatory power for the volatility risk premium (Carr and Wu, 2009). Second, the volatility risk premium is intimately related to the payoff and P&L on volatility swaps and hence reflects the market price that investors are willing to pay to hedge away uncertainty about future realized volatility (Demeterfi et al., 1999; Carr and Wu, 2009; Egloff et al., 2010; Wu, 2011; Aït-Sahalia et al., 2012). Third, the volatility risk premium is, on average or in expectation, *negative* – the volatility implied from option prices tends to be higher than the ex-

[1] As we will discuss in Section 2, most (but not all) existing studies use an ex-ante estimate of the expected future realized volatility estimated from an autoregressive model. However, this gives rise to statistical problems of overfitting and model misspecification, which we wish to avoid.

pected realized volatility – which serves as evidence that investors will *pay* to hedge volatility risk (Bakshi and Kapadia, 2003; Carr and Wu, 2009). Fourth, the volatility risk premium is highly correlated with market-wide risk aversion (Bollerslev et al., 2009, 2011; Bekaert et al., 2013; Bekaert and Hoerova, 2014). Lastly, investors are net buyers of index options (Gârleanu et al., 2009) and the volatility risk premium represents market makers' willingness to absorb inventory and provide liquidity (Nagel, 2012).

Our study builds upon these insights, focusing on how the market prices volatility within the context of intermediaries that provide liquidity to investors who seek to hedge their downside tail risk; that is, a supply and demand framework. We begin with a statistical analysis of the data, using large-scale data collected from several different sources. This introduces the additional challenge of having to collate data from multiple platforms, in different formats, often on differing time scales. Such challenges are commonplace in "Big Data" research and give rise to problems such as spurious correlation, time asynchronicity, and noise accumulation (see Fan, 2013; Fan et al., 2014). However, using tools from high-frequency finance and big data analytics, we are able to obtain a clean and more precise estimate of the true integrated volatility without any reliance on a model or parametric assumptions.Our consistently estimated realized volatility is then compared to a model-free measure of implied volatility, proxied by the VIX index to get our time-series of the volatility risk premium. We make the distinction between this ex-post formulation of the volatility risk premium with other studies and note that our resulting time series is the daily realization of the volatility risk premium. We still interpret this as the market price of volatility risk, but consistent with our data-driven philosophy, frame our measure of the volatility risk premium as a systematically priced bias.

When analyzing a statistical bias, sometimes it is insightful to analyze the magnitude of the bias separate from the direction. In fact, based on previously observed features of the volatility risk premium in the literature, we are able to construct several testable hypotheses about the underlying drivers of the magnitude (absolute value) and direction (sign) of the volatility risk premium. We are the first to perform a comprehensive analysis on the magnitude and direction of the systematically priced bias associated with the volatility risk premium, which is a significant contribution of our work. The findings are much stronger, both statistically and economically, when looking at the determinants of the absolute value of the volatility risk premium. First of all, while in theory or in expectation, the volatility risk premium should be negative, the realization of the volatility risk premium can be (and sometimes is) positive. This makes sense when viewed in terms of the P&L on volatility swaps; otherwise one side of the volatility swap would always make money which would be inconsistent with no arbitrage. However, this raises a paradox with our observed results and the stylized facts from the existing literature: how does one interpret long stretches of a positive volatility risk premium? Does this mean that investors go from being risk-averse to risk-loving? Or that they have to be paid to hedge volatility risk? Whereas common sense may hint that these cannot be right, our analysis confirms that is not what is happening. For example, we find that the biggest spike and largest stretch of a positive volatility risk premium occurs at the depth of the Financial Crisis – right after Lehman Brothers failed in Fall of 2008. Herein, our data-driven approach shows its true value. We find that the magnitude of the volatility risk premium, that is the absolute value of the bias between the realized and implied volatility, reflects supply and demand imbalances in the index option market where investors buy protection on downside tail risk. The demand effects are captured by a statistically significant relationship between open interest on S&P 500 index put options and the magnitude of the volatility risk premium. The supply effects, are captured by the *TED* spread and the credit spread. The former often appears as a proxy for liquidity (or illiquidity) in financial markets and is viewed a general measure of financial instability. The latter, which is most significant during the Financial Crisis, interestingly has the interpretation of dealers deleveraging and shrinking their balance sheets by selling off risky positions. This is consistent with the evidence provided by Adrian and Shin (2010).

As for the direction of the volatility risk premium, practitioners believe that the volatility risk premium's sign is indicative about future levels of realized volatility.

When the volatility risk premium is negative, implied volatility is higher than realized volatility and market participants believe that volatility is likely to increase in the future. On the contrary, when the volatility risk premium is positive, implied volatility is less than realized volatility and market participants believe that volatility is likely to decrease in the future. This is related to the contentious idea of implied volatilitys ability to forecast future realized volatility in the literature and, more recently, the "Expectation Hypothesis" discussed in Aït-Sahalia et al. (2012). In fact, Aït-Sahalia et al. (2012) provides a way to test this hypothesis about the direction of the volatility risk premium. An alternative explanation in the finance literature says that the sign of the volatility risk premium represents the gains or losses on market makers delta-hedged positions. This was first proposed by Bakshi and Kapadia (2003) within the context of stochastic volatility and jump-diffusion models.[2] Using our model-free, data-driven analysis, we are able to find much stronger evidence in favor of this explanation. This is also consistent with the realized P&L on a long position in a volatility swap.

The remainder of this paper is structured as follows. The next section, Section 2, provides more detail on the volatility risk premium and why implied volatility can, and will, deviate from realized volatility. We first make the distinction between the ex-ante formulation of the variance and volatility risk premium, which requires the use of an econometric (AR) model to forecast expected future realized variance/volatility from lagged values, and our ex-post realization of the volatility risk premium and its interpretation as a systematic bias that is priced by the market. In Section 2 we also propose our hypotheses. Then, in Section 3, we review methods for computing realized volatility with emphasis on the estimation of integrated volatility with high-frequency data. Section 4 details the data collection and methodology for constructing the volatility risk premium. Our empirical analysis and results are presented in Section 5. Section 6 concludes. We have two Appendices: Appendix A covers the technical details on the Fourier transform method that we use to address the microstructure

[2]In a Black-Scholes-Merton world (Black and Scholes, 1973; Merton, 1973) gains or losses on the delta-hedge perfectly offset the loss or gain on the short option position. However, that is not the case with jumps and, in fact, there is evidence that downward jumps explain much of the pricing of volatility risk.

noise in our estimation of integrated volatility with the ultra-high-frequency data; Appendix B presents simulations that demonstrate the benefit of using ultra-high-frequency data in estimating integrated volatility and the extent to which our method performs better than alternatives.

2 Implied Volatility and the Volatility Risk Premium

The notion of implied volatility is well understood and widely used by options traders and financial engineers. In this section we briefly review the concept of implied volatility and discuss in greater detail the volatility risk premium which is imbedded in implied volatilities but not included in realized volatilities. Therefore, if we wish to look at the difference between realized volatility and implied volatility, it gives us a measure of the volatility risk premium. We further view this deviation between realized and implied volatilities in an ex-post sense as a priced bias in the options markets.

While it is well known that the Black-Scholes-Merton option pricing model relies on unrealistic assumptions and therefore cannot reasonably price options, the model is still widely used by traders to infer the level of volatility associated with a particular option on a given asset (e.g. stocks, indexes, or currencies). The idea is that since volatilities are unobservable but option prices are observed and convey traders expectations about the future riskiness of the underlying asset over the life of the option, the financial engineer is faced with a mathematical inverse problem. Given the output of the model – option price – solve for the value of the volatility parameter that sets the model value equal to the market price of the option. This is the *implied volatility*. Now, of course, any option pricing model can be used; it is just that the Black-Scholes-Merton model is the most basic and convenient. Regardless, implied volatilities computed in this fashion are by definition model-dependent and constrained by parametric restrictions.

Researchers including Rubinstein (1994), Dupire (1994), and Derman and Kani (1994) have, to much success, extended this idea of implied volatility to extract market information across entire classes of options on a given asset (i.e. different strikes and/or maturities; often referred to as the "volatility smile" or "volatility surface") and fit a deterministic function of asset price, strike price, and time to expiry.

More recently, the work of Britten-Jones and Neuberger (2000) and Jiang and Tian (2005) broke away from the reliance on models and derived and implemented a *model-free implied volatility* using only current option prices. It is along these lines that we would like to use implied volatility in our data-driven analysis.

The implied volatility is forward-looking and represents the market's expectation of volatility over the life of the option. Mathematically, it can be thought of as an expectation under the risk-neutral or pricing measure. The volatility as computed from the underlying asset price movement can be thought of as being generated under the physical or statistical measure. The difference between the two represents the market price of volatility risk, or what is referred to as the "volatility risk premium". In some instances it is easier to work with the squared volatility, which leads to the variance risk premium defined as follows.

$$VRP_t = \mathbb{E}_t^{\mathbb{P}}\left[\int_t^T \sigma_u^2 du\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T \sigma_u^2 du\right]. \tag{1}$$

If you take the square root of each of the expectations in Equation (1) you would get the *volatility* risk premium which we will denote as a lowercase vrp_t. While we present all of our results in terms of the *volatility* risk premium, everything is robust with respect to the variance risk premium. The reason we chose to use the former is because the results are easier and more natural to interpret in terms of vol units. Evidence that the two may be used interchangeably is found throughout the literature. A recent example is Drechsler and Yaron (2011) where volatility is the object of interest, but the quantity used in the analysis is the variance risk premium. They define the "variance premium" as the difference between *VIX* squared and the conditional expectation of the realized variance. Conceptually, this definition follows the idea that, for a financial instrument, the risk premium is the difference between the price of the contract (VIX squared) and the expected payoff of the contract (realized variance). In fact, in most volatility (variance) risk premium studies the expectation under the risk-neutral measure will be proxied by the *VIX* volatility index (*VIX* squared). Using *VIX* will be nice for our purposes as it is closely related to the aforementioned model-free implied volatility.[3] We will discuss *VIX* more in Section 4.2. The reader may note that the Drechsler and Yaron (2011) setup is the reverse of our definition given in Equation (1). The decision of how to sign the volatility or variance risk premium is a matter of personal preference and perspective. We follow Carr and Wu (2009) who take the perspective that the negative sign reflects investors' willing to *pay* to hedge their volatility risk. The idea of hedging downside tail risk will become a central theme in explaining our empirical results. However, we further assert that the sign of the volatility risk premium plays a secondary role to the *magnitude* of the volatility risk premium when trying to find the underlying economic determinants over time.

In many studies the first term in Equation (1) is computed as an ex-ante conditional expectation of the future realized volatility or variance given the current value through an autoregressive model. We intentionally choose not to do this, but rather use an ex-post measure of the realized volatility. The reason is computing the ex-ante conditional expected realized volatility introduces model error and possible misspecification bias. Instead we compare the ex-post realized volatility (averaged over a one month period) to the model-free implied volatility (covering the same horizon) to get a realization of the volatility risk premium on each trading day over the sample period. In symbols, this is

$$vrp_t = \sqrt{\int_t^T \sigma_u^2 du} - \sqrt{\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T \sigma_u^2 du\right]}. \tag{2}$$

[3] See Carr and Wu (2006) and Jiang and Tian (2007) for more on the relationship between the model-free implied volatility and the *VIX* volatility index.

Therefore another distinction from the uppercase VRP_t in Equation (1) and the lowercase vrp_t in Equation (2) is that the latter is our *bias* representation of the volatility risk premium. Furthermore, once we substitute the level of the *VIX* volatility index in for the risk-neutral expectation, it would be the realization of the volatility risk premium which is similar to the realized P&L on a volatility swap (see Demeterfi et al., 1999).

Recent studies have been able to establish some interesting empirical properties of the volatility/variance risk premium. Carr and Wu (2009) note that traditional risk factors have very little explanatory power for the variance risk premium (we are able to confirm this in our empirical analysis of the volatility risk premium). They suggest that there is an independent risk factor that is driving the principally negative variance risk premium. Furthermore, they find evidence that the *VRP* is time-varying. Bollerslev et al. (2009) study the predictability of the variance risk premium on stock market returns from 1990 to 2005. They find that there is a strong, statistically significant positive relationship between the *VRP* and quarterly future stock returns. They note that the predictive power is better than other financial and macroeconomic factors that are typically used in stock market return forecasting. Bollerslev et al. (2011) examine the volatility risk premium and its relation to several macro-financial state variables. They find that the *vrp* exhibits significant temporal dependencies related to the macro-finance state variables and is also able to help predict stock market returns. Zhou (2011) studies the predictability of the variance risk premium across financial markets through equity returns, bond returns, and credit spreads. He observes that the *VRP* predictability maximizes typically in the one to four month horizon, and the short-run risk premium dynamics can be interpreted within a general equilibrium model which prices stochastic economic uncertainty. The calibrated model can help explain the equity premium puzzle and the credit spread puzzle in the short-run. However, it remains a challenge to incorporate long-run predictability patterns of consumption growth and asset returns found in literature.

Several studies have examined the role that jumps play. Using high-frequency index futures data, Wu (2011) computes maximum likelihood estimators of the instantaneous realized return variance. His analysis shows that both the jump arrival rate and the absolute value of the negative variance risk premium are proportional to the variance level. This last finding, in terms of the absolute value, will be very relevant in developing our hypotheses below and in interpreting the results of our analysis. Specifically, it is evidence that when volatility is high, the volatility risk premium is either very positive or very negative; that is the bias between realized and implied volatility increases when the level of uncertainty is heightened.

Todorov (2010) analyzes the variance risk premium under a semi-parametric stochastic volatility model with the inclusion of price jumps. The model parameters are estimated by *GMM* with high-frequency data on the five-minute return of S&P 500 index futures contract from 1990 to 2002. The results provide empirical evidence that investors are willing to pay for protection against jumps, especially when preceded by recent jumps, which supports the hypothesis that risk aversion is time-varying and that the volatility risk premium represents the cost of protection against market crashes.

The main takeaways are that the volatility risk premium appears to be a priced risk factor in the capital markets (both equity and credit) and investors are willing to pay a premium to hedge their downside risk, especially when uncertainty is high. However, our knowledge is still very limited about the determinants of the volatility risk premium and we do not have sound empirical evidence documenting what exactly the volatility risk premium says about the mechanics of the market for pricing and hedging risk.

Bollerslev et al. (2011) refer to their estimate of the *vrp* as a "risk aversion index". It seems that many practitioners agree with this interpretation of the volatility risk premium, and we are able to find some evidence that supports this point. This leads to something of a paradox: suppose we choose to define the volatility risk premium such that it is typically negative, thereby indicating that market participants are willing to *pay* to hedge their volatility risk. Then, when the *vrp* gets *more negative* it indicates that investors are becoming *more risk averse*. But then how do we explain the occurrence of a large positive spike in the volatility risk premium; investors becoming *risk loving*? As we will show, the data indicates several instances over our five year sample period where the *vrp* turns positive, most notably during the Financial Crisis. Surely, investors did not become risk loving during the Financial Crisis.

Our prior is that the positive spikes in the volatility risk premium reflect liquidity conditions in the financial markets. Consequently, several recent papers in the financial economics literature have linked the *vrp* to liquidity, intermediation, and hedging demand. These papers provide the conceptual underpinning that we use to construct stylized facts and testable hypotheses about the economic meaning of the volatility risk premium. First, the volatility risk premium represents option market makers' willingness to absorb inventories and provide liquidity (Gârleanu et al. (2009), Nagel (2012)). Also, investors are net buyers of index options (Gârleanu et al. (2009)). To the extent that investors use index put options to hedge their downside tail risk, then we should be able to use option market data to draw inferences about investors' demand for hedging downside tail risk and intermediaries' willingness to meet this demand (i.e. provide liquidity). The volatility risk premium can, therefore, naturally be interpreted as the compensation that option market makers receive for this intermediation and liquidity provision to meet hedging demand. Adrian and Shin (2010) find evidence of this interpretation in the expansion and contraction of financial intermediaries' balance sheets.

Even within this conceptual framework of intermediation and liquidity provision, the existence of a positive *vrp* is still a bit puzzling. Does this mean that periods of positive *vrp* indicate that sellers of volatility have to pay hedgers in order to meet their demand? Rather, perhaps the direction (sign) contains different information than the magnitude of the *vrp*. By some accounts, traders view the sign of the volatility risk premium as indicative of nothing more than the market's expectation of future levels of volatility. It is then the magnitude of the volatility risk premium that represents the actual *price* of volatility risk. The magnitude of the volatility captures the extent to which market makers are willing to absorb inventory, provide liquidity, and meet hedging demand. When demand for hedging downside tail risk increases, market makers will take the short side (sell volatility) but must be compensated appropriately. The price of volatility increases and implied volatility rises relative to realized levels. When demand for hedging downside tail risk decreases, there will be a selloff of volatility and market makers will take the other side, but only at a substantial discount. Implied volatility falls relative to realized levels. Therefore, the magnitude captures the extent to which market makers must be compensated to provide liquidity to the options markets, either as a premium or discount if intermediaries are selling volatility to meet hedging demand or buying it back in response to a reduction in hedging demand.

Taking the view of the volatility risk premium as a *systematically priced bias* we decompose the *vrp* in Equation (2) into sign and magnitude as follows

$$vrp_t = \left| \sqrt{\int_t^T \sigma_u^2 du} - \sqrt{\mathbb{E}_t^{\mathbb{Q}} \left[\int_t^T \sigma_u^2 du \right]} \right| \times \text{sgn}(vrp_t). \tag{3}$$

.

We then perform several statistical and econometric tests on each of the components of *vrp* in Equation (3). First we must introduce our testable hypotheses about both the magnitude and direction of the volatility risk premium.

Magnitude Hypotheses:

H1: *The magnitude of the volatility risk premium reflects investors' demand for hedging tail risk.*

H2: *The magnitude of the volatility risk premium reflects the willingness of option market makers to absorb inventory and provide liquidity.*

Direction Hypotheses:

H3: *The sign of the volatility risk premium contains information about future levels of realized volatility relative to implied volatility.*

H4: *The sign of the volatility risk premium reflects the delta-hedged gains or losses for option market makers.*

We should note that the hypotheses need not be mutually exclusive. Hypothesis H1 can be thought of as demand-side effects and Hypothesis H2 can be thought of as supply-side effects. We may, therefore, find that supply and demand forces work with or against each other to determine the magnitude of the *vrp* at a given time.

In order to econometrically test the volatility risk premium, we must come up with an accurate and clean measure of the actual volatility in the market. With the growth of high-frequency financial data and the application of continuous time finance to the analysis of such data, the tools for estimating the integrated volatility of a price process have become plentiful. In the next section, we first review various estimators and then present our methodology for computing the volatility risk premium using ultra-high-frequency data.

3 Realized Volatility and Estimating Integrated Volatility with High-Frequency Data

In this section we review some of the existing methodologies for estimating volatility, with emphasis on recent advances in the use of high-frequency data. Typically, the modeler will assume that the latent true (log)price X_t follows an Ito process

$$dX_t = \mu_t dt + \sigma_t dW_t \tag{4}$$

where W_t is a standard Brownian Motion and μ_t and σ_t are time-varying drift and volatility, respectively, that may or may not follow stochastic processes themselves.[4] However, what we observe is the transaction price, or its logarithm, Y_t at times $\{t_i\} \in [0, T]$, which are related to X_t according to

$$Y_{t_i} = X_{t_i} + \epsilon_{t_i}, \tag{5}$$

[4]Some methodologies can also be applied to jump-diffusion processes rather than just pure diffusion processes.

Note that the ϵ_{t_i} in Equation (5) represents *microstructure noise.*

The goal is to use observable price data to estimate the volatility σ_t in Equation (4). It is important to note the different assumptions on the structure of σ_t and ϵ_t as quite often these are the subtleties that set one method apart from another.

First, suppose we observe regularly spaced Y_{t_i}, where $t_i - t_{i-1} = \Delta$. Then, let us define $n = \frac{T}{\Delta}$; i.e. n is the number of sampled data points. If σ_t is modeled parametrically as constant σ, and the noise distribution is assumed to be Gaussian with mean 0 and variance a^2, then the log-likelihood function of $\delta Y_i = Y_{t_i} - Y_{t_{i-1}}$ is

$$l(\sigma^2, a^2) = -\frac{1}{2}\log\det(\Omega) - \frac{n}{2}\log(2\pi) - \frac{1}{2}\delta Y' \Omega^{-1} \delta Y, \tag{6}$$

where Ω is the covariance matrix of δY and Ω^{-1} can be calculated explicitly.

Choosing σ and a to maximize Equation (6) gives the Maximum Likelihood Estimator, or *MLE*, of volatility. It can be shown that the *MLE* is consistent for both the volatility component σ^2 and the noise component a^2 at rates $O_p(n^{-1/4})$ and $O_p(n^{-1/2})$, respectively. Moreover, misspecification of the marginal distribution of ϵ does not have adverse consequences. (Aït-Sahalia et al. (2005), Xiu (2010))

The assumption that volatility is constant is probably not very reasonable. There is considerable evidence of time-varying volatility which means that we have to come up with a way to estimate the instantaneous volatility process σ_t either parametrically or nonparametrically (see, e.g., Andersen et al. (2004)). Quite often we are interested in estimating the *integrated volatility* over a period of time. This is done by making use of the *quadratic variation*, $\langle X, X \rangle_T$, of the stochastic process described by Equation (4). The quadratic variation is

$$\langle X, X \rangle_T = \int_0^T \sigma_t^2 \; dt. \tag{7}$$

We want to estimate this quantity using the observable price data. A naïve estimator would be the Realized Volatility (*RV*) estimator

$$[Y, Y]_T = \sum_{i=1}^{n} (Y_{t_{i+1}} - Y_{t_i})^2, \tag{8}$$

which is a consistent estimator in a noise-free model. However, since the observable price process given by Equation (5) is contaminated by the microstructure noise this *RV* estimator is both biased and inconsistent.

Statistical theory indicates that we should be able to improve the accuracy and precision of our estimate by increasing the rate at which we sample the data; hence the value of ultra-high-frequency data.[5] This would be the case if we could observe X_t directly; but microstructure noise introduces an added dimension of complexity to the problem. In fact, assuming iid noise, the bias of the *RV* estimator is $2n\mathrm{E}[\epsilon^2]$. This tells us that as we increase the frequency of the price data, the effect from noise becomes more overwhelming.

One way to address the problem of noise when sampling at too high of a frequency is to sample sparsely and use the corresponding *RV* estimator. This practice, known as the *subsampling* approach, was first introduced by Zhou (1996). However, even when sampling sparsely at the optimally-determined frequency, the fact that large portions of data are discarded violates basic statistical principles. Furthermore, Zhang et al. (2005) argue that sampling over longer horizons merely reduces the impact of microstructure, rather than quantifying and correcting its effect for volatility estimation.

One of the earliest solutions to incorporate the full data sample is Two Scales Realized Volatility (*TSRV*) as proposed by Zhang et al. (2005). The *TSRV* estimator is based on subsampling, averaging, and bias-correction. They sample sparsely over subgrids of n observations to get K subsamples on a slower time scale. For each such sample the *RV* estimator is $[Y, Y]_T^{(sparse,k)}, k = 1, \cdots, K$. Averaging them yields the

[5] See Appendix B where we use simulations to illustrate the benefit of using ultra-high-frequency data.

estimator $[Y,Y]_T^{(avg)}$, and the final de-biased estimator is:

$$\widehat{\langle X, X\rangle}_T^{(tsrv)} = [Y,Y]_T^{(avg)} - \frac{\bar{n}}{n}[Y,Y]_T, \tag{9}$$

after accounting for the bias, where $\bar{n} = \frac{n}{K}$. Choosing the optimal sampling step $K = cn^{2/3}$ yields the convergence rate $n^{-1/6}$. The *TSRV* estimator is shown to outperform the standard *RV* estimator empirically in the study by Aït-Sahalia and Mancini (2008).

A closely related estimator is Multiple Scale Realized Volatility (*MSRV*), which is proposed and derived in Zhang (2006). As a generalization of *TSRV*, the *MSRV* estimator combines M different time scales with weights, when chosen optimally, can achieve the optimal convergence rate $n^{-1/4}$. Based on a different smoothing idea, Fan and Wang (2007) introduces a different estimator achieves the same rate, but allows jumps in the price processes.

Realized kernels, which are based on linear combination of autocovariances, represent another popular class of estimators. Barndorff-Nielsen et al. (2008) designed several realized kernels which are robust to endogenous sampling and noise. The realized kernel estimators can achieve convergence rates up to that of *MSRV*. Barndorff-Nielsen et al. (2009) discuss details of implementing the realized kernel methodology to estimate integrated volatility with high-frequency data. Barndorff-Nielsen et al. (2011) are able to achieve consistency for one of the more problematic realized kernel estimators by making use of subsampling.

The pre-averaging approach of Jacod et al. (2009) uses all or most of the data, but averages over a moving window. The averages are used to compute the realized volatility, which then have to be adjusted by an additive term to eliminate bias. The result is a rate optimal (with convergence rate $n^{-1/4}$) consistent estimator of integrated volatility in the presence of microstructure noise. In many ways, one can think of the pre-averaging approach as removal of microstructure noise by local average smoothing. Additionally, pre-averaging is an effective method of data cleaning.

Returning to the parametric approach, if we do not assume that volatility is constant and noise normally distributed with variance a^2, but nevertheless use the log-likelihood function in Equation (6), the resulting estimator is the quasi maximum likelihood estimator (*QMLE*). Interestingly, it is still a consistent estimator at the most efficient rate $n^{-1/4}$. Statistical properties of the *QMLE* are derived in Xiu (2010).

Finally, another interesting approach involves working on the frequency domain rather than the time domain. As such, it relies on the Fourier transform (see Olhede et al. (2009)). The procedure computes a consistent and unbiased estimator of integrated volatility at ultra-high-frequencies under very general specifications of the microstructure noise process. This is the methodology that we employ to estimate integrated volatility for our study.

With the frequency domain method, integrated volatility is estimated through the variance of the Fourier transform of the increment process. Under the rationale that the high-frequency coefficients are more heavily contaminated by the noise, the de-bias procedure is done locally at each frequency. The unknown parameters involved in the de-bias are estimated through *MLE* using a Whittle likelihood function. This frequency domain methodology allows us to easily model autocorrelated noise as a moving average process, and then disentangle the noise effect at each frequency in the same way. The order of the moving average process – i.e. the appropriate number of lags in the autocorrelated noise – is determined through model selection using the corrected Akaike information criteria (*AICC*). Technical details regarding this procedure can be found in our Appendix A.

The frequency domain method for estimating integrated volatility has several desirable features, both in terms of the statistical properties and the practicality in applying to real financial data. From the financial modeling and data analysis point-of-view, working in the frequency domain provides an elegant way to address more general specifications of the microstructure noise process. For the most part, the other methods discussed above assume that the noise process, ϵ_t, is iid or uncorrelated. However, in practice this is an unreasonable assumption.[6] Autocorrelated microstructure noise may be a more reasonable assumption, since large disturbances this second may be highly correlated with large disturbances last second, especially if there is a lot of noise in the market. This may give the impression that the market is more turbulent or volatile, when in fact the persistent volatility in our observed time series, Y, is coming from the microstructure noise. Thus, we need a clean way to strip away the true volatility of the price process in the presence of microstructure noise at ultra-high-frequencies. This is why we use the frequency domain estimation method in computing integrated volatility. In fact, we find that the data indicates the latent noise process has on average lag-1 autocorrelation, and the time varying order of autocorrelation ranges from 0 to 5.

4 Data and Methodology

4.1 Data Collection

The data used for our empirical analysis came from several different sources, on multiple platforms, and were analyzed using a variety of softwares. This is a common feature of Big Data analytics and requires careful processing and collating to ensure that the data are in consistent formats, with large-scale computations often being done in parallel (see Fan et al., 2014). First, we began by cleaning and processing the ultra-high-frequency transaction data for the SPDR ETF. Then we used the cleaned price data estimate the integrated volatility on the frequency domain via the Fast Fourier Transform (FFT) algorithm. The computed integrated volatility was then merged with a daily time series of the *VIX* index, and the difference between the two time series gave us the volatility risk premium. Finally, we had to collect, clean, and merge with the economic, financial market, and risk factor variables from

[6]While most previous approaches assume iid microstructure noise, recent work by Aït-Sahalia et al. (2011) addresses the complicated issue of estimating volatility from ultra-high-frequency data with dependent microstructure noise. Our methodology is similar in that it also permits estimation of integrated volatility for more general classes of microstructure noise, but with less parametric restrictions.

their respective databases. This data was used in our econometric analyses of the determinants and drivers of the volatility risk premium.

Transaction price data for the SPDR ETF (ticker *SPY*) was obtained from the TAQ database within WRDS. The sample period we studied goes from July 2006 to June 2011. Over these five years there were a total of 523,814,632 trades. For our integrated volatility estimation method to work best, we need as many observations as possible.[7] Trade volume decreases considerably as we go further back in time which is why we stop at 2006. The first year of data (2006-2007) has approximately one-quarter the number of trades as the final year of data (2010-2011). Additionally, this sample period contains about the same number of observations Pre-Crisis, Crisis, and Post-Crisis for better comparison across subperiods.

For data cleaning and processing purposes, we filtered the data based on the "Correction Indicator" (CORR) and "Sale Condition". We kept only transactions where CORR=00; these represent regular trades that were not cancelled or corrected. This resulted in only 0.003% of the data being removed from the sample, leaving us with 523,796,850 trades remaining. We also eliminated any "Special Condition Trades" which introduced suspicious and irregular patterns in the transaction price sequences (i.e. large jumps that were immediately reversed). This resulted in 1.8% of the data being removed from the sample leaving us with 514,270,624 trades remaining.

Since multiple trades can occur in any given second, we next introduced an aggregation step in the data processing. This would allow us to have a second-by-second time series of *SPY* prices. We tried two methods for aggregation: median and size-weighted average price and did not find significant aberrations. Finally, we had to include an expansion step to account for seconds where no trades were executed. To address these instances we used piecewise constant interpolation; i.e. if there was no trade at second t then we filled it with the last executed price $t-1$ ("last tick").

[7] We illustrate this principle with simulations in Appendix B. The simulations show that, under our method, sampling at higher frequencies allows for the most precise estimation of integrated volatility. Our method performs better than naïve subsampling rules that are typically used in high-frequency studies, and as noted earlier, has the added benefit that the microstructure noise can be autocorrelated and so we need not restrict ourselves to the case where microstructure noise is independent over time.

This resulted in 29,461,859 second-by-second data points covering 1,259 trading days. This was the data that was used to compute our daily time series of monthly realized volatility (on a rolling 21 trading day basis) via the frequency domain estimation methodology.

The daily opening level of the *VIX* volatility index was obtained from the CBOE database. As discussed in Section 2, the *VIX* is a model-free implied volatility extracted from near-term put and call options on the S&P 500 index.[8]

The explanatory variables in our regressions also come from multiple sources. First, we have the traditional risk factors from the Fama-French Three Factor Model (Fama and French, 1993); the data for the Fama-French factors are available from Kenneth French's website.[9] We also include the credit spread, also known as the default risk premium, which is the difference in yield on Baa-rated and Aaa-rated corporate debt. The yields on corporate debt, by Moody's rating, are available from the FRED database maintained by the Federal Reserve Bank of St. Louis. Use of the credit spread as a risk factor in asset pricing studies goes back to Chen et al. (1986) and has the interpretation as a measure of investor risk aversion. It has subsequently been used in volatility risk premium studies such as Zhou (2011). We will see that there is also a supply-side interpretation of the highly significant effects that the credit spread has on the volatility risk premium (and its magnitude), especially during the Crisis subperiod. The *TED* spread is included to capture liquidity effects in the financial markets and as a measure of distress in the financial system. The *TED* spread is the difference between 3-month Eurodollar rates and 3-month Treasury rates, both of which are also available through the FRED database. The interpretation of the *TED* spread follows from the logic that as uncertainty in the financial system heightens, financial institutions charge more to each other for short-term borrowing– this is reflected in Eurodollar rates; at the same time they require

[8]For details on the methodology used in constructing the *VIX* volatility index please see CBOE (2009). A similar methodology is employed in Jiang and Tian (2005) where the information content of model-free implied volatility is studied. The CBOE volatility index is studied in Carr and Wu (2006) and Jiang and Tian (2007).

[9]`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

better collateral, which drives up demand for Treasury Bills and pushes down their rates (Brunnermeier, 2009).

As a final explanatory variable we include the daily open interest for put options on *SPY*, which is obtained from the OptionMetrics database. We use this as a proxy for investors' demand for hedging tail risk, following the combined logic of several recent options studies. Cao and Han (2013) use open interest as a proxy for option demand pressure. Additionally, Gârleanu et al. (2009) find that investors are *net buyers* of index options. Since S&P 500 index put options give investors a way to hedge against market-wide crashes, the open interest provides a natural proxy for investors' demand to purchase protection and hedge downside tail risk. While options on the SPDR ETF (*SPY*) are different from index options on the S&P 500 (*SPX*), they essentially provide the same protection for investors and have some features that may make them more attractive (see Kelly et al., 2012). In fact, our choice to use SPDR options might be even more consistent with our desired proxy as demand for hedging downside tail risk, since the former are American-style options with physical settlement (*SPX* index options are European-style options with cash settlement only) and therefore give the investor more flexibility and robustness in protecting themselves against market crashes. Furthermore, when we discuss market makers' delta-hedging, SPDRs would be a more effective hedge on *SPY* put options than on *SPX* index options.

The explanatory variables are summarized in Table 1. Descriptive statistics for all variables are given in Table 2, which will be referred to throughout the discussion.

4.2 Construction of the Volatility Risk Premium

In this section we discuss our construction of the volatility risk premium from the market data. Consistent with our representation of the *vrp* as a "bias", we calculate it as the deviation of the realized volatility from the expected volatility implied by option prices. Therefore we first compute the realized volatility as the estimated integrated volatility using the Fourier method, described in Section 3 and detailed in Appendix A, on the ultra-high-frequency transaction data for the SPDR S&P 500 ETF (ticker *SPY*). This gives us a daily time series of the realized volatility; however, there is a lot of variation in the day-to-day realization of this quantity which will contribute to additional statistical noise in our attempt to quantify the systematic bias that is the *vrp*. To better represent this systematic bias we smooth the time-series of realized volatility by taking a rolling average of the next 21 trading days so as to cover the same month as the contemporaneous *VIX* index, which is our measure of model-free implied volatility. Define this rolling average as $\{\overline{RV}_t\}_{t=2006.07}^{2011.06}$, where $\overline{RV}_t = \frac{1}{21}\sum_{i=1}^{21} RV_{t+i-1}$ and RV_i represents the realized volatility for day i computed using the frequency domain methodology.

The time series of the *VIX* Open value, $\{VIX_t\}_{t=2006.07}^{2011.06}$ is then subtracted from the average realized volatility to measure the extent that the implied volatility represents a biased expectation of the future realization:

$$vrp_t = \overline{RV}_t - VIX_t. \tag{10}$$

We use the Open value (rather than the Close) of *VIX* so as to be consistent with our realized volatility estimate in terms of the 21-trading-day period for which we are looking at on any given day.[10]

A time series of our computed *vrp* over the sample period is plotted in Figure 1. Looking at the Figure, two things stand out immediately: first, the risk premium is negative throughout most of the sample period; second, there are a few pockets where the *vrp* goes positive– most notably in the third quarter of 2008. That large positive spike which extended for a period of more than two months seemed to be anomalous to what most of the literature says. To the extent that the negative *vrp* represents investors being risk averse, does a positive *vrp* mean that investors went from risk averse to risk loving during this time? That certainly does not seem right, since that period includes the failure of Lehman Brothers and the plunging of the global economy into the worst financial crisis in history. So perhaps it means that investors

[10]For robustness, we ran the regressions using the *VIX* Close value and found that the results (sign, significance) are consistent for the variables of interest in our study.

who typically pay to hedge volatility risk were no longer willing to, but rather required compensation (i.e. be paid) to enter into any volatility related transaction? A similar story was told about the negative yields on T-Bills during the Fall of 2008, but it doesn't seem to fit with what is going on in our data. So perhaps it is a fictitious by-product of the frequency domain estimation methodology? Fortunately, that was not the case as we were able to confirm positive *vrp* including the third quarter of 2008 using other methods for computing realized volatility (i.e. the *TSRV* estimator of Zhang et al. (2005) and the pre-averaging estimator of Jacod et al. (2009)).

One possible explanation for this large positive spike is that the option markets underpriced the actual volatility level during that period of time. Unexpected shocks such as Lehman Brothers' failure and subsequent government interventions kept the markets on edge, and it was impossible to know the magnitude of such a market tsunami and its impacts on realized volatilities, a priori. The idea that the government would provide a backstop against any large financial catastrophe, known as the "Fed put", was arguably priced into the market keeping implied volatilities low relative to realized volatilities. Therefore, there was a strong bias in one direction with implied volatility underestimating the realized volatility. This eased a bit after Lehman did fail, but when officials were quick to step in thereafter, it remained to keep implied volatilities lower than perhaps they should have been given the circumstances. After the government programs such as TARP and QE were in place and it became clear there would be no "quick fix", implied volatilities rose relative to realized levels (even though both were rising steadily during this entire period because of the high degree of overall uncertainty) thus reversing the bias.

5 Empirical Analysis

5.1 Preliminary Regressions

We start with a couple of baseline regressions. First, we ran a standard regression with the traditional risk factors from the Fama-French Three Factor Model (Fama and French (1993)).[11] The Fama-French regression is

$$vrp_t = \beta_0 + \beta_1 Mkt_t + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_t. \tag{11}$$

The results for the Fama-French regression can be found in Table 3.[12] Note that although the market risk premium is highly significant (*HML* is also significant at the $< 5\%$ level), the R-squared is very small (less than 3%) indicating that traditional risk factors have very little explanatory power for the volatility risk premium. This is consistent with previous findings in other studies.

We now introduce additional risk factors that may have theoretical links to the volatility risk premium as indicated by the literature (e.g. Gârleanu et al. (2009), Nagel (2012), etc.). We want to be able to capture demand for hedging tail risk, liquidity provision, and the overall stability of the financial system. We proxy demand for hedging tail risk with the open interest on *SPY* put options. The *TED* spread is viewed by many as a proxy for liquidity risk and a measure of distress in the financial sector (see Brunnermeier, 2009). The credit spread, or the default risk premium, can be viewed as a measure of macro-level risk aversion, but also has a nice interpretation in terms of liquidity provision capturing the de-leveraging and risk reduction that occurred after the onset of the Financial Crisis. This next regression is specified as

$$vrp_t = \beta_0 + \beta_1 Mkt_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 CS_t + \beta_5 TED_t + \beta_6 POI_t + \epsilon_t. \tag{12}$$

The results for this regression over the whole sample period are reported in Table 4. Here we see that the inclusion of the additional factors – credit spread, *TED* spread, and put option open interest – improve the explanatory power substantially as the R-squared is over 40%. The market risk premium remains significant at the $< 1\%$ level; the credit spread and put option open interest are also significant at the $< 1\%$ level. *TED* spread is not significant.

[11] Data for the Fama-French factors are available from Kenneth French's website: `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[12] All regressions results are reported with Newey-West robust standard errors.

Continuing with our view that the *vrp* is a priced bias, we next decompose it into magnitude and sign and test the hypotheses proposed in Section 2. We start by examining the magnitude of the bias and its relationship to supply and demand imbalances in the market for hedging tail risk.

5.2 Magnitude Regressions

Our next econometric specification is to use the explanatory variables from Equation (12), but now regressing the *magnitude* component of the volatility risk premium on them to see what additional insights might be obtained within the context of our hypotheses. The magnitude regression is specified as

$$|vrp_t| = \beta_0 + \beta_1 Mkt_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 CS_t + \beta_5 TED_t + \beta_6 POI_t + \epsilon_t. \quad (13)$$

The results of the magnitude regression for the whole sample period are shown in Table 5. By regressing the absolute value of *vrp* on the explanatory variables, the R-squared increases to 58.32% compared to 41.55% in the original *vrp* regression specified by Equation (12). That is, that over our entire sample period the factors are able to explain more than half of the variation in the magnitude of the volatility risk premium. Since the most significant variables are the credit spread and the put option open interest, much of the explanatory power may be attributed to these supply and demand effects after controlling for the traditional risk factors.

Additional evidence for the value of studying the size of the bias apart from the direction can be seen by comparing the sign and economic meaning of the coefficient estimates in the original regression (Table 4) with those in the magnitude regression (Table 5). Both regressions have credit spread as one of the most significant factors. Note, however, that the p-value associated with the magnitude regression is much smaller than that on the original *vrp* regression. Furthermore, the coefficient of *CS* in the magnitude regression is positive, but negative in the original regression. Given the popular interpretation of the *vrp* as the amount that investors *pay* to hedge their volatility risk, the difference in signs is perfectly explainable, but only during those times when the *vrp* is negative. The coefficient estimate of -6.9265 indicates that a 100bps widening of the credit spread results in an increase of approximately 693bps in the cost of hedging such risk (i.e. more negative). This is consistent with our interpretation of the coefficient estimate of 5.1703 in the magnitude regression.[13] Since, under our hypotheses, it is the magnitude of the deviation of realized volatility from implied volatility that represents the market price of volatility, the interpretation is that a 100bps widening of the credit spread results in an increase of approximately 517bps in this price, holding the effects from the other factors constant. However, this intuition falls apart when we consider the times when *vrp* is positive (e.g. December 2007 through January 2008, and August 2008 through October 2008). The coefficient estimate of -6.9265 basically says that a 100bps widening of the credit spread results in a *decrease* in the positive *vrp* of 693bps. This interpretation is not economically justified, especially within the context of the framework we propose. Aside from the R-squared and coefficient signs, analyzing the magnitude of the bias is provides more statistical clarity since allowing the volatility risk premium to change sign introduces additional noise that cannot be explained by the data. This can easily be seen in Table 2, where the standard deviation of *vrp* is uniformly higher than abs(*vrp*) across the entire sample (23% higher) and all subperiods (39% higher during the Crisis).

The other highly significant variable in the magnitude regression is the put option open interest, which is our measure of demand for hedging downside tail risk. For the entire sample period, the estimated coefficient of put option open interest is 6.7138×10^{-7} and is statistically significant at the $< 1\%$ level. The interpretation and economic significance of this result is that, over the entire sample period, a 1 million unit increase in hedging demand results in an increase in the magnitude of the volatility

[13] Also, when credit spreads widen, other variables can change too. Therefore, it might be insightful to also perform marginal regressions on the individual factors. The logic behind marginal regressions is as follows. Suppose $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \epsilon$; when X_1 increases from x_1 to $x_1 + \Delta_1$, a more plausible scenario is for X_2 to change from $\mathrm{E}[X_2|X_1 = x_1]$ to $\mathrm{E}[X_2|X_1 = x_1 + \Delta_1]$ rather than staying the same. If we treat the conditional expectation $\mathrm{E}[X_2|X_1]$ as linear in X_1 (i.e., $\mathrm{E}[X_2|X_1] = a + bX_1$), then the change in y is $\Delta_y = (\beta_1 + \beta_2 b)\Delta_1$ rather than just $\beta_1 \Delta_1$, which is equivalent to the marginal regression of Y on X_1. That is, if we regress Y on X_1 only, then the coefficient before X_1 is exactly $\beta_1 + \beta_2 b$. We ran marginal regressions of this type on all of the factors for the entire sample period and each of the subperiods for robustness. We find that the results are consistent with those reported for the multivariate regressions. Some marginal regression results are reported in the paper.

risk premium of approximately 67bps. This is what can be considered a demand-side effect, holding supply constant. If we were to allow the supply and other factors to change, then the marginal regressions tell a similar story: a 1 million unit increase in hedging demand results in an increase in the magnitude of the volatility risk premium of approximately 57bps.[14] For the entire sample period, the data supports H1, both statistically and economically.

Next, we performed the magnitude regression in Equation (13) for three subperiods – "Pre-Crisis", "Crisis", and "Post-Crisis" – to see if we can identify any patterns that might coincide with the dramatic changes in financial markets as a result of the Financial Crisis of 2007-2009. It is difficult to assign start and end points to financial crises, since they are not as clearly defined as business cycles. To best address this, we use the official NBER recession dates which puts our "Crisis" period from December 2007 to June 2009. The results for the magnitude regressions for the three subperiods are reported in Tables 6, 7, and 8, for "Pre-Crisis", "Crisis", and "Post-Crisis", respectively.

An initial comparison of the results in Tables 6, 7, and 8 with those in Table 5 reveals two preliminary observations. First, there is considerable variation in the explanatory power of the factors over the three subperiods. The R-squared in the "Pre-Crisis" subperiod is only slightly lower than the R-squared for the whole sample period; the R-squared for the "Crisis" subperiod is higher than the whole sample period; and the R-squared for the "Post-Crisis" subperiod is lower than the whole sample period. We might, therefore, interpret the R-squared for the whole sample period as representative of the average explanatory power of our factors with respect to the magnitude of the volatility risk premium. We also observe that, over the entire sample period, the most significant explanatory variables are credit spread and put option open interest (which serves as compelling evidence in favor of the hypotheses H1 and H2); and there seems to be a trend that these two factors become increasingly more significant over time. In the "Pre-Crisis" subperiod, neither the credit spread

[14] The coefficient on the *POI* marginal regression is 5.6972×10^{-7} and is significant at the $< 1\%$ level.

nor put option open interest are significant, but the *TED* spread is the sole significant explanatory variable (at the $< 1\%$ level). During the "Crisis" subperiod, the credit spread is highly significant (at the $< 1\%$ level) and open interest is significant at the 10% level. In the "Post-Crisis" subperiod, both credit spread and put option open interest are highly significant (at the $< 1\%$ level), and the *TED* spread reappears as minimally significant ($< 10\%$ level). When examined together, the results render an interesting story that supports our hypothesis that the magnitude of the volatility risk premium (i.e. the bias between realized and implied volatilities) represents the price that options dealers require to provide liquidity and investors pay to hedge their tail risk. In fact, the trend seems to indicate that the increase in high-frequency data will allow for more precise and accurate measurement of the volatility risk premium and that going forward the results will provide even stronger support for this data-driven analysis of the priced bias in options markets.

During the "Crisis" subperiod, the R-squared of the magnitude regression indicates that the factors are able to explain more than 63% of the variation in the magnitude component of the volatility risk premium. The most significant explanatory variable is the credit spread, which not only provides validation for the risk aversion interpretation of the volatility risk premium, but also supports the hypothesis that the magnitude of this bias reflects the willingness of market makers to absorb inventory and take risk onto their balance sheet. This is what we refer to as the supply-side effect.

The supply-side effect is, holding demand for hedging tail risk constant, if market makers are less willing to take on additional risk then as a result the price for volatility risk will increase. There is evidence of this through the credit spread variable. In addition to being used in the literature as a proxy for risk aversion, more recently some papers have suggested that the credit spread represents "global risk appetite" (Bekaert et al., 2009, 2011). Professional traders view the magnitude of the volatility risk premium as a reflection of the risk tolerance of market makers, along the lines of this "risk appetite" interpretation. We further believe that credit spreads, or the difference between yields on speculative and investment grade debt, captures a related supply-side effect that is present for the entire sample period and seems to dominate during the "Crisis" subperiod. When a financial institution is concerned with the risk on its balance sheet, one way to reduce the overall risk exposure is to *de-leverage*. De-leveraging can be achieved through the right-hand-side of the balance sheet by altering the capital structure: buying back debt, issuing equity, or both. However, there is evidence that large financial institutions have a preference to de-leverage through the left-hand-side of the balance sheet: i.e. reducing its holding of risky assets (Adrian and Shin, 2010). Bai and Collin-Dufresne (2011) show that the credit spread actually picks up this effect quite well, particularly during the 2007- 2009 Financial Crisis. They present evidence that the large financial institutions classified as primary dealers in the credit markets sold off their holdings of risky corporate debt which would have exerted downward pressure on speculative grade bond prices and increases in yields relative to investment grade debt.

This de-leveraging justifies the dominant impact of credit spreads in explaining the volatility risk premium during the "Crisis" subperiod in our results, since there is a high degree of overlap in the set of institutions that serve as primary dealers in the credit markets and those that are market makers in index options.[15] During the "Crisis" subperiod, recall that the most statistically significant explanatory variable was the credit spread (see Table 7). The economic significance of the coefficient estimate is that for a 100bps widening of the credit spread we would expect a 624bps increase in the price of volatility risk. As can be seen from Figure 2, credit spreads increased dramatically at the end of 2008. While this can certainly be viewed as an increase in risk aversion or decrease in global risk appetite, it is also reflective of the massive de-leveraging that occurred after the failure of Lehman Brothers. As large financial institutions reduced their holdings of speculative grade debt they were also reluctant to take on additional risk in other markets. It is reasonable to conclude that during this time dealers in index options increased the price at which they were willing to make a market for hedging downside tail risk, thus increasing the magnitude

[15]Compare the Federal Reserve Bank of New York's list of primary dealers at `http://www.newyorkfed.org/markets/pridealers_current.html` with the members of the Options Clearing Corporation that are dealers of index options at `http://www.optionsclearing.com/membership/member-information/`.

of the volatility risk premium during this time. Because of the inherent leverage in option positions, it is not surprising to see a multiplier effect to the order of 5 to 7 times that of what occurs in the credit markets. This supply-side effect was clearly the driving force behind the increase in the magnitude of the volatility risk premium, not only because it is the only statistically significant explanatory variable during the "Crisis" subperiod, but also because there were no sharp increases in demand for put options on SPRDs as can be seen in Figure 3 where open interest fluctuated between 5,000,000 and 10,000,000 when the range for the entire sample period is 2,000,000 to nearly 16,000,000; this can also be seen in Table 2 by comparing the standard deviation for the put open interest variable across subperiods. While it may seem curious that during the most uncertain point in the financial crisis investors were not aggressively trying to hedge their downside tail risk, but there is an intuitive explanation for this: the so-called "Fed put". After the failure of Lehman Brothers and the subsequent bailout of the financial industry, it became clear that the federal government would provide a backstop either implicitly or explicitly. Therefore, there was little need for investors to pay the high price to hedge their downside tail risk. After the financial crisis, however, demand for hedging downside tail risk returned (see Figure 3) and, indeed, put option open interest was highly significant in the "Post-Crisis" subperiod (see Table 8). This is in addition to the credit spread which still represents the supply-side effect.

Of course, we can also examine what impact this supply-side effect would have on the magnitude of the volatility risk premium if we allow the demand-side effect and other factors to change with it. The marginal regressions indicate that a 100bps increase in the *CS* results in a 550bps increase in the price of volatility risk over the entire sample period.[16]

Lastly, there is the "Post-Crisis" subperiod. Although the R-squared (27.75%) drops off substantially compared to either of the other subperiods or the entire sample period, first note that both the credit spread and the put option open interest are highly significant. The economic interpretation of the coefficients is consistent with

[16]The coefficient on the *CS* marginal regression is 5.5003 and is significant at the $< 1\%$ level.

the findings from the whole sample period and supports our hypotheses. The R-squared is interesting since it indicates perhaps a structural change after the financial crisis. Since the R-squared measures the proportion of the variation in the y-variable (here the magnitude of the *vrp*) that is explained by the x-variables (the risk factors and economic variables) it makes sense to simply look at the descriptive statistics to see if any patterns emerge to which we might attribute a structural change. In the "Post-Crisis" subperiod, both the *vrp* and its magnitude have less variation (standard deviations of 4.19 and 3.52, respectively) than the entire sample period (standard deviations of 6.91 and 5.63, respectively), so a possible explanation is that some factor(s) lost relatively more of their variability and became less correlated with the *vrp*. We know that early on in the sample period, the *TED* spread appeared to have good explanatory power with respect to the *vrp* and even more so for the magnitude, but that the significance seems to disappear over time. We also see that the mean and standard deviation of the *TED* spread become very small in the "Post-Crisis" subperiod (mean of 38bps and standard deviation of 17bps). The *TED* spread was our proxy for liquidity risk and financial market stability, which has a secondary or tertiary effect when analyzing the regression results. However, it is possible that after the Crisis, the low *TED* spread, with its minimal variation, ceased to be a good proxy for liquidity risk and financial stability. We also saw that over time the put option open interest, our proxy of demand for hedging tail risk, seemed to become increasingly more important. While the mean put option open interest increased over time (4.13mm, 6.59mm, and 10.51mm, for the "Pre-Crisis", "Crisis, and "Post-Crisis" subperiods, respectively), which could just be indicative of the growing market for index options and options on ETF's, its standard deviation falls during the "Crisis" subperiod (from 1.57 to 0.92) and then more than doubles in the "Post-Crisis" period (2.05). It is possible that because of the government's implicit and explicit backstop – the so-called "Fed put" – made demand play less of a role during the Crisis, but as we emerged, new demand for hedging tail risk came to be a key driver in explaining the market price of volatility risk. Therefore, our structural change could be the increased role of put option open interest and investors' demand for hedging tail risk rather than liquidity and overall financial stability as being a key factor. Additionally, since the quality and quantity of data increased over this period, it might also suggest the additional economic insight that can be obtained from in-depth analysis of large scale financial data.

Wu (2011) shows that the absolute value of the variance risk premium is proportional to the level of volatility in the market. Therefore, the variance risk premium (and consequently the volatility risk premium) is either very negative or very positive when volatility levels are high. This provides added justification for examining the absolute value of the volatility risk premium to make inferences about the price of volatility risk. More specifically, it suggests that uncertainty in the market increases the bias between realized and implied volatility and the price that must be paid to hedge this risk. In this section we were able to provide empirical evidence explicitly linking this to demand for hedging tail risk and liquidity provision to the volatility market (hypotheses H1 and H2). Next we examine the direction, or sign, of the volatility risk premium.

5.3 Sign Tests

While a negative volatility risk premium is justified theoretically and for the most part supported by the data, the large positive spike in *vrp* in the middle of the Financial Crisis as well as several other positive spikes throughout the entire sample period provides a paradox. To help reconcile this paradox within our price bias interpretation, we now examine the information in the direction, or sign, of the bias as proposed by Hypotheses H3 and H4. Recall, H3 echos the view of some derivative traders that the direction of the volatility risk premium reflects the market's expectation of future changes in volatility. When the *vrp* is negative, then the realized volatility in the equity market is less than the implied volatility extracted from option prices. Volatility is priced higher in the forward-looking options market indicating that market participants expect realized volatility to increase in the future. When the *vrp* is positive, then the realized volatility in the equity market is greater than the implied volatility extracted from option prices. Volatility is priced lower in the forward-looking options market indicating that market participants expect realized volatility to decrease in the future.

We seek to test Hypothesis H3 using a modified version of the regression proposed in Aït-Sahalia et al. (2012) to test the Expectation Hypothesis. First, to establish a baseline we run the following specification of the Expectation Hypothesis:

$$\overline{RV}_t = \alpha + \beta_1 VIX_{t-21} + \epsilon_t. \tag{14}$$

Here, the y-variable is the average ex-post realized volatility using the frequency domain estimation methodology for the current 21 trading-day period; the x-variable is the 21 trading-day *lagged VIX*. The idea is to see whether implied volatility doesconvey information about the market's expectations about the future realized volatility levels. If implied volatilities are unbiased and efficient estimates of future realized volatility, then we could use the *VIX* index to predict what the future level of realized volatility will be one month in the future. This is the essence of the Expectation Hypothesis.

The results for this regression can be found in Table 9. The coefficient β_1 is significant at the $< 1\%$ level and the R-squared indicates that the lagged *VIX* is able to explain 32.56% of the future realized volatility. This suggests that implied volatility does have some predictive power for realized volatility, or in other words, it represents to some extent the market's expectation about future realized volatility. We note that the t-statistic and p-value associated with β_1 just tells us that the coefficient is statistically different from zero. It is easy to show that β_1 is also statistically less than 1.[17] The implication is that given the current level of *VIX*, to predict the future realized volatility you would first discount the level of *VIX* (by approximately 0.45) and then add a constant ($\alpha = 5.28\%$). Furthermore, the results in Table 9 indicate that for $VIX > 9.629$, implied volatility tends to *overestimate* future realized volatility; this implies that, except when *VIX* is very low, there should be a *negative* volatility risk premium. However, we know from examining our time series, that the

[17]Construct the new test statistic $\hat{t} = \frac{\beta_1 - 1}{se_{\beta_1}} = \frac{0.45133-1}{0.06634} = -8.27$.

vrp was positive when *VIX* was approaching historical highs. Therefore, perhaps a simple linear model such as in Equation (14) does not tell the entire story about the Expectations Hypothesis when it comes to volatilities.

To test hypothesis H3, we introduce a binary variable, sgn(*vrp*) which equals -1 if $vrp < 0$ and +1 if $vrp > 0$ and run the modified regression specified as:

$$\overline{RV}_t = \alpha + \beta_1 VIX_{t-21} + \beta_2 \text{sgn}(vrp_{t-21}) + \epsilon_t. \tag{15}$$

This attempts to identify whether or not the direction, or the sign, of the volatility risk premium provides additional information about future levels of realized volatility. The results in Table 10 indicate that including the sign of the *vrp* improves the explanatory power as the adjusted R-squared is 54.7% with everything – α, β_1, β_2 – significant at the < 1% level. Note that β_1 is still statistically less than 1, but the relationship between the lagged level of *VIX* and the future level of realized volatility is no longer linear.

The results can be used to predict next month's realized volatility in terms of the sign of the current volatility risk premium and conditional on the current level on *VIX*. Suppose the *VIX* is currently 21 (roughly the median value for our entire sample period). Then our prediction for next month's realized volatility level depends on whether the *vrp* is currently positive or negative. If the *vrp* is negative, then the results of the regression specified by Equation (15) predicts next month's realized volatility to be approximately 13.41%; and, thus, when the *vrp* is negative, implied volatility *overestimates* expected future realized volatility. However, if the *vrp* is positive, the forecast changes to 29.95% and now implied volatility *underestimates* expected future realized volatility. This appears to confirm hypothesis H3; similar analysis of higher and lower *VIX* values (e.g., using 30 and 13, roughly the median value for the "Pre-Crisis" and "Crisis" subperiods, respectively) leads to the same conclusion.

These results should be met with some degree of skepticism. Aït-Sahalia et al. (2012) show that as the forecasting horizon increases, the parameter estimates of the Expectation Hypothesis regression become biased and inefficient. So, while our results do perhaps provide some evidence in favor of hypothesis H3, the statistical inferences may not be sound. Furthermore, it does not say much about the economic meaning of a positive or negative volatility risk premium.

Alternatively, Bakshi and Kapadia (2003) provide an explanation that is more consistent with our motivating theme about market-making and intermediation in the options market. They show, both theoretically and with empirical evidence on index options, that a negative volatility risk premium is representative of the underperformance of a delta-neutral portfolio, where the trader sells calls and purchases Δ units of the underlying as a hedge (or sells puts and short sells Δ units of the underlying as a hedge). Since we are examining the *vrp* in terms of market makers who provide liquidity to investors that wish to hedge downside tail risk with put options on the market (the S&P 500 index or SPDR ETF), the market maker must short sell the underlying (e.g. SPDRs) in order to maintain delta-neutrality. Consequently, the market maker will have a gain on the delta-hedge when the S&P 500 is down and a loss on the delta-hedge when the S&P 500 is up. Table 11 shows the annualized returns on the S&P 500 index when the volatility risk premium is positive (Panel A) and negative (Panel B). We can see that for every period that the *vrp* is positive, the S&P 500 has negative returns. This is consistent that, in the less frequent instances when the *vrp* is positive, traders making a market in *SPY* put options have a profit on their delta-neutral hedge. More often than not, when the *vrp* is negative, traders making a market in *SPY* put options are losing money on their delta-neutral hedge as evidenced by the majority of periods that show positive returns on the S&P500, as well as the overall average annualized return of 10.99% when the *vrp* is negative. This suggests strong evidence in favor of H4.

In order to give more econometric rigor and statistical significance of this relationship posited by hypothesis H4, we ran the following regression:

$$S\&P\_return_t = \beta_0 + \beta_1 \mathrm{sgn}(vrp_t) + \epsilon_t \tag{16}$$

where the dependent variable, $S\&P\_return_t$, is the annualized daily return on the S&P500 index and on day t and the independent variable, $\text{sgn}(vrp_t)$, is the sign of the contemporaneous volatility risk premium; The results for the regression specified by Equation (16) can be found in Table 12. The coefficient estimate for β_1 is negative and significant at the $< 5\%$ level. We can therefore conclude from this regression that returns on the S&P 500 index statistically depend on whether the volatility risk premium is positive or negative. Furthermore, the negative coefficient estimate supports the delta-hedged gains argument of Bakshi and Kapadia (2003), but within the context of put options on the market: when the *vrp* is positive, returns on the S&P 500 (or SPDRs) can be expected to be negative (and the delta hedge of being short the underlying will make money), whereas when the *vrp* is negative, returns on the S&P 500 (or SPDRs) can be expected to be positive (and the delta hedge of being short the underlying will lose money).

In sum, we find that there is evidence, albeit weak, in favor of H3. We find more economic significance supporting H4. There is also very strong statistical and economic relationships found in the data to support H1 and H2, and the marginal regression results confirm that these relationships still hold and need not be mutually exclusive.

6 Conclusion

"Big Data" has the potential to transform research in many areas, including financial economics and risk analysis. We use a massive data set, collected from numerous sources, to perform a unique study of how the market prices volatility. Our research questions equate the volatility risk premium to a systematically priced bias between ex-post realized volatility and ex-ante expected volatility implied by options. Unlike most other studies of the volatility risk premium, rather than start with a theoretical model of volatility, we begin with intensive data-driven methods leveraging the insights of Big Data analytics.

First, we collected price and volume data on every transaction in SPDRs, the ETF that tracks the S&P 500 index, over a five year period from 2006 through 2011 yielding over half a billion observations. We then use a novel technique to estimate the integrated volatility using the processed ultra-high-frequency data in the frequency domain. This methodology allows us to distinguish the true volatility of the price process and the microstructure noise, even when the noise is correlated over time. The result is a consistent, de-biased estimate of the integrated volatility as our measure of realized volatility. In constructing the time series of the volatility risk premium, we smooth the daily realized volatility by taking a 21-trading-day rolling average and subtract from it the daily value of *VIX* volatility index for the same month.

Insofar as the option implied volatility represents the market's expectation of future volatility, this formulation of the volatility risk premium is very much like a statistical bias. We decompose this bias into magnitude (absolute value) and direction (sign) and analyze them separately. Based on stylized facts about the volatility risk premium, we construct four testable hypotheses about its economic meaning and determinants. The general theme is that the volatility risk premium cannot be explained by traditional risk factors, but rather are related to supply and demand forces in option markets and the role of market makers in providing liquidity to investors who seek to hedge their downside tail risk. This is all viewed within the lens of the volatility risk premium being systematically priced bias.

The results indicate that the size of this bias represents the price that market makers require to meet the demand of investors who wish to hedge their downside tail risk and compensates for supply and demand imbalances in this market. In fact, we find compelling evidence that during the Financial Crisis, supply-side forces dominated as financial intermediaries shed risky positions and were reluctant to take more risk onto their balance sheets. Demand-side forces dried up as the implicit guarantees and "Fed put" made hedging tail risk less attractive for investors. This is reflected in the highly significant credit spread, which reflects market makers deleveraging during the Crisis, and reduced significant in put option open interest, which is our proxy for investors' demand for hedging tail risk.

Practitioners view the sign of the volatility risk premium, on the other hand, to the market's expectation about future levels of volatility. This is similar to the Expectation Hypothesis discussed in Aït-Sahalia et al. (2012) and, while we are able to find some evidence in favor of this hypothesis in the data, statistical issues raise doubt on the validity of the inferences and there is no clear economic interpretation within the conceptual framework we established. An alternative hypothesis links the sign of the volatility risk premium to the gains and losses on traders' delta-hedged positions when making a market for index options. Bakshi and Kapadia (2003) were the first to propose this interpretation, and we are able to find fairly conclusive evidence in favor of it in the market for S&P 500 put options. That is, market makers provide liquidity to investors seeking to hedge their downside tail risk – via put options on SPDRs – will delta-hedge these positions by shorting shares of the underlying ETF. We find that returns on the S&P 500 are negative over all consecutive trading days where the volatility risk premium is positive, indicating a delta-hedged gain for the market maker. We find that the returns on the S&P 500 are positive over all but two series of consecutive trading days where the volatility risk premium is negative, indicating a delta-hedged loss for the market maker.

Overall, the ability of our data-driven analysis to identify economic insights into how the market prices volatility is very encouraging for researchers interested in using similar approaches for other quantitative studies in financial economics and risk analysis. While we do utilize the results from the existing literature along with economic intuition to highlight some stylized facts about the volatility risk premium and come up with testable hypotheses, our analysis does not rely on any specific theoretical model and has minimal parametric assumptions. The trend over our sample period seems to indicate that the increase in high-frequency data will allow for more precise and accurate measurement of the volatility risk premium as a systematically priced bias. This will then allow for even better identification of the determinants of the volatility risk premium and highlight role that intermediation in the market for volatility and hedging downside risk as well as the role that supply and demand imbalances play in driving the deviation between realized and implied volatilities.

Appendix

A The Fourier Transform Method

A.1 Frequency Domain Representation

First define the discrete Fourier transform of the increment process $\Delta U_{t_j} = U_{t_{j+1}} - U_{t_j}$ of a sample from a generic time series $U_{t_j}, j = 1, \cdots, N$,

$$J_k^{(U)} = \sqrt{\frac{1}{N}} \sum_{j=1}^{N} \Delta U_{t_j} e^{-2\pi i t_j f_k}, f_k = \frac{k}{T}. \tag{17}$$

Assume that the latent true (log)price X_t follows an Ito process as in Equation (4). To simplify the notation, we assume the drift term is zero as it does not affect the asymptotic behaviour (for a more complete version, see Olhede et al. (2009)). The frequency domain estimator uses the fact that the integrated volatility can be written in terms of the variance of $J_k^{(X)}$. It can be shown that

$$\int_0^T \mathrm{E}\{\sigma_s^2\} ds = \sum_{k=0}^{N-1} \mathrm{E}|J_k^{(X)}|^2 + O(\Delta t). \tag{18}$$

However, what we observe is the transaction price Y_t at times $\{t_i\} \in [0, T]$ as in Equation (5), so at each frequency there is a noise contribution,

$$\sum_{k=0}^{N-1} \mathrm{E}|J_k^{(Y)}|^2 = \sum_{k=0}^{N-1} \left(\mathrm{E}|J_k^{(X)}|^2 + a^2 |2 \sin(\pi f_k \Delta t)|^2 \right), \tag{19}$$

where for now we assume ϵ_t is a white noise process with variance a^2.

A.2 The De-biased Estimator

The frequency domain representation gives us a nice way to disentangle the the microstructure noise. If we could shrink by

$$L_k = \frac{\mathrm{E}|J_k^{(X)}|^2}{\mathrm{E}|J_k^{(X)}|^2 + a^2|2\sin(\pi f_k \Delta t)|^2} \tag{20}$$

at each frequency, an oracle estimator would be $\widehat{\langle X, X\rangle}_T^{(L_k)} = \sum_{k=0}^{N-1} L_k |J_k^{(Y)}|^2$. It remains the task to estimate the multiscale ratio L_k. The unknown quantities in Equation (20) can be estimated by the Whittle log-likelihood,

$$l\left(\sigma_X^2, a^2\right) = -\sum_{k=1}^{N/2-1} \log\left(\sigma_X^2 + a^2|2\sin(\pi f_k \Delta t)|^2\right) - \sum_{k=1}^{N/2-1} \frac{|J_k^Y|^2}{\sigma_X^2 + a^2|2\sin(\pi f_k \Delta t)|^2}, \tag{21}$$

therefore

$$\widehat{L_k} = \frac{\hat{\sigma}_X^2}{\hat{\sigma}_X^2 + \hat{a}^2|2\sin(\pi f_k \Delta t)|^2}. \tag{22}$$

The final de-biased estimator is

$$\widehat{\langle X, X\rangle}_T^{(\widehat{L_k})} = \sum_{k=0}^{N-1} \widehat{L_k} |J_k^{(Y)}|^2. \tag{23}$$

It can be shown that this is a consistent estimator of the integrated volatility,

$$\widehat{\langle X, X\rangle}_T^{(\widehat{L_k})} = \int_0^T \sigma_t^2 \; dt + O_p(\Delta t^{1/4}). \tag{24}$$

A.3 Autocorrelated Noise

If we assume that ϵ_{t_j} is an autocorrelated stationary time series, it is convenient to model it as a moving average process of order q,

$$\epsilon_{t_j} = \eta_{t_j} + \sum_{k=1}^{q} \theta_k \eta_{t_{j-k}}, \tag{25}$$

where $\{\eta_{t_j}\}$ is a white noise process with variance σ_η^2. This MA(q) specification leads to a new likelihood function

$$\begin{aligned} & l\left(\sigma_X^2, \sigma_\eta^2, \{\theta_k\}_{k=1}^q\right) \\ = \; & -\sum_{k=1}^{N/2-1} \log\left(\sigma_X^2 + \sigma_\eta^2 |1 + \sum_{k=1}^{q} \theta_k e^{2i\pi f k}|^2\right) - \sum_{k=1}^{N/2-1} \frac{|J_k^Y|^2}{\sigma_X^2 + \sigma_\eta^2 |1 + \sum_{k=1}^q \theta_k e^{2i\pi f k}|^2}, \end{aligned} \tag{26}$$

and therefore the multiscale ratio is defined as

$$\widehat{L_k} = \frac{\hat{\sigma}_X^2}{\hat{\sigma}_X^2 + \hat{\sigma}_\eta^2 |1 + \sum_{k=1}^q \hat{\theta}_k e^{2i\pi f k}|^2}. \tag{27}$$

To determine the order q in Equation (25), we minimize the corrected Akaike information criterion ($AICC$),

$$AICC(q) = -2l\left(\hat{\sigma}_X^2, \hat{\sigma}_\eta^2, \{\hat{\theta}_k\}_{k=1}^q\right) + 2\frac{(q+2)N}{N-q-3}. \tag{28}$$

B Simulations

We compare the performance of Fourier method and naive subsampling at different sampling frequency on simulated data using a Heston (1993) model:

$$\begin{aligned} dX_t &= (\mu - \nu_t/2)dt + \sigma_t dB_t, \\ d\nu_t &= \kappa(\alpha - \nu_t)dt + \gamma \nu_t^{1/2} dW_t, \end{aligned}$$

where $\nu_t = \sigma_t^2$. The parameters are set as follows: $\mu = 0.05$, $\kappa = 5$, $\alpha = 0.04$, $\gamma = 0.5$, and the correlation between the two Brownian motions B_t and W_t is $\rho = -0.5$.[18] The initial values are $X_0 = 0$ and $\nu_0 = 0.04$. We take T as one day, and simulate data with $\Delta_t = 0.1$s, which yields a sample path of length $N = 234,000$ in one trading day. We first calculate the underlying true integrated volatility by a Riemann sum approximation of the integral, i.e.: $\frac{T}{N}\sum_{i=1}^{N}\sigma_i^2 = \int_0^T \sigma_t^2 dt$. Then we add $AR(2)$ noise $\epsilon_i = 0.6\epsilon_{i-1} - 0.4\epsilon_{i-2} + \eta_i$, to get the observed data $Y_i = X_i + \epsilon_i$, whereη_i's are i.i.d. $\mathcal{N}(0, \sigma_\eta^2)$ and we set $\sigma_\eta = 5 \times 10^{-4}$.

We estimate the integrated volatility using two methods, the Fourier method and the naive subsampling, which yields $< X, X >_T^{Fourier}$ and $< X, X >_T^{subsampling}$. We calculate the RMSE (root-mean-square error) of the estimates to the truth over 200 simulated sample paths. To further illustrate the effect of high frequency data, we evaluate two methods from $\Delta_t = 1$s up to $\Delta_t = 150$s. Figure 4 shows the RMSE of the Fourier method and the naive subsampling against decreasing sampling frequencies.

The takeaway of this figure is two folds. First, the Fourier method can effectively filter the correlated microstructure noise, and works better than naive sampling method. We did not implement other more sophisticated methods for comparison, as the simulation is not to illustrate the superiority of the Fourier method, but rather to justify the use of high frequency data. Second, if we can filter the microstructure noise, higher frequency gives us a better estimate as we are able to utilize more data hence more information.

[18]These are the same as those used in the Olhede et al. (2009) simulations.

References

Adrian, T. and Shin, H. S. (2010), “Liquidity and leverage,” *Journal of Financial Intermediation*, 19, 418 – 437.

Aït-Sahalia, Y., Karaman, M., and Mancini, L. (2012), “The Term Structure of Variance Swaps, Risk Premia and the Expectation Hypothesis,” Available at SSRN: http://ssrn.com/abstract=2136820.

Aït-Sahalia, Y. and Mancini, L. (2008), “Out of sample forecasts of quadratic variation,” *Journal of Econometrics*, 147, 17 – 33.

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005), “How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise,” *Review of Financial Studies*, 18, 351–416.

— (2011), “Ultra high frequency volatility estimation with dependent microstructure noise,” *Journal of Econometrics*, 160, 160 – 175.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2004), “Parametric and Nonparametric Volatility Measurement,” in *in L.P. Hansen and Y. Ait-Sahalia (eds.), Handbook of Financial Econometrics*, North-Holland.

Bai, J. and Collin-Dufresne, P. (2011), “The Determinants of the CDS-Bond Basis During the Financial Crisis of 2007-2009,” Available at SSRN: http://ssrn.com/abstract=2024531.

Bakshi, G. and Kapadia, N. (2003), “Delta-Hedged Gains and the Negative Market Volatility Risk Premium,” *Review of Financial Studies*, 16, 527–566.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008), “Designing Realized Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise,” *Econometrica*, 76, 1481–1536.

— (2009), “Realized kernels in practice: trades and quotes,” *Econometrics Journal*, 12, C1–C32.

— (2011), “Subsampling realised kernels,” *Journal of Econometrics*, 160, 204 – 219.

Bekaert, G., Harvey, C. R., Lundblad, C. T., and Siegel, S. (2011), “What Segments Equity Markets?” *Review of Financial Studies*, 24, 3841–3890.

Bekaert, G. and Hoerova, M. (2014), “The VIX, the variance premium and stock market volatility,” *Journal of Econometrics*.

Bekaert, G., Hoerova, M., and Lo Duca, M. (2013), “Risk, uncertainty and monetary policy,” *Journal of Monetary Economics*, 60, 771–788.

Bekaert, G., Hoerova, M., and Scheicher, M. (2009), “What Do Asset Prices Have to Say about Risk Appetite and Uncertainty?” ECB Working Paper No. 1037., Available at SSRN: http://ssrn.com/abstract=1358848.

Black, F. and Scholes, M. (1973), “The pricing of options and corporate liabilities,” *Journal of Political Economy*, 81, 637–654.

Bollerslev, T., Gibson, M., and Zhou, H. (2011), “Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities,” *Journal of Econometrics*, 160, 235 – 245.

Bollerslev, T., Tauchen, G., and Zhou, H. (2009), “Expected Stock Returns and Variance Risk Premia,” *Review of Financial Studies*, 22, 4463–4492.

Britten-Jones, M. and Neuberger, A. (2000), “Option Prices, Implied Price Processes, and Stochastic Volatility,” *The Journal of Finance*, 55, 839–866.

Brunnermeier, M. K. (2009), “Deciphering the Liquidity and Credit Crunch 2007-2008,” *Journal of Economic Perspectives*, 23, pp. 77–100.

Canina, L. and Figlewski, S. (1993), “The informational content of implied volatility,” *Review of Financial studies*, 6, 659–681.

Cao, J. and Han, B. (2013), “Cross section of option returns and idiosyncratic stock volatility,” *Journal of Financial Economics*, 108, 231–249.

Carr, P. and Wu, L. (2006), “A Tale of Two Indices,” *The Journal of Derivatives*, 13, 13–29.

— (2009), “Variance Risk Premiums,” *Review of Financial Studies*, 22, 1311–1341.

CBOE (2009), “The CBOE Volatility Index - VIX,” White paper, Chicago Board Options Exchange, Chicago, IL.

Chen, N.-F., Roll, R., and Ross, S. A. (1986), “Economic Forces and the Stock Market,” *The Journal of Business*, 59, pp. 383–403.

Christensen, B. J. and Prabhala, N. R. (1998), “The relation between implied and realized volatility,” *Journal of Financial Economics*, 50, 125–150.

Demeterfi, K., Derman, E., Kamal, M., and Zou, J. (1999), “A guide to volatility and variance swaps,” *The Journal of Derivatives*, 6, 9–32.

Derman, E. and Kani, I. (1994), “Riding on a smile,” *Risk*, 7, 32–39.

Drechsler, I. and Yaron, A. (2011), “What’s Vol Got to Do with It,” *Review of Financial Studies*, 24, 1–45.

Dupire, B. (1994), “Pricing with a smile,” *Risk*, 7, 18–20.

Egloff, D., Leippold, M., and Wu, L. (2010), “The term structure of variance swap rates and optimal variance swap investments,” *Journal of Financial and Quantitative Analysis*, 45, 1279–1310.

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

Fan, J. (2013), “Features of big data and sparsest solution in high confidence set,” *Past, Present, and Future of Statistical Science*, 507–523.

Fan, J., Han, F., and Liu, H. (2014), “Challenges of Big Data analysis,” *National Science Review*, 1, 293–314.

Fan, J. and Wang, Y. (2007), "Multi-Scale Jump and Volatility Analysis for High-Frequency Financial Data," *Journal of the American Statistical Association*, 102, 1349–1362.

Gârleanu, N., Pedersen, L. H., and Poteshman, A. M. (2009), "Demand-Based Option Pricing," *Review of Financial Studies*, 22, 4259–4299.

Heston, S. L. (1993), "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *Review of Financial Studies*, 6, 327–343.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009), "Microstructure noise in the continuous case: The pre-averaging approach," *Stochastic Processes and their Applications*, 119, 2249 – 2276.

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

— (2007), "Extracting Model-Free Volatility from Option Prices," *The Journal of Derivatives*, 14, 35–60.

Kelly, B. T., Lustig, H., and Van Nieuwerburgh, S. (2012), "Too-systemic-to-fail: What option markets imply about sector-wide government guarantees," Fama-Miller Working Paper; Chicago Booth Research Paper No. 11-12. Available at SSRN: http://ssrn.com/abstract=1762312.

Merton, R. C. (1973), "Theory of rational option pricing," *The Bell Journal of Economics and Management Science*, 4, 141–183.

Mixon, S. (2009), "Option markets and implied volatility: Past versus present," *Journal of Financial Economics*, 94, 171–191.

Nagel, S. (2012), "Evaporating Liquidity," *Review of Financial Studies*, 25, 2005–2039.

Olhede, S., Sykulski, A., and Pavliotis, G. (2009), "Frequency Domain Estimation of Integrated Volatility for Ito Processes in the Presence of Market-Microstructure Noise," *SIAM - Multiscale Modeling and Simulation*, 8, 393–427.

Rubinstein, M. (1994), "Implied binomial trees," *The Journal of Finance*, 49, 771–818.

Todorov, V. (2010), "Variance Risk-Premium Dynamics: The Role of Jumps," *Review of Financial Studies*, 23, 345–383.

Wu, L. (2011), "Variance dynamics: Joint evidence from options and high-frequency returns," *Journal of Econometrics*, 160, 280 – 287.

Xiu, D. (2010), "Quasi-maximum likelihood estimation of volatility with high frequency data," *Journal of Econometrics*, 159, 235 – 250.

Zhang, L. (2006), "Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach," *Bernoulli*, 12, pp. 1019–1043.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, pp. 1394–1411.

Zhou, B. (1996), "High-Frequency Data and Volatility in Foreign-Exchange Rates," *Journal of Business & Economic Statistics*, 14, pp. 45–52.

Zhou, H. (2011), "Variance Risk Premia, Asset Predictability Puzzles, and Macroeconomic Uncertainty," Available at SSRN: http://ssrn.com/abstract=1783624.

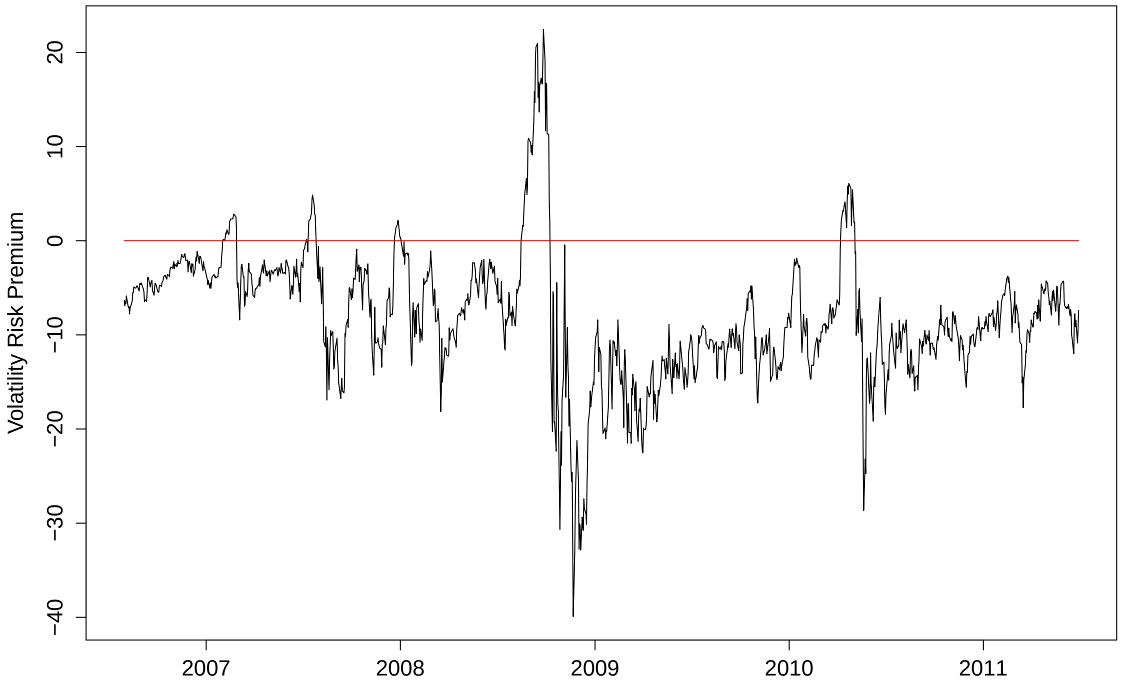

Figure 1: Time Series of the Volatility Risk Premium

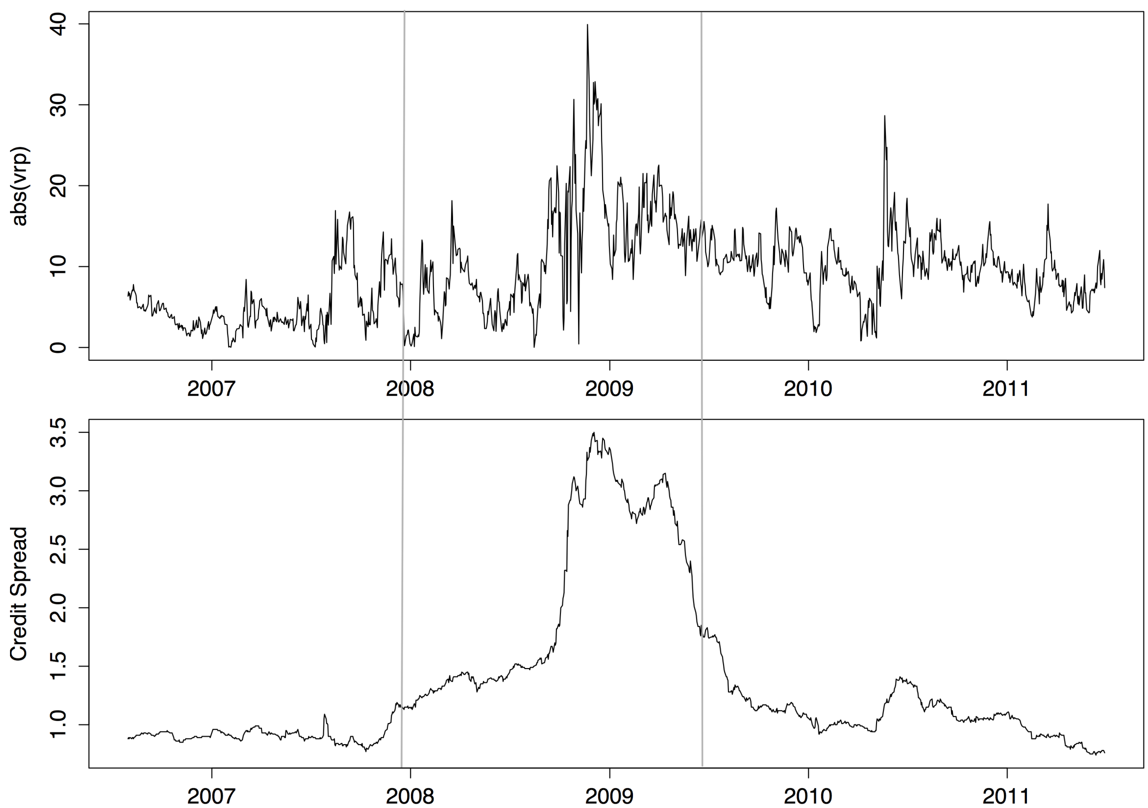

Figure 2:
Top: Time series of the absolute value of the volatility risk premium or the *magnitude component* of the *vrp*
Bottom: Time series of the credit spread, defined as the yield on Baa-rated and Aaa-rated corporate debt.
Note: The vertical lines separate the three subperiods.

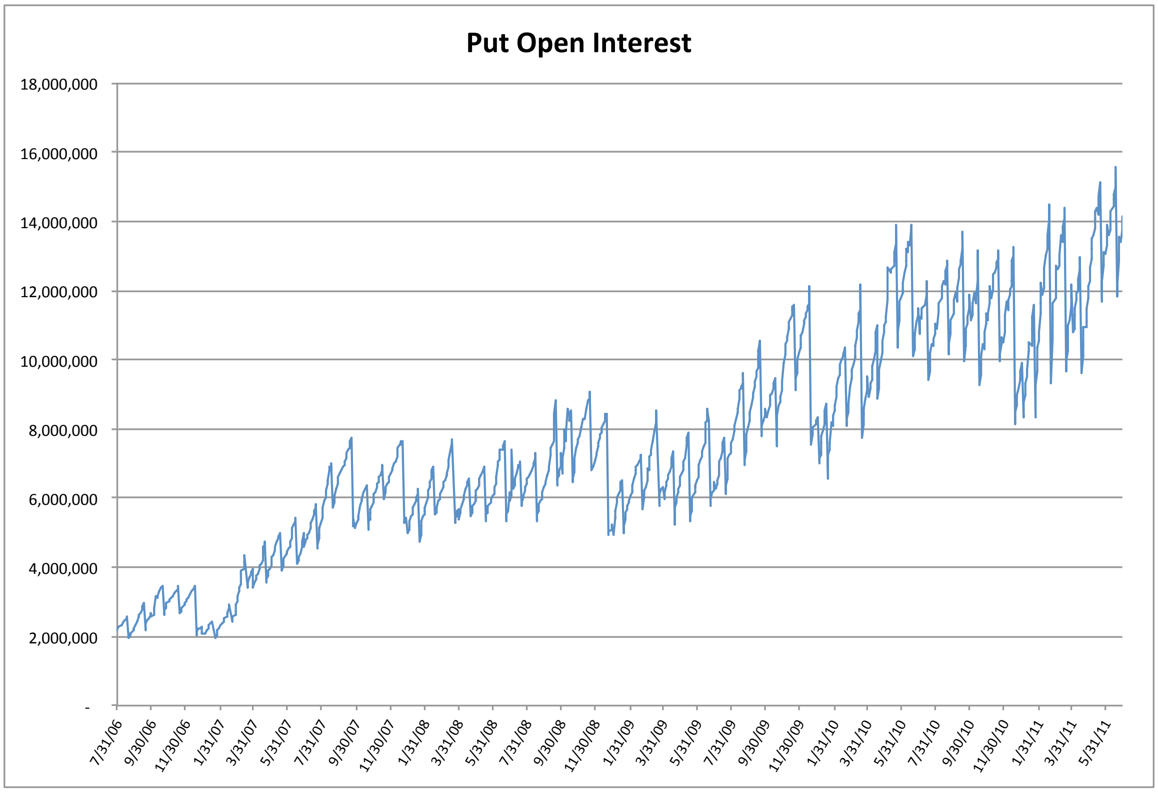

Figure 3: Open interest for put options on the S&P500 index over the sample period. Put open interest is our proxy for investor demand for hedging downside tail risk.

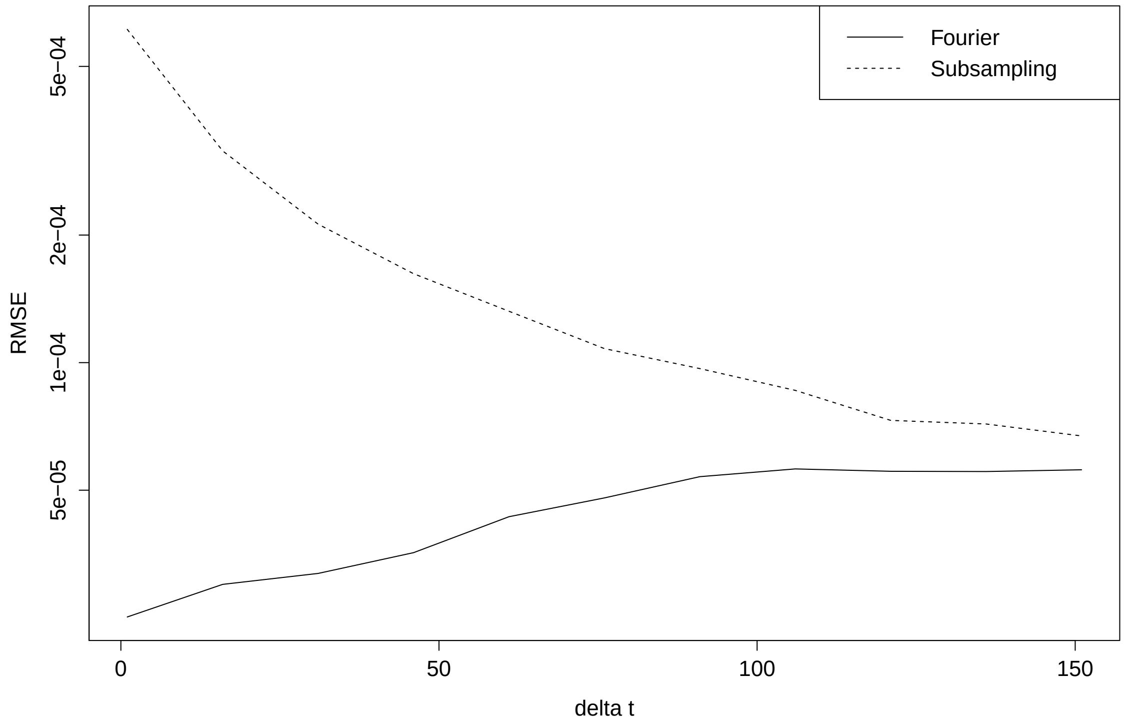

Figure 4: RMSE (shown in logarithmic scale) of two methods

Variable Name	Description
$r_m - r_f$ (*Mkt*)	Fama-French market risk factor; market return minus risk-free rate
SMB	Fama-French size factor
HML	Fama-French value factor
Credit Spread (*CS*)	Difference in yield on Baa-rated and Aaa-rated corporate debt
TED Spread (*TED*)	Difference between 3-month Eurodollar rate and 3-month Treasury rate
Put Open Interest (*POI*)	Daily open interest for put options on *SPY*

Table 1: Explanatory variables

	Whole-Sample (N = 1238)			Pre-Crisis (N = 338)			Crisis (N = 375)			Post-Crisis (N = 525)		
	Mean	Med	Std	Mean	Med	Std	Mean	Med	Std	Mean	Med	Std
vrp	-8.10	-8.32	6.91	-4.53	-3.86	3.82	-9.17	-9.05	10.20	-9.63	-10.00	4.19
abs(*vrp*)	9.03	8.68	5.63	4.86	3.88	3.39	11.58	10.58	7.34	9.90	10.00	3.52
VIX	24.20	21.88	11.70	15.56	13.54	5.16	34.58	28.19	14.68	22.34	21.84	4.86
Mkt	0.02	0.10	1.60	0.04	0.11	0.86	-0.09	-0.01	2.45	0.09	0.14	1.10
SMB	0.016	0.02	0.64	-0.02	-0.04	0.42	-0.02	-0.03	0.88	0.03	0.05	0.54
HML	-0.007	-0.01	0.73	-0.03	-0.03	0.26	-0.004	0.01	1.13	0.007	-0.01	0.55
CS	1.35	1.09	0.68	0.90	0.90	0.05	2.07	1.56	0.80	1.12	1.07	0.25
TED	0.87	0.44	0.88	0.61	0.40	0.51	1.78	1.38	1.03	0.38	0.33	0.17
POI($\times 10^6$)	7.58	7.04	3.15	4.13	3.86	1.57	6.59	6.50	0.92	10.51	10.65	2.05

Table 2: Descriptive statistics for all variables across all periods.

Variable	Estimate	Std Error	t-value	Pr(> $\|t\|$)	Significance
(Intercept)	-8.074632	0.329221	-24.5265	< 2.2e-16	***
Mkt	-0.876107	0.190903	-4.5893	4.902e-06	***
SMB	0.049744	0.365566	0.1361	0.8918	
HML	1.008269	0.478214	2.1084	0.0352	**
Adjusted R^2	0.02711				
F-statistic	12.49 on 3 and 1234 DF				
p-value	$4.759e - 08$				

Table 3: Regression of *vrp* on Fama-French risk factors. Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are < 1%, < 5%, and < 10% for ***, **, and *, respectively.

Variable	Estimate	Std Error	t-value	Pr(> $\|t\|$)	Significance
(Intercept)	4.2742e+00	1.4152e+00	3.0201	0.002579	***
Mkt	-7.4433e-01	1.5077e-01	-4.9368	9.033e-07	***
SMB	2.3838e-01	2.4281e-01	0.9817	0.326418	
HML	7.1323e-01	3.7520e-01	1.9009	0.057546	*
CS	-6.9265e+00	1.2609e+00	-5.4935	4.785e-08	***
TED	2.0051e+00	1.4945e+00	1.3416	0.179965	
POI	-6.2710e-07	9.9694e-08	-6.2902	4.396e-10	***
Adjusted R^2	0.4155				
F-statistic	147.6 on 6 and 1231 DF				
p-value	$< 2.2e - 16$				

Table 4: Regression of *vrp* on explanatory variables: Whole Sample Period. Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are < 1%, < 5%, and < 10% for ***, **, and *, respectively.

Variable	Estimate	Std Error	t-value	Pr(> $\|t\|$)	Significance
(Intercept)	-3.6404e+00	1.0045e+00	-3.6239	0.000302	***
Mkt	4.2400e-01	9.4296e-02	4.4965	7.560e-06	***
SMB	-1.1227e-01	1.5503e-01	-0.7242	0.469092	
HML	-1.3273e-01	1.4148e-01	-0.9381	0.348354	
CS	5.1703e+00	6.4763e-01	7.9833	3.243e-15	***
TED	6.8857e-01	5.3153e-01	1.2955	0.195408	
POI	6.7138e-07	8.3925e-08	7.9998	2.856e-15	***
Adjusted R^2	0.5832				
F-statistic	289.5 on 6 and 1231 DF				
p-value	$< 2.2e-16$				

Table 5: Regression of $|vrp|$ on explanatory variables: Whole Sample Period. Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are $< 1\%$, $< 5\%$, and $< 10\%$ for ***, **, and *, respectively.

Variable	Estimate	Std Error	t-value	Pr(> $\|t\|$)	Significance
(Intercept)	-4.7395e+00	7.4185e+00	-0.6389	0.5233	
Mkt	1.5573e-01	1.5360e-01	1.0139	0.3114	
SMB	1.6859e-01	3.2744e-01	0.5149	0.6070	
HML	4.1528e-01	6.1978e-01	0.6700	0.5033	
CS	7.5490e+00	8.0908e+00	0.9330	0.3515	
TED	5.6288e+00	8.7777e-01	6.4126	4.934e-10	***
POI	-1.6247e-07	2.9512e-07	-0.5505	0.5823	
Adjusted R^2	0.5789				
F-statistic	78.23 on 6 and 331 DF				
p-value	$< 2.2e-16$				

Table 6: Regression of $|vrp|$ on explanatory variables: Pre-Crisis Subperiod. Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are $< 1\%$, $< 5\%$, and $< 10\%$ for ***, **, and *, respectively.

Variable	Estimate	Std Error	t-value	Pr(> $\|t\|$)	Significance
(Intercept)	-1.0040e+01	4.0327e+00	-2.4897	0.01322	**
Mkt	4.8153e-01	1.1335e-01	4.2481	2.735e-05	***
SMB	1.4010e-01	1.7774e-01	0.7883	0.43105	
HML	-1.6113e-01	1.8898e-01	-0.8527	0.39440	
CS	6.2419e+00	7.3952e-01	8.4405	7.350e-16	***
TED	8.9576e-01	6.6878e-01	1.3394	0.18127	
POI	1.0826e-06	5.8060e-07	1.8647	0.06302	*
Adjusted R^2	0.6366				
F-statistic	110.2 on 6 and 368 DF				
p-value	$< 2.2e-16$				

Table 7: Regression of $|vrp|$ on explanatory variables: Crisis Subperiod. Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are < 1%, < 5%, and < 10% for ***, **, and *, respectively.

Variable	Estimate	Std Error	t-value	Pr(> $\|t\|$)	Significance
(Intercept)	-6.3951e+00	3.1377e+00	-2.0381	0.042046	**
Mkt	5.2553e-02	1.6511e-01	0.3183	0.750387	
SMB	-2.5424e-01	2.6633e-01	-0.9546	0.340230	
HML	2.9242e-01	3.3875e-01	0.8632	0.388413	
CS	1.1665e+01	2.2727e+00	5.1329	4.042e-07	***
TED	-5.2204e+00	3.0118e+00	-1.7333	0.083630	*
POI	4.9471e-07	1.6949e-07	2.9188	0.003666	***
Adjusted R^2	0.2775				
F-statistic	34.55 on 6 and 518 DF				
p-value	$< 2.2e-16$				

Table 8: Regression of $|vrp|$ on explanatory variables: Post-Crisis Subperiod. Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are < 1%, < 5%, and < 10% for ***, **, and *, respectively.

Variable	Estimate	Std Error	t-value	Pr(> $\|t\|$)	Significance
(Intercept)	5.28323	1.70561	3.0976	0.001996	***
VIX_{t-21}	0.45133	0.06634	6.8033	1.601e-11	***
Adjusted R^2	0.3256				
F-statistic	588 on 1 and 1215 DF				
p-value	$< 2.2e-16$				

Table 9: Regression to test the Expectation Hypothesis. Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are < 1%, < 5%, and < 10% for ***, **, and *, respectively.

Variable	Estimate	Std Error	t-value	Pr(> $\|t\|$)	Significance
(Intercept)	11.61774	3.27874	3.5434	0.0004101	***
VIX_{t-21}	0.47898	0.05470	8.7565	$< 2.2e-16$	***
$\text{sgn}(vrp_{t-21})$	8.26995	3.13237	2.6402	0.0083930	***
Adjusted R^2	0.547				
F-statistic	735.2 on 2 and 1214 DF				
p-value	$< 2.2e-16$				

Table 10: Regression to test the Modified Expectation Hypothesis (H3). Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are < 1%, < 5%, and < 10% for ***, **, and *, respectively.

Positive *vrp* Dates	Trading Days	Avg(vrp)	S&P500 Return
2007/02/02 - 2007/02/27	17	1.49	-47.81%
2007/07/12 - 2007/07/26	11	2.76	-53.34%
2007/12/21 - 2008/01/02	7	1.17	-30.65%
2008/08/15 - 2008/10/08	38	11.69	-169.62%
2010/04/08 - 2010/05/04	19	3.63	-8.37%
Average S&P500 return when *vrp* is positive:			-88.13%

(a) Panel A: S&P500 returns (annualized) when *vrp* is positive

Negative *vrp* Dates	Trading Days	Avg(vrp)	S&P500 Return
2006/07/31 - 2007/02/01	128	-3.88	24.50%
2007/02/28 - 2007/07/11	93	-3.72	22.89%
2007/07/27 - 2007/12/20	103	-8.18	-1.71%
2008/01/03 - 2008/08/14	156	-6.52	-15.89%
2008/10/09 - 2010/04/07	375	-13.40	18.56%
2010/05/05 - 2011/06/29	291	-9.88	10.42%
Average S&P500 return when *vrp* is negative:			10.99%

(b) Panel B: S&P500 returns (annualized) when *vrp* is negative

Table 11: Test of Delta-Hedged Gain/Loss Hypothesis

Variable	Estimate	Std Error	t-value	$\Pr(>\lvert t\rvert)$	Significance
(Intercept)	-0.38570	0.23188	-1.6633	0.09650	*
sgn(*vrp*)	-0.49555	0.22375	-2.2148	0.02696	**
Adjusted R^2	0.00347				
F-statistic	5.308 on 1 and 1236 DF				
p-value	0.0214				

Table 12: Regression of S&P 500 return on the direction of *vrp*. Standard errors and t-statistics are computed using the Newey-West correction. Significance levels are $< 1\%$, $< 5\%$, and $< 10\%$ for ***, **, and *, respectively.