The Dynamics of the Variance Risk Premium: Refining Volatility Forecasts and Portfolio Returns*

Fredrik Øien Iversen[a] Andreas Rückert[b]

Master Thesis in Finance
Summer 2014

Abstract

In this paper, we investigate the dynamics of the variance risk premium and whether it can be used to achieve incremental predictability of future volatility on the S&P 500 index. Previous studies have focused on the usefulness of implied volatility in volatility forecasting. However, these studies do not take into account the distortion in implied volatility caused by the risk premium investors demand, despite that such a distortion has been suspected for decades. We use a parsimonious model that accounts for the time varying risk premium embedded in implied volatility. We come up with the instantaneous variance premium (IVP) variable designed to capture the time variability of the variance risk premium. Our proposed forecasting model is able to outperform prevailing models both in-sample and out-of-sample. A trading strategy with exposure to variance risk is created based on the results, and it shows that the novel variables are able to capture the dynamic properties of the variance risk premium. The trading strategy shows an improved performance compared to a benchmark, which provides evidence that the IVP variable can be used to accurately predict the variance risk premium. Our analysis uncovers useful information about the behavior of the VRP both in the context of volatility forecasting and portfolio optimization.

JEL classification: C22, C53, G17
Keywords: Volatility forecasting, variance risk premium, implied volatility, realized kernel

---

* We would like to thank our supervisor Assistant Professor Michael Halling at Stockholm School of Economics for his useful comments and guidance. Furthermore, we want to thank our parents who built the foundation of our education. Of course, they are directly responsible for any mistakes, errors, sins, or omissions herein.

[a] 40409@student.hhs.se
[b] 40406@student.hhs.se

Contents

List of Figures

List of Tables

1. Introduction

"A ship is safe in harbor, but that's not what ships are for." – William G.T. Shedd.

Risk is a necessary evil we must face to move forward, whether to move forward as a society or to cross the ocean. We take particular risk on the expectation that it will bring sufficient contribution, while optimal behavior implies taking risks that prove worthwhile. This is the central paradigm of finance; reward must be proportionate to risk. The expectations of future risk and reward thus dictates behavior and hence, everything in the financial markets. Accordingly, the purpose to which everyone abides is to maximize rewards and minimize risks. In the financial markets, risk is associated with the degree to which the rewards fluctuate, known as *volatility*.

Hence, the modeling and forecasting of volatility is a key aspect of finance. For instance, asset-pricing models employ measures of volatility as a proxy for risk that in turn are key to most asset pricing procedures, while in portfolio management it is crucial to have reliable volatility estimates in order to properly hedge risk. Unsurprisingly, this very active research area in finance has resulted in the development of many models seeking to address this objective. The most popular representatives in literature are the Autoregressive Moving Average (ARMA), the (Generalized) Autoregressive Conditional Heteroscedasticity (ARCH/GARCH), the Stochastic Volatility (SV) and the Regime Switching models.

In this thesis, we devise a more accurate model for forecasting volatility relative to conventional models. The intuitive nature of our model implicates, if its assumptions hold, not only an increased accuracy for forecasting volatility, but also a valuable economic interpretation. A trading strategy with exposure to volatility is designed on the basis of our results to validate our model and its interpretations, and to further demonstrate its applications.

An extensive effort has been put into uncovering the dynamic properties of volatility and the findings show some characteristic properties, such as persistence and long memory (Andersen & Bollerslev, 1997). As a result, past values of volatility can be useful when forecasting future volatility. Some forecasting models include both past values of volatility and implied volatility (henceforth IV, a list of acronyms can be found at the beginning of the Appendix) derived from option prices (Busch, Christensen, & Nielsen, 2011). In a semi strong efficient market, current market prices should contain all available information from historical data (Fama, 1970). Hence, the known dynamic properties of volatility – which past values of volatility tries to capture – should already be subsumed in IV, assuming the semi-strong efficient market hypothesis hold. In this thesis, a volatility model without past values of volatility is proposed. Instead, the model is based on forward-looking information such as IV and other risk neutral moments derived from option prices.

An inaccuracy of using IV in forecasting stems from the fact that IV is estimated in a risk-neutral world (Chernov, 2007). It is not real probabilities, but risk neutral probabilities that are the ingredients of risk neutral valuations. Hence, option prices, which are the source of IV, are potentially affected by the risk aversion of the market participants and not only the future expectations of volatility. Therefore, in order to get a more accurate forecast of volatility, an adjustment for this risk premium to the IV should be made. The risk premium connecting the option implied variance with the true variance is entitled variance risk premium (VRP) (Carr & Wu, 2009). To elaborate, the IV reflects the price for portfolio protection against volatility. This price depends on the expectations about the future volatility, in addition to a risk premium that investors pay to benefit from the unlimited upside and limited downside the option offers. Thus, the VRP is the difference between implied variance and variance, and similarly the square root of VRP is the difference between IV and volatility. The VRP can be thought of as the insurance premium of the financial markets. Since volatility, and hence option prices, is related to the probability distribution of the returns of the underlying, it is possible to back out these probabilities if option prices can be observed. Like the IV, these probabilities are distorted by the VRP. For options, attention has been devoted to extracting real probabilities from option-implied probabilities by studying the relation between the risk neutral density and physical density. The topic still remains a challenge for researchers and practitioners (Poon & Granger, 2003).

The model proposed in this study is based on a fundamentally different intuition than conventional volatility models: one that adjusts for the VRP (strictly speaking, the square root of the VRP). The intuition behind the model is that option implied information is, assuming semi-strong form efficiency, the best estimate for the future volatility – once the distortion related to the VRP is accounted for. Accordingly, we employ a modified version of the recent HAR-RV model of Corsi (2009). In its conventional format, the model uses past values of realized volatility aggregated over different time scales to forecast volatility - essentially moving averages over different lengths. In this way the model can capture dynamic properties of volatility such as persistence and mean reversion.

Instead of using past values of realized volatility, we introduce the instantaneous variance (risk) premium (IVP) variable. The IVP variable is a proxy for the VRP, and in spirit of Corsi (2009) we use aggregations of this proxy over three different time scales. The three aggregations are essentially moving averages of the IVP variable of different lengths. We exclude past values of realized volatility from the model on the basis that they are not helpful in the analysis of the model and the variables. The intuition of the model, when excluding past values of volatility, is consistent with a semi-strong form efficient market, although this is not the motivation for the exclusion. Strictly speaking, a model containing both IV and past values of volatility may not violate the efficient market view, since if the past values of volatility actually predict a time varying discount factor (increased risk), it may very well be consistent with an efficient market too. The reason why the variables are excluded is to make it easier to interpret and analyze our newly introduced IVP variable. To explain in further detail, remember that historical values are redundant in a risk neutral and semi-strong form efficient market. In such a case, including past values of realized volatility, which is highly correlated with IV, will only serve to bias the coefficients of the model. Distinctive from most volatility models, the shorter and neater model obtained by excluding past values of volatility has the feature of enabling valuable interpretations. Key to understanding this is seeing that any additional predictability beyond IV is a prediction of the VRP (since the square root of VRP is the difference between IV and volatility). In other words, the other explanatory variables in a regression model including IV should predict the VRP, assuming the coefficients are true and not biased. The assumption does not hold for common volatility forecasting models, as estimation of coefficients of the realized volatility variables interacts with the coefficient for IV, because of high correlation among the variables. In our proposed model, the coefficient of IV happens to be rather insensitive to the exclusion of the other variables in the model, and this makes interpretation easier.

Furthermore, we test the impact of adding additional risk neutral moments to our model. Including the risk neutral skewness and kurtosis to the model is not really consistent with the efficient market view, and may actually cause some of the aforementioned problems to reoccur, although likely not to a severe extent. Moreover, potential discoveries on the information content of skewness and kurtosis may contribute to the literature, and hence there is a good reason for including them.

The IVP variable is designed to adjust for the VRP to improve the accuracy of forecasts based on IV. Rather than capturing the dynamic properties of past-realized volatility – which should already be incorporated in the IV – the purpose is to capture any systematic time variation in the VRP. In contrast to the use of past values of (monthly) VRP, the key innovation of the IVP variable is its rapid reaction to market information. In essence, the idea of the IVP variable is to be an instantaneous measure of the (square root of) VRP. Whereas estimating the monthly VRP requires one month of volatility data, the IVP variable requires only one day of volatility data. In other words, the IVP variable is a proxy for the VRP that is first observable in the data one month later. The proximity in time, between the lagged proxy of the VRP and today's IV, is favorable in the context of capturing the dynamic properties of the VRP, since we want to capture the premium component of today's IV. The effectiveness of the model boils down to whether the proposed variables are able to capture the dynamics of the VRP, and thereby improve the volatility forecast by having predictive power over the VRP component of IV. As of today, attempts to discover the systematic behavior of the VRP have produced little evidence on how to predict it (Carr & Wu, 2009; Todorov, 2010; Vilkov, 2008). Time varying risk premium has nevertheless become a well-known concept in the equity markets. Over time, research has increasingly focused on the time variation of risk premium, a development prompted by persistent empirical regularities: the well-known "puzzles" and "anomalies" can practically all be traced back to discount factor variation rather than problems of information incorporation (see work by Bansal & Yaron (2004) and Hansen, Heaton, & Li (2008), among others). By the same token, the VRP is suggested to be time varying, although any major systematic behavior remains largely undetected.

Our model, denoted the *HAR-IVP-RNM* model (where RNM refers to risk neutral moments), successfully accomplishes its purpose by showing improvement in forecast accuracy for both the daily, weekly and monthly horizons. Most notably, for the daily horizon, the adjusted R-squared is over 8 % higher than the second best model. While our model is superior even on a standalone basis, it can also be merged with existing models and potentially improve the accuracy further.

As a final step, the interpretation of the IVP variable is being used in a trading strategy based on predicting the VRP. The intention is that a successful strategy will support the interpretation/acceptance of the model. The results show a magnificent risk adjusted outperformance - with more than a doubling of the Sortino ratio compared to the benchmark, which suggests that the IVP variable indeed can be used to predict the VRP.

For the remainder of this study we will refer to the difference between IV and volatility also as the (square root of) VRP. Moreover, we name our variable Instantaneous Variance Premium rather than instantaneous volatility premium, even though it is the difference between volatilities. We do this in order emphasize that these expressions all refer to the same type of risk premium, and to avoid creating confusing by introducing several words for a risk premium of very similar nature, particularly because many readers may associate volatility risk premium with low volatility stocks.

1.1 Outline

We will commence this study by covering relevant areas in the literature. Most prominently, this involves literature on volatility, forecasting models, risk neutral moments and the VRP. Next, we go into further detail of the methodology applied in our analysis. A short statement of the purpose of the analysis institutes the section, before the procedures applied for this motive are outlined. The section entails one part of definitions and derivations, and one part that describes the empirical analysis conducted. Finally, we present the results and conclude our analysis.

2. Literature Review

The volatility of asset prices is one of the most essential topics in the financial literature. Accordingly, there exists a vast literature on the subject. Notwithstanding the recognition and importance of volatility, even finance professionals have difficulties dealing with it, e.g. Goldstein & Taleb (2007). One common problem of volatility is the combination that volatility itself cannot be observed (only the realizations) and that it varies over time - the latter already indicated by Mandelbrot (1963) and Fama (1965).

The following will provide a brief overview of the most important papers that are related to our topic. This part of the thesis is organized as follows. The first section mainly covers research on volatility. It summarizes the different measures of volatility, various forecasting models and also the predictive power of IV/risk-neutral moments on volatility. Section 2 sketches the research results on the VRP, whereas section 3 is mainly concerned about the existing literature on risk neutral moments. One external input to calculate the risk neutral moments is the risk-free rate. Since our approach differs from the one of traditional papers, we conclude this section with an overview of OIS discounting and its most important papers.

2.1 Volatility and Forecasting

Research on volatility has developed rapidly and experienced a new dawn with the advent of high frequency data. Because the true latent volatility is not known, there is an ever-growing need for an accurate measure, given the progression and development of complex financial instruments. Andersen & Bollerslev (1998) were the first to propose the so-called realized volatility as an alternative method of volatility estimation. A simple form of the realized volatility is calculated by the square root of realized variance, which is measured as the sum of intra-day squared returns. The choice of interval frequency of the returns is a trade-off between microstructure noise, that can arise e.g. through the bid-ask bounce or liquidity issues, and accuracy, which is theoretically optimized using the highest possible frequency. A volatility estimator considered to be robust to noise is the so-called realized kernel (that we apply here). The estimator is typically computed with high frequency return data, such as second-by-second returns (Barndorff-Nielsen, Hansen, Lunde, & Shephard, 2011). Hansen & Lunde (2005) show that the use of a noisy proxy, in place of the true volatility, can severely distort a comparison of volatility models. A thorough review about realized volatility can be found in the article by McAleer & Medeiros (2008).

A particular area that has attracted recent interest is the one concerned with continuous volatility and jumps. It has been shown that, in theory, volatility can be decomposed into two components: a diffusion component and a jump component. Bipower variation was introduced by Barndorff-Nielsen & Shephard (2004) as a measure of the continuous or diffusion component, as it is robust to the presence of jumps. More recently, an even more robust estimator has been developed called threshold bipower variation by Corsi, Pirino, & Renò (2010).

2.1.1 Volatility Forecasts Models

Various models have been adopted with the purpose of forecasting volatility. Some of the most prominent are the Generalized Autoregressive Conditional Heteroscedasticity models (GARCH-models), which are extensions of the ARCH model described by Engle (1982). Bollerslev (1986) later extended Engle's model to GARCH. Common for volatility forecasting models is the aim to capture the dynamic properties of volatility. The ARCH model is based on the concept of volatility clustering – that volatility one period is auto-correlated with volatility in proceeding periods. ARCH/GARCH models incorporate this feature while also solving the problem of heteroscedasticity present in other traditional econometric models when modeling volatility.

Since the original models of Engle (1982) and Bollerslev (1986), a number of variants of the models have been developed. One example is the NGARCH introduced by Engle & Ng (1993), which incorporates the leverage effect (asymmetric volatility phenomenon). Several other types of models have also been proposed to model volatility, such as stochastic volatility models (Hull & White, 1987; Satchell & Knight, 2007; Taylor, 1986) and long memory models (Baillie, Chung, & Tieslau, 1996; Granger & Joyeux, 1980). The model of Heston (1993) is among the most popular stochastic volatility models.

Later, a new set of models have emerged belonging to the Autoregressive Fractionally Integrated Moving Average (ARFIMA) type, which was first introduced in 1999 and published by Andersen, Bollerslev, Diebold, & Labys (2003). The model estimates a tri-variate long-memory Gaussian VAR for the logarithmic realized volatilities, and has proven useful in modeling time series with long memory (slower than exponential decay of deviations from mean). Oomen (2001) confirms this view by finding that the ARFIMA model outperforms that of conventional GARCH.

A more recent model for volatility forecasting, and an alternative to ARFIMA, is the Heterogeneous Autoregressive model of Realized Volatility (HAR-RV) of Corsi (2009). The beauty of the model is its simplicity, yet remarkably strong performance. In this model the long memory is approximated by aggregation of different time scales (daily, weekly, monthly) of past volatility. Based on the HAR-RV model, refined models add IV as an explanatory variable (HAR-RV-IV). To show the rationale behind it, a compact literature overview about the information content of IV and risk neutral moments follows this section.

2.1.2 The Information Content of Implied Volatility and Risk Neutral Moments

Since the Black-Scholes option pricing formula was derived, IV moved to the center of attention. To calculate the option price via Black & Scholes (1973), the current stock price, the strike price of the option, the risk-free rate, the option time to maturity and the standard deviation of the return of the underlying asset were required. Once an option is traded, all parameters are easily observable except the standard deviation. Thus, one can obtain the IV by solving for the standard deviation. Merton (1973) generalized the mentioned model and allowed for time-varying, non-stochastic volatility.

One of the first studies that focused on the predictive power of IV, is Latané & Rendleman (1976). In a simple model, where they take the weighted average of Black-Scholes call option implied volatilities, they discovered that this is a better predictor of future volatility than using a specific historical model. Canina & Figlewski (1993) challenge this view in their paper. The authors examine the information content of option implied volatility and discover that the option implied forecasts are inferior to historical forecast techniques of volatility. Fleming (1998) reaches a different conclusion in a subsequent study. While applying a refined methodology, Christensen & Prabhala (1998) confirmed the findings of Fleming (1998). Both studies find that forecasts based on IV outperform those based on historical volatility.

Guo (1998) and Jorion (1995) are the first who investigate the foreign exchange market and find similar results. The latter reports that historical time-series models cannot cope with option implied volatility forecasts even if you allow calibration over the whole sample. In a related study, Martens & Zein (2004) analyze time-series models that use high-frequency data. They discover that these models sometimes outperform option-implied forecasts. Charoenwong, Jenwittayaroje, & Low (2009) are able to show that the predictive power of IV is irrespective of the trading venue.

Szakmary, Ors, Kyoung Kim, & Davidson (2003) are the first – to our knowledge – who compare at-the-money IV to GARCH and simple moving average models in the commodity market. They document the superiority of at-the-money implied forecasts. In a related study, Agnolucci (2009) investigates a more advance set of GARCH models and compare it to the predictive power of at-the-money IV. Contrary, he finds that the examined time-series models provide a better forecast in the crude oil market.

A milestone in the literature sketches the work of Jiang & Tian (2005). In their work, they propose a model that uses the model-free IV. Consequently, their model does not depend on any specific option-pricing model. In subsequent papers based on these findings, Frijns, Tallau, & Tourani-Rad (2010) and Cheng & Fung (2012) were able to show, that the model-free IV is superior to time series models in the stock market of Australia and Hong Kong respectively. In a comprehensive investigation, Taylor, Yadav, & Zhang (2010) provide an analysis of the performance of ARCH, GJR, GARCH, at-the-money IV and model-free IV and report that the time-series models are inferior to option implied forecasts.

Another series of papers investigate the information content of risk neutral moments. The work of Dennis & Mayhew (2002) can be seen as the first in this area. They mention a negative relation between systematic risk and risk-neutral skewness in their work. They conclude that market risk is reflected in the risk-neutral skewness extracted from the option prices. The notion has since been substantiated by Doran, Peterson, & Tarrant (2007) and Han (2007), who discover that risk-neutral skewness has strong predictive power on the stock market. These results are confirmed by Rehman & Vilkov (2012) who report that risk-neutral skewness is positively related to future stock returns. However, Chang, Christoffersen, Jacobs, & Vainberg (2011) are the first who show that the risk-neutral skewness can be used to improve the predictive power of future volatility. Very recently, Byun & Kim (2013) proposed a novel model (HAR-RV-IV-SK) to forecast volatility. Their attempt takes the risk-neutral skewness into account. They discover that the risk-neutral skewness contains additional information, and with it, they are able to slightly improve the accuracy of forecasting future volatility of the S&P 500.

The aforementioned papers can be criticized on the grounds that they ignore the impact of the VRP in their empirical studies. The work of Chernov (2007) shows that the risk premium leads to a disparity between realized volatility and IV. He concludes that the existence of VRP drives the biasedness of volatility forecasts using IV. Thus, it would be reasonable to take the VRP into account while predicting volatility.

2.2 Variance Risk Premium

The difference between the implied and expected variances is called the VRP, as found in Bollerslev (2009). Carr & Wu (2009) propose to use the difference between realized variance and a synthetic variance swap rate derived from options to quantify the premium. The VRP is related to the compensation for taking on variance risk. Traditional risk factors fail to explain the VRP, and the majority of the market VRP seems to be generated by an independent variance risk factor (Carr & Wu, 2009).

One view is that of Rosenberg & Engle (2002), Bakshi & Madan (2006), and Bollerslev, Gibson, & Zhou (2011), who have interpreted it as a measure of aggregate risk aversion. Another view is that of Bollerslev, Tauchen, & Zhou (2009) and Drechsler & Yaron (2010), who see it as a factor of aggregate economic uncertainty. While the interpretations of the VRP are inconclusive, many studies have focused on the properties and characteristics of it.

In the papers of Carr & Wu (2009) and Leippold, Wu, & Egloff (2007), they investigate the historical behavior and the pricing of the VRP and find a strong presence of the VRP for 100 different indexes including the S&P500 (meaning you obtain a positive return from being exposed to variance risk). Carr & Wu (2009) suggest that the VRP is time varying and a somewhat inconclusive notion that it is correlated with the variance swap rate (the level of IV).

Bollerslev et al. (2009) as well as Zhou (2010) find that the VRP is able to explain a significant fraction of the time series variation in returns of the aggregate S&P500 composite index, whereas a high VRP predicts high future returns and vice versa. The predictive power is strong in the short run, where it dominates other popular predictors like the P/E ratio, the consumption-wealth ratio (CAY), and the default spread. Bollerslev et al. (2011) confirm these results and conclude that the VRP helps to predict future stock market returns.

Based on previous results, Vilkov (2008) studies the dynamics and cross-sectional properties of the VRP, whereas Drechsler & Yaron (2010) show the conditions under which the VRP displays significant time variation and return predictability.

Bollerslev, Marrone, Xu, & Zhou (2012) demonstrate that these empirical findings about the predictive power of the VRP cannot be explained by statistical finite sample biases. On the contrary, they were able to show that the VRP in stock markets in France, Germany, Japan, Switzerland and the U.K. have similar predictive power.

2.2.1 Volatility Forecast using Variance Risk Premium

Lamoureux & Lastrapes (1993) tested the semi-strong efficient-market hypothesis on IV, i.e. forecasts from time-series models of stock-return processes should not have any predictive power given the market forecast as embodied in option prices. They are able to reject the null and mention as one possible reason that volatility risk is priced into the IV. Despite the findings that the risk premium is time-varying in the stock market and that it could be a source of distortion in the implied forecast embedded in the option prices, the topic has not been in the center of attention for further academic research and this conjecture remains largely unexplored.

Hence, there are only a very few influential studies that examine this conjecture further. Poteshman (2000) uses a parametric approach to calculate the IV from the Heston (1993) model, which allows taking a predetermined premium into account. With his approach, he reduces the biasedness of IV and concludes that allowing for a risk premium is the reason for it.

The first to truly include the VRP in volatility forecasting were Prokopczuk & Simen (2014). With their non-parametric approach, they are able to improve the forecasting performance of IV in the commodity market. However, the model lacks in flexibility to react on different market environments by using a one-year average of the VRP.

2.3 Risk Neutral Moments

Generally, the literature suggests different techniques to derive information about the risk-neutral distribution and moments. Breeden & Litzenberger (1978) and Banz & Miller (1978) show that the risk-neutral density is the second derivative of the option pricing function with respect to the strike price. An extremely valuable feature of this procedure is that it is model-free, and thus, the risk neutral distribution does not depend on any particular pricing model.

The prevailing practice among researchers, cf. Hansis, Schlag, & Vilkov (2010), Neumann & Skiadopoulos (2013), Conrad, Dittmar, & Ghysels (2013), Byun & Kim (2013) and more, is to use the method of Bakshi, Kapadia, & Madan (2003) to extract the risk-neutral variance, skewness, and kurtosis from option prices. Their method is also based on the result of Breeden & Litzenberger (1978) and similarly it relies on a continuum of strikes and does not incorporate specific assumptions on an underlying model. The neat approach of Bakshi et al. (2003) will be the adopted in this study.

Another approach to derive risk-neutral moments is to first derive the risk-neutral distribution and, based on these results, determine the risk-neutral moments. A lot of literature is available in this area. Jackwerth (2004) and Cont (1998) provide a good review on how to estimate the risk-neutral distribution. In general, there are two different ways to derive the risk neutral distribution, a parametric and a non-parametric approach.

Parametric methods postulate a particular form for the risk neutral density and fit the parameters to the observed option prices. A drawback of using this method is that the parametric probability distribution might not be flexible enough for matching the observed option prices. A pioneer in this area is Rubinstein (1994) who proposes an algorithm that constructs an implied binomial tree reflecting the risk-neutral density.

Non-parametric methods do not make any assumption on the form of the risk-neutral density. Hence, they allow for more general functions. These approaches achieve greater flexibility at the cost of requiring a lot of data. Furthermore, the methods bear several problems since scaling is required in order for the probability distribution to sum up to 1, to be positive, and to exhibit some smoothness. A very thorough study is found in Figlewski (2010), who evaluates previous models, combines and modifies them. Thus, he offers a neat and subtle approach to extract the risk-neutral density from option prices.

2.3.1 OIS – Risk-Free Rate

The risk-free term structure of interest rates is a key input to the risk-neutral valuation of derivatives. It is the most important ingredient for defining the expected growth rates of asset prices in a risk-neutral world and for determining the discount rate for expected payoffs in this world.

The literature before the financial crisis in 2007 such as Ron (2000) or Boenkost & Schmidt (2005) was mainly focused on bootstrapping the yield curve and determining a single curve. Whereas in more recent papers, e.g. Mercurio (2009), Bianchetti (2010) and Fujii, Shimada, & Takahashi (2011) focus on the construction and application of multiple curves, as a consequence of LIBOR no longer being considered risk-free.

Hull (2014) adopts the OIS rate as a proxy for the risk-free rate in the 9th edition of his book, "Options, Futures and Derivatives". In addition, he argues in Hull & White (2013) that there is no "perfect" risk-free rate, but that the OIS rate is the best proxy currently available. He advocates that OIS rates should always be used as the risk-free rate in risk neutral valuations.

3. Data

The core data in this study consists of daily option prices on the S&P 500 in the period 2001-2013. Daily closing records of the S&P 500 index is also incorporated in the dataset. S&P 500 is the most commonly used benchmark for the U.S. equity market, and the most popular underlying for U.S. equity derivatives. Public holidays that fall on weekdays were omitted from the data set, resulting in a full sample of 3,267 observations. For our estimations, we used the first 22 observations to calculate necessary variables, creating an effective forecast sample of 3245 observations. To study the anomalies and the impact around the financial crisis (2007-2009) on our analysis, we divide our data set into three different subsamples wherever it makes sense: the pre-crisis (January 1, 2001 to August 8, 2007), the banking crisis (August 9, 2007 to August 8, 2009), and the post-crisis (August 9, 2009 to December 31, 2013) resulting in 1,658, 503, and 1,106 observations, respectively. Researchers are dating the start of the banking crisis from August 2007 (Gorton, 2009) with the jump of repo haircuts and the dramatic increase in the OIS-LIBOR spread. We take the latter event to define the start of the banking crisis and let it last for a duration of two years in order to capture the core of the crisis. For the out-of-sample forecast, we create a 2-year rolling estimation window starting from January 1, 2001 to December 31, 2002.

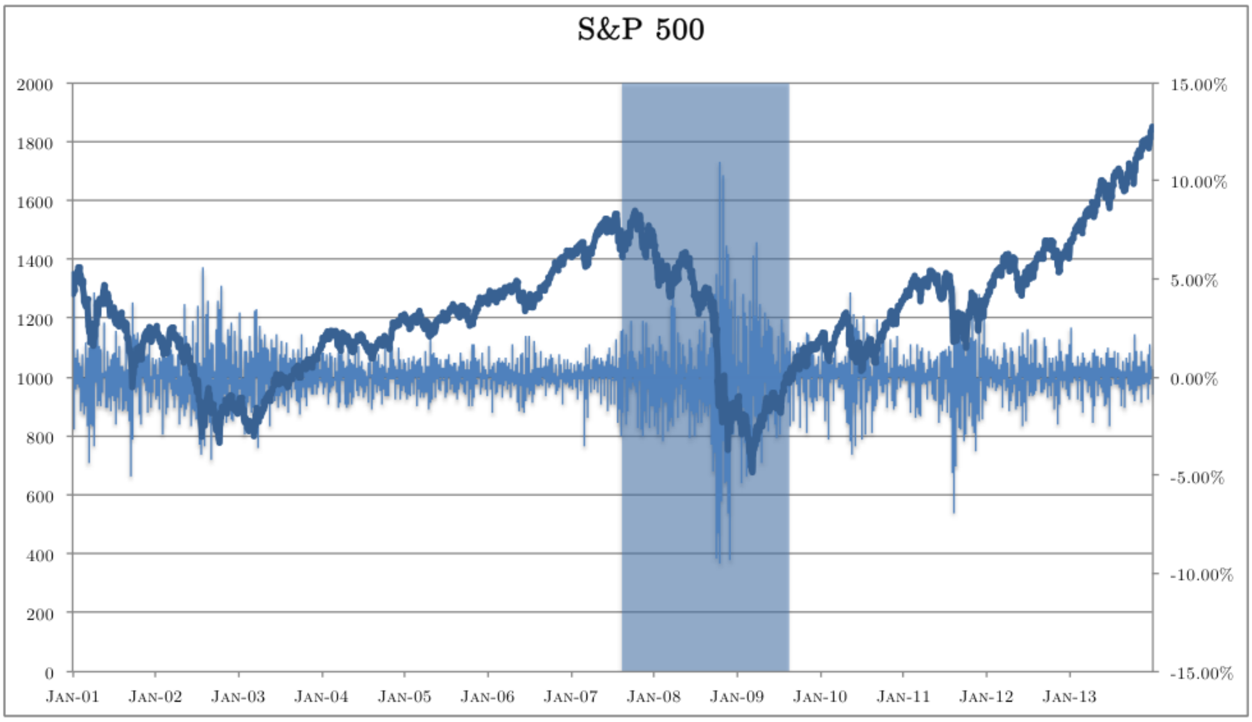


Figure 1 The graphs above show the performance (left scale) and the daily returns (right scale) of the S&P 500. The blue shaded area marks the financial crisis.

3.1 Option Data

In our study we focus on the option prices on the S&P 500 of the Chicago Board Options Exchange (CBOE). Hereby, the option prices are matched to the closing time of the S&P 500. With synchronous closing times, we prevent potential bias in our IV, which Jorion (1995) has estimated to be 1.2% in a back-of-the-envelope calculation. The raw data included all types of option contracts on the S&P 500, for every day of the period 2001-2013. This comprises a dataset with about 5 million rows for 17 different data points. The data set was filtered to include only the 3 first monthly maturities for standard index options for any given day, with information like closing price (bid and ask), strike price, trading day, expiration, option type (put-call) and closing price of the underlying. The relevant data was acquired from *Market Data Express*.

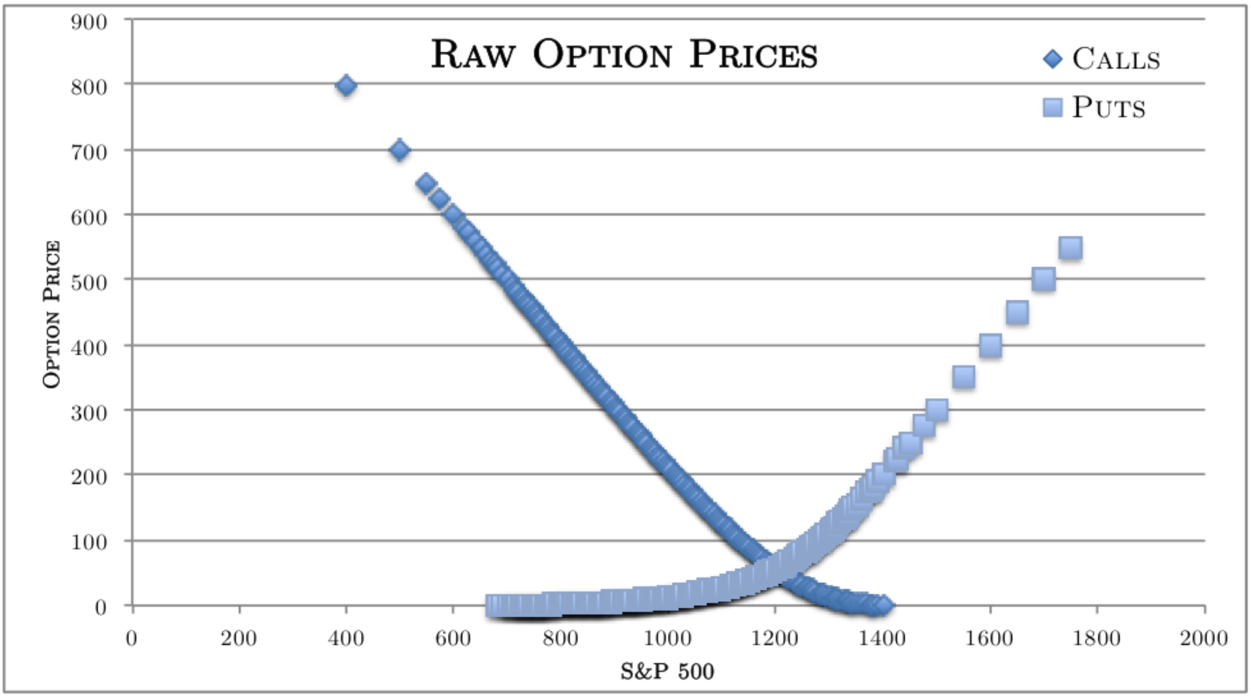


Figure 2. The graph shows the raw option prices of European options for an example day in 2011.

3.2 Volatility Data

We obtain the volatility data set, like realized volatility, bipower variation, and semivariance, from the "Realized Library" of the Oxford-Man Institute of Quantitative Finance.[1] This library is based on high-frequency data, which the institute obtains through *Reuters Data Scope Tick History*. The data set also includes closing prices for the S&P 500.

Table 1. Descriptive Statistics of the Realized Volatility (Square Root of the Realized Kernel)

	Full Sample	Pre-Crisis	Banking-Crisis	Post-Crisis
Observations	3,267	1658	503	1106
Max	1.844	0.721	1.844	1.010
Min	0.029	0.044	0.067	0.029
Mean	0.177	0.158	0.296	0.151
Median	0.146	0.137	0.232	0.131
Range 5%-95%	0.073-0.39	0.074-0.328	0.118-0.688	0.066-0.296

3.3 Risk-Free Rate Data

LIBOR was widely regarded as a good proxy for the risk-free rate. During the banking crisis from 2007 the LIBOR-OIS-spread[2] reached its maximum with 364 BPS. The gap between the two rates

1 Obtained from library version 0.2, Heber, Gerd, Asger Lunde, Neil Shephard and Kevin Sheppard (2009) "Oxford-Man Institute's realized library", Oxford-Man Institute, University of Oxford.

2 The LIBOR-OIS-spread determines the difference between the 3M LIBOR rate and the Overnight Interest Swap rate.

was seen negligible before the credit crunch, but became highly relevant due to their enormous influence on the pricing of derivatives after the financial crisis. Even so, in option pricing, the interest rate is still devoted little attention. While the interest rate has only a small impact on option prices, even tiny discrepancies in the price may transform into larger changes in IV. The choice of risk-free rate may possibly influence the analysis of the VRP.

Hence, as a proxy for the risk-free rate, we use a combination between the London Interbank Offered Rate (LIBOR) and the Overnight Indexed Swap (OIS). Before the banking crisis (2007-2009), it was best practice to discount cash flows with the 3M-LIBOR-rate, since the liquidity of it was high, and the risk was perceived as low that an AA-rated financial institution would default on a 3M-LIBOR loan.

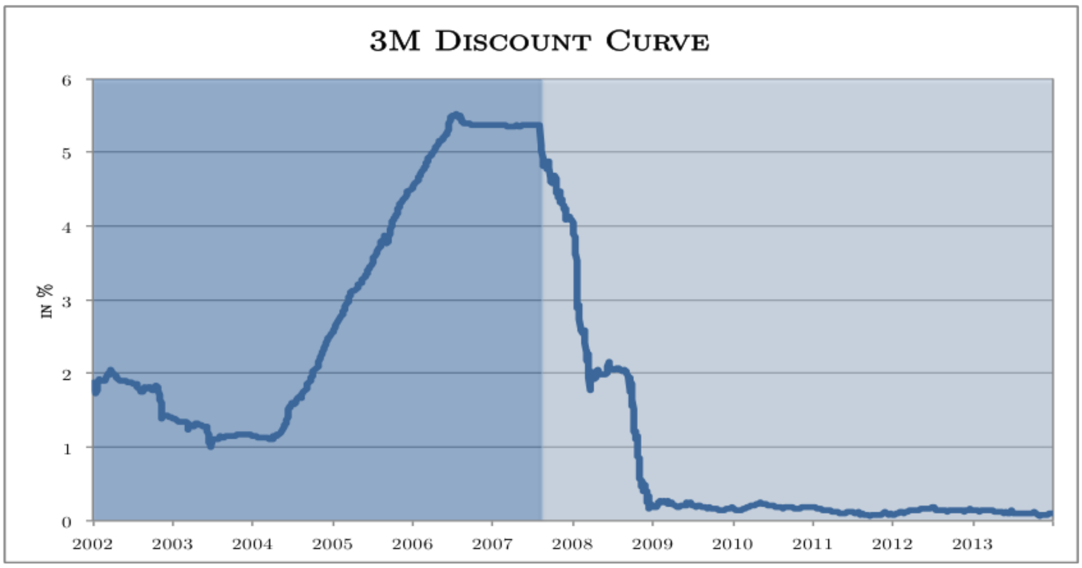


Figure 3 The graph shows a generic discount curve for the 3 months tenor. The area on the left indicates the use of LIBOR, the area on the right the use of OIS.

In this study the LIBOR is used until August 9, 2007, which is the day when the spread between the LIBOR and OIS no longer can be seen as negligible (see Figure 20 in the Appendix). Thus, from this point in time until December 31, 2013, we use the OIS as new proxy for the risk-free rate. OIS are traded with maturities starting at 1 week to up to 30 years. We use the federal funds rate as our starting point to conclude the OIS discount curve. To create a continuous yield curve, the cubic-spline method for LIBOR and OIS is applied. Hence, we are able to interpolate between the different tenors. The source for our obtained data is the financial data provider *Bloomberg*.

4. Methodology

In the following, the methodology applied to analyze the *HAR-IVP-RNM* volatility forecasting model will be explained. The *HAR-IVP-RNM* model*,* alternatively spelled *HAR-IVP-IV-SK-KU,* consists of forward-looking risk neutral moments (IV, skewness (SK), kurtosis (KU)) and the IVP variable aggregated over three time scales. The latter variables that are intended to increase the accuracy by adjusting for the bias in IV caused by the imbedded risk premium. The emphasis of the study lies within two main areas of interest:

(1) The accuracy of the forecasting model

(2) The components of the forecasting model and their interpretations

Where the components of the model can be differentiated between risk neutral moments (RNM) and variables intended to adjust for bias in the IV (IVP component). With regards to the risk neutral moments, it is mainly the improvement in forecast accuracy by using the information content of kurtosis and skewness that will be assessed. For the IVP component, the study will underscore their use as proxy and their abilities to predict the VRP. The main emphasize of the study will lie on IVP, and the insight it reveals about the VRP.

The first part of the methodology covers the calculation of the variables used in the forecasting models. Thus, we will show the methodology used to derive the risk neutral moments in addition to various volatility measures. Then, the methodology of the HAR-RV model will be described, before we later present the necessary methodology to fully grasp the *HAR-IVP-RNM* model. After this is done, the applied methodology to test the two aforementioned areas of interest is covered. The various forecasts performed in the study are presented. The forecasts aid both in assessing the accuracy and in analyzing the components in the model – albeit the latter in an insufficient way. Therefore, as a final step, a trading strategy (decision-rule) is constructed to further investigate the IVP component, to test its predictive abilities of the VRP and to uncover potential dynamics of the VRP.

4.1 Risk Neutral Moments

The procedure outlined in Bakshi et al. (2003) is used to derive the risk neutral moments. As previously stated, a valuable feature is that the procedure is model-free. This means that the method does not depend on any particular valuation model. The procedure requires a continuum of option prices.

Since the procedure to extract the risk neutral moments are not straightforward, we have to deal with several issues along the way. One, among many, is that there is no continuum of option prices across strikes. To obtain such a continuum interpolation is required. Moreover, in this thesis, a similar concept like the one to calculate VIX is adopted in order to have prices that always correspond to a maturity of 30 days. In other words, the process consists of (1) interpolating and extrapolating the option prices in the moneyness dimension, and (2) interpolating prices in the maturity dimension. The method applied in this study is more sophisticated than in standard papers deriving risk neutral moments, concerning the calculating of the 30 days average measure. Most papers estimate a 30 days average of the risk neutral moments as a final step. This is done by calculating the moments for different maturities and then interpolating them. Constructing prices that correspond to 30 days maturity, which is done here, is unprecedented and hence a departure from previous literature. This continuum of prices corresponding to 30 days maturity has useful properties for many purposes – extracting the corresponding 30 days risk neutral density is one of them. In short, the method can be considered more general in terms of applicability, and arguably also more accurate.

4.1.1 Moneyness Dimension: Constructing a Continuum of Strikes

The procedure outlined in the following is conducted for options of the three nearest monthly maturities on the S&P 500. In essence, the procedure deals with how to create a (close to) continuous range of prices. This is accomplished through interpolating in the range where market data exists, and extrapolating outside the range of available data. Noisy market data in the available set of traded contracts are a pervasive problem that must be dealt with before interpolating. The option prices are calculated as the mid-price between bid and ask prices for each strike. To avoid irregularities and inaccuracies, contracts with bid prices below $ 0.5 are excluded, as the spread in these cases become extreme relative to the price. Furthermore, since put-call parity does not seem to hold for strikes in the money or out of the money, only out of the money options are used, as these are the most liquid (and also the input required in the calculation of the moments). If the number of available strikes is less than 10, the maturity is dropped out of the calculation. However, by virtue of the maturity exclusion criteria (see maturity dimension below) and using data from S&P 500 (the most liquid equity option market) the number of available strikes is in general satisfying.

Interpolating in the moneyness dimension by interpolating option prices directly has proven unfruitful, as it tends to give extreme, even negative, implied volatilities. Slight irregularities in prices may transform into abnormal volatilities, in which case signifies that the price reflects an unrealistic view. Therefore, using the Black-Scholes formula to smooth and interpolate option values has become a common practice. The IV can be backed out of Black-Scholes since all other inputs are known. By using the Black & Scholes (1973) equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \qquad (1)$$

we obtain the Black & Scholes (1973) formula for European Call options

$$C(S,t) = N(d_1)S - N(d_2)Ke^{-r(T-t)} \qquad (2)$$

where $d_1$and d_2 is specified as:

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right] \qquad (3)$$

$$d_2 = \frac{1}{\sigma\sqrt{T-t}}\left[ln\left(\frac{S}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)\right] = d_1 - \sigma\sqrt{T-t} \qquad (4)$$

and by obeying the rules of the Put-Call parity we receive

$$P(S,t) = Ke^{-r(T-t)} - S + C(S,t) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S \qquad (5)$$

For both, as above:

- N(•) is the cumulative distribution function of the standard normal distribution
- T-t is the time to maturity
- S is the spot price of the underlying asset
- K is the strike price
- r is the risk-free rate
- σ is the volatility of returns of the underlying asset.

Further details on the Black-Scholes can be found in the original paper. Paradoxically, the idea contradicts one of the assumptions of the model of Black & Scholes (1973) (constant volatility) and is known as "practitioners Black-Scholes". The idea is to calculate an IV for each strike, corresponding to each option price. This method yields a volatility smile, which for equity options commonly is referred to as a "smirk" because of the skewed smile. Shimko (1993) was the first to propose the idea of converting prices into volatility space before interpolating, and then converting the interpolated implied volatilities back into option prices. Interestingly, Black-Scholes is merely used to convert between price space and volatility space for interpolating purposes; the overall procedure is not dependent on the *correctness* of Black-Scholes.

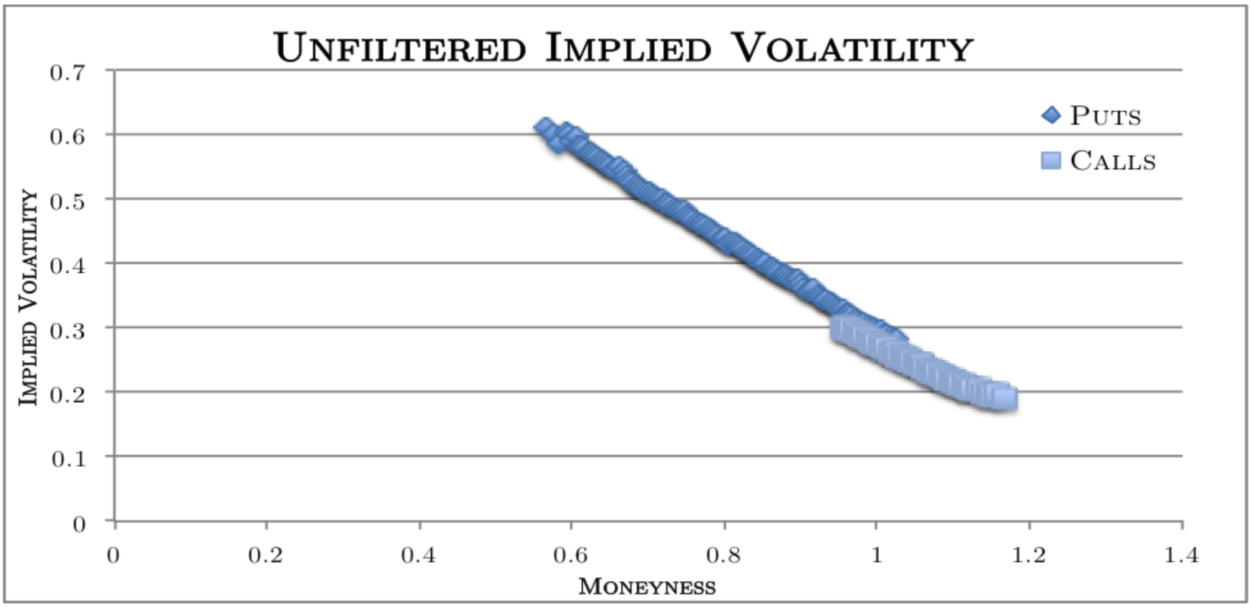


Figure 4 The plot illustrates the implied volatilities of maturity 2 after using Black-Scholes to convert the option prices. Prices corresponding to negative implied volatilities – typically far ITM options – are automatically excluded.

Various interpolation techniques have been used in the literature. Cubic spline, which connects a line through all points in a price-volatility space by estimating a cubic polynomial between every point, has been widely used. Bliss & Panigirtzoglou (2004) shows that a smoothing spline is successful for the purpose of interpolating the smile, since it is more robust to noise in the data. In this thesis a smoothing spline is used. In essence, a smoothing spline is a cubic spline with the addition of a smoothing parameter, which controls a tradeoff between deviating from observations and the smoothness of the curve. This parameter may be adjusted to achieve the appropriate level of smoothing. At this stage of the procedure, it is desirable to keep the noise instead of smoothing it out. Hence, the penalizing parameter is set with very limited smoothing, only to the extent that extreme outliers - which are obvious errors in the data - will not adversely effect the interpolation.

Before the interpolation is performed a small adjustment is made. To get a more smooth transition between put and calls in the volatility smile, a weighting of put and call prices is done for a small range around ATM. Before interpolating, the strikes are converted to a standardized measure (moneyness, $\frac{K}{S}$) in order enable comparison of data across maturities and time.

Once the interpolation is completed, the next step is to deal with the ends of the smile where no market data exists. This corresponds to the tails of the RND. Again, various approaches have been applied in the literature. Carr and Wu (2009) use constant volatility as extrapolation. This gives a smile, which yields lognormal tails of the RND. Considering the empirical evidence of fat tails, this is unlikely to be a satisfactory approximation also for the RND if the tails are of importance

(Figlewski, 2010). Bliss & Panigirtzoglou (2004) employ a linear extrapolation outside the range of the available strikes, thus achieving a fatter tail of the corresponding RND. Since the main purpose of this exercise is to get inputs for forecasting, it is desirable to get theoretically precise risk neutral moments.

The idea behind the extrapolation technique applied in this thesis is making a small adjustment to the linear extrapolation used by Bliss & Panigirtzoglou (2004) in order to get an extrapolated part that is slightly more conservative, but still capturing the shape of the market data. In addition to the linear extrapolated component, the idea is to add a component that essentially extends the smile by letting the first and second derivative of the smile continue the trend set forth by the tradable strikes in the ends of the smile. This was attempted by estimating a second-degree polynomial from traded strikes near the end point, which was connected at the end point of the smile. The final extrapolated curve was a weighted average of the linear line, the fitted second-degree polynomial and a non-linear curve designed to capture the convexity of the whole smile. In practice, various constraints and measures needed to be applied to avoid issues with the extrapolation technique. The weights in the non-linear curves were set somewhat conservative for the same purpose. Extreme outliers were excluded from the estimation of the extrapolation curves. Constraints were set so that the extrapolation would not significantly overestimate of underestimate the IV. For example, constraints prevented the endpoints of the smile to take too extreme values relative to the end point of the smile calculated from market data. The refinement of the extrapolation in order to obtain satisfying smiles is just as much a work of art as science. The method proved robust enough to yield nicely shaped tails of the RND for all days in our sample. Similar to other papers, discrete strikes with increments of 0.001 moneyness were used to calculate 1401 strikes as an approximation for a continuum of strikes.

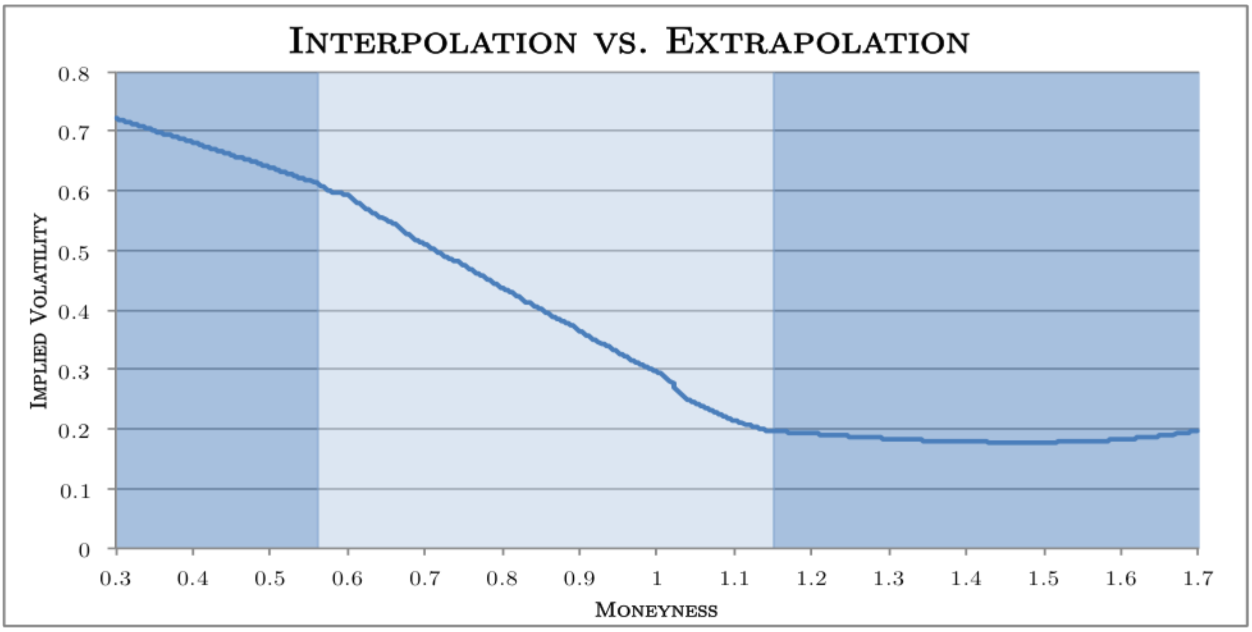


Figure 5. The plot shows the smile of maturity 2 after the interpolation (light blue area) and extrapolation (blue area) procedure are completed

4.1.2 Maturity Dimension: Constructing the 30-Days Smile

The smiles from the three nearest monthly maturities are used in order to obtain a 30 days measure. If the nearest maturity is below 10 days it is excluded to avoid unwanted effects from market microstructure of soon to expire options, a common practice among researchers. Similarly, maturities of longer than 120 days are excluded on the basis that interpolating with maturities far beyond 30 days has low validity.

Resembling the methodology of VIX, a linear interpolation of variances is conducted across maturities. Time scaling of volatility linearly in variance only holds when prices follow a Geometric Brownian motion, which implies that prices are log normally distributed and returns are normally distributed (Shreve 2004). However, the inaccuracy of interpolating between two values where the true value presumably lies in between is limited. Instead of only interpolating the final risk neutral moments, we interpolate the variance for each moneyness. By estimating values for each moneyness, rather than performing the interpolation at an aggregated level, ensures that the non-normality is better accommodated for, as the precision across the whole range of strikes is preserved. If anything, this should lead to increased accuracy of the estimated 30 days risk neutral moments, as the order of interpolation presumably matters (just like: $average(4^2 + 8^2) \neq average\ (4 + 6)^2$). Additionally, the linear interpolation of skewness and kurtosis that would otherwise have to be conducted might not be accurate. Furthermore, obtaining 30 days prices makes it possible to check whether the interpolated values corresponds to densities with abrupt changes and large negative sections etc. The weights in the variances are found in the following way:

$$\omega_{1,2} = \frac{30-T_2}{T_1-T_2} \qquad\qquad \omega_{2,3} = \frac{30-T_3}{T_2-T_3}$$

Where T_1 is time to expiration of the first maturity, $\omega_{1,2}$ is the weight in maturity 1 (the first maturity) and $(1-\omega_{1,2})$ is the weight in maturity 2, when only these two maturities are used in the calculation of the 30 days average. If the nearest maturity is over 30 days, the later maturity will have negative weights. With three smiles available we use all the smiles in the calculation, instead of only two (like VIX). A particular weighting method is used with the purpose of smoothing out the transition between the maturities used in the calculation. In the VIX calculation, where this is not the case, abrupt changes in the values may occur when the nearest maturity is dropped from the calculation in favor of the longer maturity. The idea applied here is simply adding an additional weighting, transition weight (ψ), of the original weights that smoothes the transition between $\omega_{1,2}$ and $\omega_{2,3}$:

$$\psi = \left(\frac{(T_1-10)1.25}{30}\right)^{0.25} \tag{6}$$

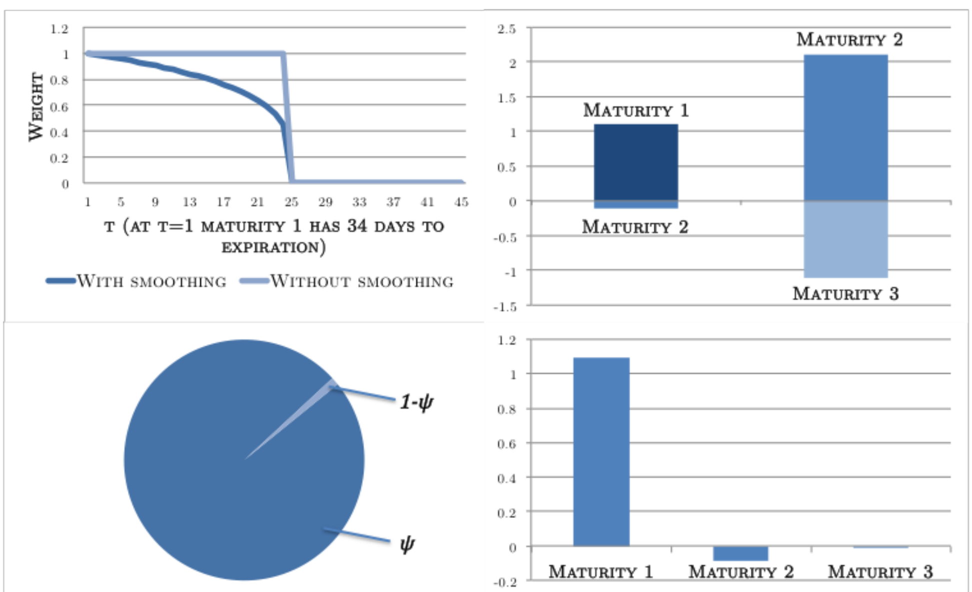
$$\omega_1 = \omega_{1,2}\psi \tag{7}$$

$$\omega_2 = (1-\omega_{1,2})\psi + \omega_{2,3}(1-\psi) \tag{8}$$

$$\omega_3 = (1-\omega_{2,3})(1-\psi) \tag{9}$$

$$\omega_1 + \omega_2 + \omega_3 = 1 \tag{10}$$

Where ω_1 is the weight in maturity 1 relative to total weight, and ψ is the weight in $\omega_{1,2}$ and $(1-\psi)$ is the weight in $\omega_{2,3}$.



(Figure 6. See description on next page.)

Figure 6. Top left: smoothed weighting vs non-smoothed weighting by showing ψ - the proportion of weight in $\omega_{1,2}$. Top right: shows the weights in the calculation of 30 days average using maturity 1 and 2, and using maturity 2 and 3. Bottom left: shows the weight in using the two first maturities in the calculation, vs using the two last maturities. Bottom right: shows the final weights in the three maturities

Once the 30 days volatility is obtained, a smoothing spline is applied in order to smooth out the smile slightly. The intention is to deal with noisy or improper values – not all mid-prices are necessary reflecting a tradable price or credible value. A side effect is a neater theoretical smile that converts to a more nicely shaped RND. In general, we do not want one or two outliers to have too much effect on the estimation of the smile and thus risk neutral moments. Since we do not know the magnitude of the impact on the risk neutral skewness and kurtosis when having a smile that corresponds to an erratic and abnormal RND, it can be seen as a sanity check to ensure that the RND looks somewhat realistic. The smoothing is done as a final step in order to keep as much information as possible from all the three smiles. Some smoothing will naturally happen, without information loss, in the weighting process. Performing the smoothing in the final step will, however, limit the impact of smoothing when it comes to loss of information content.

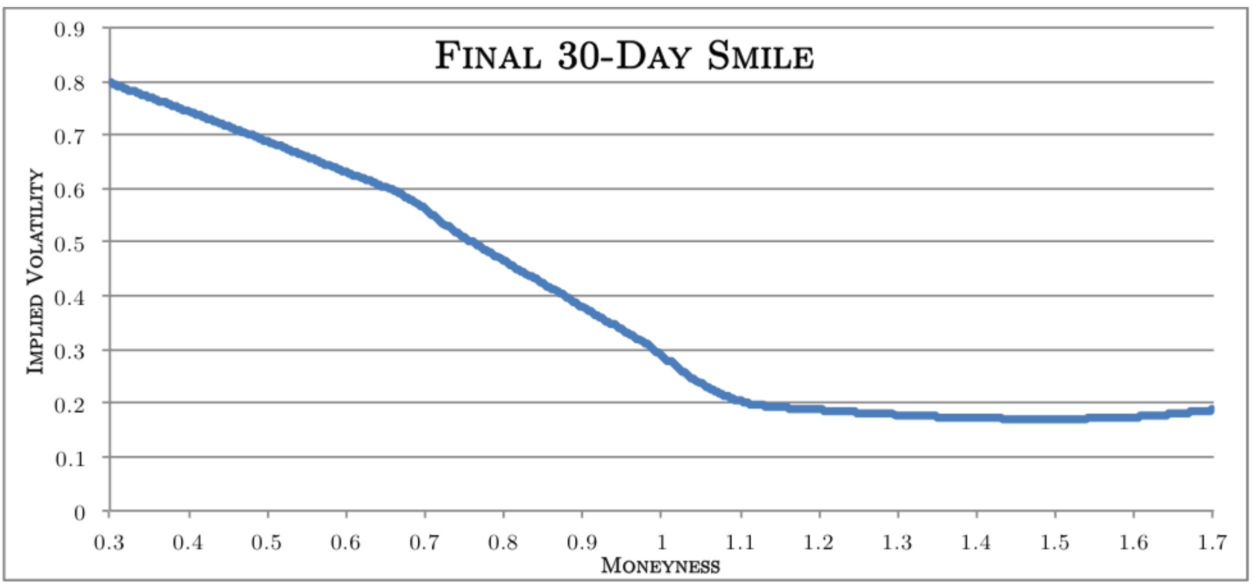


Figure 7. The volatility smile above corresponds to 30 days maturity, found by interpolating all the smiles across maturity.

In the end the implied volatilities are converted back into prices by using Black-Scholes. The implied volatilities are converted to both put and call prices, where in the end only the OTM prices are being used.

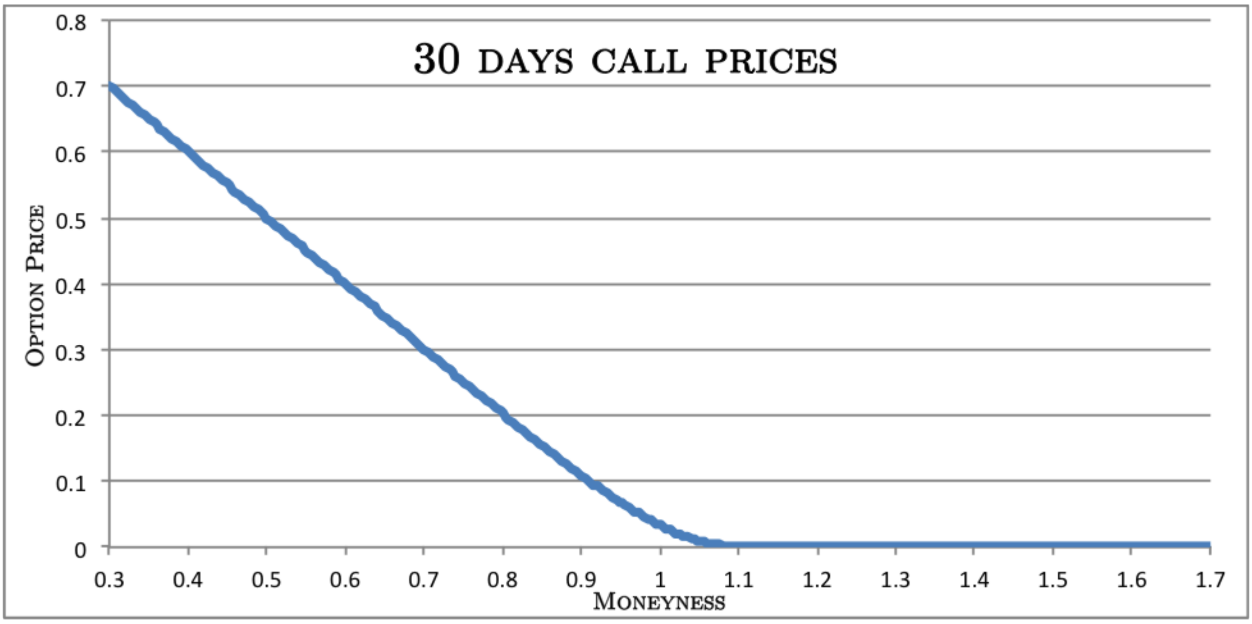


Figure 8. The plot shows the obtained call prices from converting the 30 days implied volatilities. Put prices were obtained in similar fashion.

The chain of 1401 option prices is used to derive the risk neutral moments and the RND. Further explanation of the calculation of the risk neutral moments and a derivation of the RND can be found in Appendix A and B.

4.2 Volatility Measures

As a measure of realized variance we use the realized kernel, which is more accurate than the simple realized variance using the squared intraday returns. For doing so, we follow the approach of Shephard & Sheppard (2010). The analysis is based on daily financial returns

$$r_1, r_2, r_3, \dots, r_T$$

and its corresponding sequence of daily realized measures denoted as RM_t

$$RM_1, RM_2, RM_3, \dots, RM_T$$

where realized measures are theoretically sound high frequency, nonparametric based estimators of the variation of an asset during the times at which this asset is traded frequently on an exchange. Realized measures have the favorable property that it ignores the variation of prices overnight. Furthermore it may neglect the high variation in the beginning of the trading day when recorded prices comprise errors.[3]

Realized kernel (Barndorff-Nielsen, Hansen, Lunde, & Shephard, 2008) is used as an estimator for the realized variance in combination with a Parzen weight function where the

[3] For more detailed information about the background of realized measures see Andersen, Bollerslev, & Diebold (2010) and Barndorff-Nielsen & Shephard, (2005)

bandwidth choice of H follows the recommendation of Barndorff-Nielsen, Hansen, Lunde, & Shephard (2009).

$$RM_t = \sum_{h=-H}^{H} k\left(\frac{h}{H+1}\right)\gamma_h \qquad (11)$$

$$\gamma_h = \sum_{j=|h|+1}^{n} x_{j,t}x_{j-|h|,t} \qquad (12)$$

where k(x) is the Parzen kernel function

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3, & 0 \le x \le \frac{1}{2} \\ 2(1-x)^3, & \frac{1}{2} \le x \le 1 \\ 0, & x > 0 \end{cases} \qquad (13)$$

An important property of this realized kernel is that it cannot be negative. For clarity and simplicity, we refer to the realized kernel when we talk about realized variance and realized volatility.

Other measures we make use of in our analysis are realized semivariance and bipower variation. Contrary to the realized variance, semivariance is entirely based on downward moves using high frequency data. Bipower variation is a model free estimator of integrated variance in stochastic volatility models and is robust to the presence of jumps (Barndorff-Nielsen & Shephard, 2004). To illustrate the concept of integrated variance, suppose a price process $P_t = exp(p_t)$ is given by:

$$p_t = p_0 + \int_0^t \sigma_s dB_s \qquad (14)$$

where B_s is a standard (s) Brownian motion. When σ_s follows a continuous process, the integrated variance is given by:

$$integrated\ variance = \int_0^t {\sigma_s}^2 ds \qquad (15)$$

and hence the name integrated variance.

Bipower variation is calculated as the products of powers of absolute returns. Henceforth, we will refer to bipower variation and semivariance as the square root of the respective measures (i.e. volatilities). The formulas used to estimate the realized bipower variation and realized semivariance are provided in the Appendix.

4.3 The HAR-RV Model

The forecasting models applied belong to the type Heterogeneous Autoregressive (HAR) model of Corsi (2009). Henceforth, we will invent names to the models for ease of reference, rather than using the normal acronyms (stated in the parenthesis).

The standard version, HAR-RV takes the shape of a simple multiple regression model:

Conventional Model (HAR-RV)

$$RV_{t,t+h} = \beta_0 + \beta_{RV}^D RV_t^D + \beta_{RV}^W RV_t^W + \beta_{RV}^M RV_t^M + \varepsilon_{t,t+h}$$

Where RV_t^D is the RV measured from day t-1 to t ($RV_{t-1,t}$). The variables RV_t^W and RV_t^M are the average daily RV for the past 5 days and 22 days respectively. A list of definitions can be found in the Appendix. $RV_{t,t+h}$ is the average daily RV over time t to t+h, where h is the forecast horizon. The current time, t=0, may be thought of as the time after closing on the given day. Hence, $RV_{t-1,t}$ is the RV measured earlier the same day. IV at t=0 is based on closing prices on the given day. Hence, it essentially succeeds the RV. Three horizons are being forecasted: daily, weekly and monthly. In spite of absence of true long-memory properties, the HAR-RV model successfully achieves the purpose of reproducing the main empirical features of financial returns, like long memory, fat tails, and self-similarity, as shown by Corsi (2009). This is done through the additive cascades of past values of realized volatility. The idea is that market participants are concerned with different time horizons.

Disregarding our own model, we will work with two different versions of the HAR-RV model in the forecasting. The first is the one mentioned above, and the second is an extension.

Extended Model (HAR-RV-IV)

$$RV_{t,t+h} = \beta_0 + \beta_{RV}^D RV_t^D + \beta_{RV}^W RV_t^W + \beta_{RV}^M RV_t^M + \beta_{IV} IV_t + \varepsilon_{t,t+h}$$

The model is the one proposed by Busch et al. (2011) and has included IV as a variable with the purpose of improving forecast accuracy.

4.4 The Variance Risk Premium

A definition of the VRP is necessary before delving into the proposed model. Conceptually, the VRP may be thought of in spirit of Carr & Wu (2009) as:

$$VRP = \sigma_{Implied}^2 - \sigma_{realized}^2 \qquad (16)$$

where the horizon of implied component of the variances should correspond to the horizon of the realized component. The VRP is often defined as the difference of the monthly IV and the monthly

RV[4]. Horizons significantly shorter or longer will typically suffer from less available market data when calculating IV. Nevertheless, from a theoretical perspective it would make sense to refer to VRP as the difference between the implied variance and realized variance for any horizon, as long as both components are properly matched in time dimension. For calculating the VRP, we will use a specification where the implied component is the monthly model free implied variance, and the realized component is the average daily squared RV.

Since we are using 30 days IV, the VRP can be quantified as follows:[5]

$$VRP_t = IV_{t-23}^2 - RV_{t-22,t}^2 \qquad (17)$$

where $RV_{t-22,t}^2$ is the average squared daily realized volatility of the past month and IV_{t-23}^2 is the square of the implied volatility 23 trading days ago (remember that the IVs are measured at closing time, so that the IV 23 days ago is the most recent IV preceding the last month's realized volatility. One take on the VRP is that the aforementioned definition is only a measure of it, while in fact the latent VRP is the difference between the implied variance and the expected variance, where the latter is in line with Bollerslev (2009). On average, the expected variance and realized variance can be assumed to be equal. However, taking the view that the VRP is the premium the market is willing to pay in excess of the expectations, leads to the result that our measure of VRP is subject to noise (as the realizations and expectations are only equal on average). Hence, to get an estimate of the VRP robust to noise, a long horizon should be used. This poses a problem in our context, because in order to get a long horizon, market data from far back in time must be used.

It has previously been indicated that there is a delay in how quickly the VRP reacts to new market data. In words, the above definition is saying that the latest VRP that can be calculated today at time t, is the difference between the squared IV volatility one month ago and the squared average RV from 22 trading days ago until today. This may not seem slow in the sense that new values of RV are incorporated immediately. The problem is that the value of the IV is always about a month old, which in turn means the prevailing risk premium estimated at that time reflects the price investors were willing to pay one month ago. What we are interested in is the current VRP, and hence the more recent VRP, the better. The purpose of the IVP variable is to find a way around this issue and hence improve predictability.

[4] There exist definitions of the VRP with small differences in the literature, mostly due to how the realized component is calculated

[5] Carr & Wu (2009), among others, define the VRP the other way around, by deducting the implied variance from the realized variance. However, this yields the somewhat counterintuitive result that a negative premium gives a positive return for taking on variance risk.

4.5 The Instantaneous Variance Premium (IVP) Variable

We define the IVP variable as the following:

$$IVP_r = IV_{t-1} - RV_{t-1,t} \qquad (18)$$

The IVP variable is thus the difference between the IV one day and the RV the subsequent trading day, and not the next 22 trading days. Using the RV of only one day may be seen as a proxy for the expected variance of the remaining 21 days, in order to approximate the VRP. One way to view that the current RV is a proxy for the expected variance is that shocks to volatility are known to be persistent, and beyond that the future RV is rather unpredictable. If the approximation is successful, we believe that the IVP variable will be more effective in predicting the VRP than the monthly VRP itself. This is because the IVP variable is essentially a proxy for the previous day VRP, and not the one last month, i.e. it captures the risk component of yesterday's IV. After all, we are interested in how the current risk is priced. Instead of IV_{t-1} one could potentially use IV_t, which is a more recent value, and arguably reflecting more contemporaneous variables. However, considering that IV reflects the future, it makes sense to use a subsequent RV rather than a RV that has already passed at the time of estimate of the IV. Furthermore, our specification ensures that the information from the IVP variable can be easily used in time to trade on today's option prices. Technically speaking, when using the closing value of today's IV, it is no longer possible to trade the prices that this IV reflects. However, with intra day data of IV one could obtain contemporaneous variables and a more instantaneous measure than we use here. One could possibly argue, that if the current RV is a good proxy for the expected variance, that the variant of using contemporaneous returns better reflects the VRP than the aforementioned quantification of VRP in section 4.4 (because of the noise in the estimation). That is, if one takes the view that the latent VRP is the difference between the implied variance and the expected variance.

In other words, one way to view IVP variable is that the daily RV is a proxy for the monthly RV. Another way to view the IVP variable is that it is closely linked to the daily VRP. This is the case if the monthly IV and the daily IV is closely connected. The link to the daily VRP can be detected by substituting the monthly IV with the daily IV in the IVP variable, where you would then obtain a formula for quantifying the daily VRP. Hence, one alternative to using the IVP variable could possibly be to use IV with one-day horizon. These two views on the IVP variable are intertwined; if the first view has merits, then the short-term IV and monthly IV should logically be highly correlated.

Similar to the three different variables of RV, we construct three variables with the same aggregation technique of past values, but of IVP rather than RV. More specifically, IVP_t^D is the difference between IV_{t-1} and $RV_{t-1,t}$, while the weekly (IVP_t^W) and monthly (IVP_t^M) variables are aggregations or moving averages like earlier. More details can be found in the Appendix. For simplicity and ease of reference, the three moving averages of the IVP variables are denoted the Bermuda[6] variables, when they are jointly referred to.

4.6 The HAR-IVP-RNM Model

The *HAR-IVP-RNM* model is based on the concept of the HAR-RV model in terms of using additive cascades, or moving averages, of past values. However, the additive cascades of RV are substituted with cascades of IVP. Hence, the autoregressive terms are not of the independent variable.

Bermuda Model (HAR-IVP-IV)

$$RV_{t,t+h} = \beta_0 + \beta_{IVP}^{D} IVP_t^{D} + \beta_{IVP}^{W} IVP_t^{W} + \beta_{IVP}^{M} IVP_t^{M} + \beta_{IV} IV_t + \varepsilon_{t,t+h}$$

Where IVP_t^W and IVP_t^M is weekly and monthly moving averages like earlier. The Bermuda variables are thus trying to capture the dynamics of the VRP by being daily, weekly and monthly moving averages of the IVP variable. In order to adjust for the VRP in IV, the three variables must be able to estimate the VRP component of today's (the current) IV.

4.7 Forecasting

Every forecast is always conducted for all three forecast horizons. All regressions are estimated with ordinary least squares (OLS) and Newey & West (1987) standard errors with 5, 10 and 44 lags for the daily, weekly and monthly forecast respectively. The lag lengths are chosen on the basis of the recommendations of Andersen, Bollerslev, & Diebold (2007).

First, we run the following model with only IV as explanatory variable:

$$RV_{t,t+h} = \beta_0 + \beta_{IV} IV_t + \varepsilon_{t,t+h}$$

Next, the impact of adding risk neutral skewness and kurtosis is assessed. In total, we run three different regressions for the *Extended* model and *Bermuda* model. Then, we refine the models and construct:

Extended Model (HAR-RV-RNM)*

$$RV_{t,t+h} = \beta_0 + \beta_{RV}^{D} RV_t^{D} + \beta_{RV}^{W} RV_t^{W} + \beta_{RV}^{M} RV_t^{M} + \beta_{IV} IV_t + \beta_{SK} SK + \beta_{KU} KU_t + \varepsilon_{t,t+h}$$

Bermuda Model (HAR-IVP-RNM)*

$$RV_{t,t+h} = \beta_0 + \beta_{IVP}^{D} IVP_t^{D} + \beta_{IVP}^{W} IVP_t^{W} + \beta_{IVP}^{M} IVP_t^{M} + \beta_{IV} IV_t + \beta_{SK} SK + \beta_{KU} KU_t + \varepsilon_{t,t+h}$$

[6] Bermuda is commonly associated with the Bermuda Triangle – an area where the magnetic field distorts the compass. Hence, adjustments must be made when using a compass. Similarly, our model acknowledges the necessity of adjusting for the distortion in IV (caused by the VRP). Distortion and triangle should be easily associated with three variables which purpose is to correct a distortion.

Since the forecasting mainly focus on a comparison between the performance of the *Extended* model and the *Bermuda* model, only these two models are refined. However, the *Conventional* HAR-RV is sometimes included by way of illustration, but never with risk neutral moments as explanatory variables.

Next, the models are evaluated in-sample. The three models that are run for all remaining in-sample regressions are the *Coventional, Extended** and *Bermuda** model. We run regressions for full-sample and the three subsamples.

Third, the out-of-sample performance of the *Extended** and *Bermuda* are evaluated. For the out-of-sample forecasting the models are reinitiated using the first 500 observations representing two years of data. A rolling window of length 500 observations is used to estimate the coefficients. The forecasting performance are evaluated using two different loss functions as criteria measures:

$$MSE = \frac{1}{N}\sum_{t=1}^{n}\left(\widehat{RV}_{t,t+h}^{2} - RV_{t,t+h}^{2}\right)^{2} \qquad (20)$$

$$QLIKE = \frac{1}{N}\sum_{t=1}^{n}\left(log(RV_{t,t+h}^{2}) + \frac{\widehat{RV}_{t,t+h}^{2}}{RV_{t,t+h}^{2}}\right) \qquad (21)$$

According to Patton (2011), the applied loss functions are robust to noise in the volatility proxy. Variance, instead of volatility, is commonly being used in the loss functions with the rationale that the conditional variance is the one of interest. The difference arising from using variance is that the loss function will penalize overestimations slightly more than underestimations, as compared to using volatility.

Next, in order to further assess the risk neutral moments and the Bermuda variables, we repeat the in-sample procedure with semivariance, bipower and their difference to realized volatility.

In the end, regressions are run to uncover the predictability of the Bermuda variables on the VRP. Four in-sample regressions are run, one with all three IVP variables as independent variables, and three single regressions with each variable. The dependent variable is the square of the VRP.

$$IV_t - RV_{t,t+22} = \beta_0 + \beta_{IVP}^{D} IVP_t^{D} + \beta_{IVP}^{W} IVP_t^{W} + \beta_{IVP}^{M} IVP_t^{M} + \varepsilon_{t,t+h}$$

There are however shortcomings of testing the predictability this way; the economic significance of the results are difficult to determine.

A perhaps intuitive approach to test the underlying intuition of the Bermuda model would be to include the past values of RV, and then see whether the Bermuda variables drive them out. Considering the fact that models in previous literature commonly includes both RV and IV, it could make sense to find support for excluding them. Unfortunately, there are challenges with this seemingly simple analysis. Firstly, a forecast improvement by adding RV variables to the Bermuda model is not necessarily evidence that the market is inefficient or that the Bermuda variables fail to adjust for the VRP. This is because the forecast improvement, at least for the daily and weekly horizon, could be unrelated to the VRP. Since we are using IV of monthly horizon it is possible that the forecast improvement is an adjustment of the unsynchronized horizons between the IV and the forecast, and hence is unrelated to the VRP. Secondly, even if there is only a very limited forecast improvement, the RV variables will tend to stay significant because of the high correlation between IV and the RV. Hence, adding the RV variables will cause the significance of IV to drop, while the RV variables may still be significant, even without any forecast improvement of note. Lastly, we are mainly concerned with investigating a way to adjust for the VRP component (by using the Bermuda variables). As previously mentioned, adding many additional variables to the model would not help in this regard, because of the increased difficulties in interpreting the results. Hence, since the main purpose is not to construct a model to maximize the forecast accuracy of volatility, nor to test market efficiency, but rather to assess the VRP, we refrain from adopting this approach.

4.8 Trading Strategy

The trading strategy is paramount in testing the predictability of the IVP variables on VRP, as it is out-of-sample and gives an indication of the economic significance. Hence, the trading strategy serves to test the economic foundation and the predictive power of the Bermuda variables in the *Bermuda** model. The nature of the *Bermuda** model gives an economic interpretation that is directly applicable if the Bermuda variables work as intended, namely to predict the VRP. However, the interpretation only holds if the coefficients are true and not biased, i.e. due to multicollinearity or omitted variable bias. The trading strategy is thus likely to fail with a high degree of multicollinearity.[7] Let us assume for a moment that there is no such bias in the forecast model. Since IV is included as an explanatory variable in the forecast model, the VRP must be affected if, say, IR^W predicts lower future RV for the next month, as this represents a prediction that is not incorporated in IV. Any incremental prediction beyond IV itself should predict the VRP.

Variance swaps are appropriate for the purpose of constructing the trading strategy because they give a pure play on the VRP. A variance swap is an over-the-counter product and an agreement to exchange the realized variance for a fixed amount corresponding to the implied variance. No investment is needed to enter the contract (however, typically the parties will agree to maintain a margin). Conveniently, the model free implied variance as already calculated has been shown, by no arbitrage argument, to be similar to the fixed leg of a variance swap (Carr & Wu, 2009). A synthetic variance swap can be constructed with a static portfolio of vanilla options and a dynamic position in

[7] It is mainly high correlation between IV and other variables that poses a problem.

the underlying. A more extensive elaboration can be found in the appendix. We will use this synthetic variance swap in our trading strategy. The payoff of the variance swap used in the trading strategy is specified as:

$$Variance\ swap = N\left({RV_{t,t+22}}^2 - {IV_t}^2\right) \qquad (23)$$

where N is the notional amount. In the industry, the realized leg is often calculated as the squared daily log returns. The specification applied in this study, using realized kernel returns, is more theoretically consistent with the fixed leg we are using in the swap. Since we want to test a theoretical relationship, not practical applicability, this seems appropriate. Nevertheless, the details in the choice of specification should not alter the main results. The strategy serves as good test for the predictability of the VRP as the return is entirely determined by the VRP. The practical relevance will be discussed later.

The concept of the strategy is to vary the exposure to the VRP according to the information of one of the Bermuda variables (only one is chosen for simplicity and clarity); in this case we have chosen the IVP^W. We want to test whether the value of IVP^W can be used as a decision rule for when to take variance risk. The following formulas are used to create four different strategies, by determining the time varying exposure to the variance swap:

***Benchmark*:** $NE_{t+1} = \pm 0.02V_t$

Active Harvester: $NE_{t+1} = \pm 0.02V_t\left(1 \pm 2(IVP_t^W - \overline{IVP^W})\right)$

Tactician: $NE_{t+1} = \pm 0.02V_t\left(2(IVP_t^W - \overline{IVP^W})\right)$

Steamroller: $NE_{t+1} = \pm 0.02V_t\left(1 \pm 23.5(IVP_t^W - \overline{IVP^W})\right)$

Where NE_{t+1} is the notional exposure that is entered into on day t+1 (or at the end of day t). V is denoted as the realized value of the portfolio, implying that V will remain constant until the first swap has expired. The minus sign indicates the necessity of being on the short side (selling the fixed amount) of the swap contract in order to get (positive) exposure to the VRP. Intuitively, the strategies can be thought of as entering into a new variance swap every day with a notional amount corresponding to NE. The *Benchmark* enters a new variance swap every day with (negative) notional exposure equal to 2 % of the realized portfolio value, reaching a ceiling of 44 % after 22 days. The exposure is calculated as a portion of the realized portfolio value, and the portion (0.02) is rather arbitrarily chosen. It essentially determines the size of the buffer to cover losses from engaging in the strategies. The adequacy of the buffer will be discussed ex post. The 22 first trading days are used to build up the portfolio, while the 22 last are used to wind it down.

Whether the IVP_t^W term should be added or subtracted depends on the results of the forecast (the sign of the coefficient, β_{IVP}^W). A negative β_{IVP}^W indicates that the exposure to the VRP should be increased (by increasing negative notional exposure) if IVP_t^W is positive. It is multiplied with two in order to get a more noticeable impact on the exposure.

The $\overline{IVP^W}$ term, which is the average IVP^W over the full sample, is subtracted in order to make the strategies easier to compare. IVP^W is necessarily above zero on average because of the presence of VRP. Hence, the adjustment makes sure the active bets made by IVP^W are not increasing the average exposure. In turn, this ensures that any excess return over the *Benchmark* is not due to taking increased variance risk (or at least increased exposure to VRP). The fact that the value of IVP^Wcannot be determined ex ante is irrelevant; it does not alter nature of the active bets made by IVP^W.

There is a deviation from testing the predictability of the VRP as compared to the regressions, in that it is not the VRP that is a component in the regressions, but the square root of it. However, we refrain from squaring the IVP^W variable. Even though it could be easily done, it would mostly be a complicating aspect (the result would be similar), since it would mean that the IVP^W would have to be squared, with its sign kept, and then deduct the squared mean. More importantly, the exponential decision rule would make it hard to see the effect of the decision rule except when IVP^W takes extreme values.

The *Benchmark* represents a passive investor who wants to harvest the VRP but does not make bets to improve returns. The second strategy has varying exposure with average exposure similar to the *Benchmark* and is referred to as *Active Harvester.* The strategy essentially holds the *Benchmark* on average, but deviate according to IVP^W. Theoretically, it can be seen as consisting of a strategic allocation ($-0.02V_t$), which is the *Benchmark,* and a tactical allocation ($\pm 0.02V_t\left(2(IVP_t^W - \overline{IVP^W})\right)$), which is the decision-rule based on IVP^W. Active Harvester harvests the VRP through its strategic allocation, and makes additional active bets. Any excess return over the *Benchmark* is a sign that IVP^W has predictive power of the VRP. The third strategy, *Tactician*, has average exposure equal to zero. Here, the strategy has a strategic allocation of zero, implying that the investor do not believe in the presence of VRP. Hence, on average the *Tactician* does not want exposure to the VRP. Instead, tactical bets are made whenever the decision rule allows for it. The outcome of this strategy will show the pure effect of using IVP^W as a decision-rule to make active bets.

To get a better picture of the performance of the IVP^W variable, a fourth strategy, *Steamroller*, is constructed ex post with the same semivariance as the *Benchmark* (which happens to be when the factor of IVP^W is increased from 2 to 23.5). Thus, according to the semivariance, the two strategies have the same downside risk, which should make for an interesting comparison of the performances. It is assumed that every trading strategy starts out with an initial capital, V, of $100.

In the evaluation of the strategies we employ the Sharpe ratio and Sortino ratio:

$$Sharpe\ ratio = \frac{\overline{r_p} - r_f}{\sigma_p} \qquad Sortino\ ratio = \frac{\overline{r_p} - r_f}{\sigma_p^S}$$

where p is referring to the portfolio in question. In the Sortino ratio the standard deviation is only estimated from negative returns and is denoted σ_p^S (square root of semivariance). It is the relative performance of the strategies that is of most importance in our study, rather than the absolute performance. Since we do not add interest to the portfolio holdings when calculating returns, it may already be regarded as subtracted.

Log returns of the strategies are calculated in the following way:

$$r_t = LN\left(\frac{V_t}{V_{t-1}}\right) \qquad (24)$$

5. Results

There are two main objectives in the assessment of the results. One is to evaluate the relative forecast performance of *Bermuda**. The other is to evaluate the economic interpretation of the *Bermuda**, with particular focus on the link to the VRP. In order to get there, we will start by taking a look at the risk neutral moments we have derived. Then, we cover the relative forecast performance of the *Bermuda** model. Once the accuracy is assessed, we move on to the findings concerning the risk neutral moments and the Bermuda variables. Here, we start by taking a look at the information content of the risk neutral moments in a forecasting context. In the end, we go through a more extensive analysis and interpretation of the Bermuda variables, including the results from the trading strategy.

5.1 The Risk Neutral Moments

The smiles for every day are plotted for the whole sample period (Figure 15 in the Appendix), for a range of strikes spanning from 0.3 to 1.7. The level of the smile reflects the periods of turmoil in the market. Figure 9 shows a shorter time period for a shorter range of moneyness where the dynamics of the smiles are more visible.

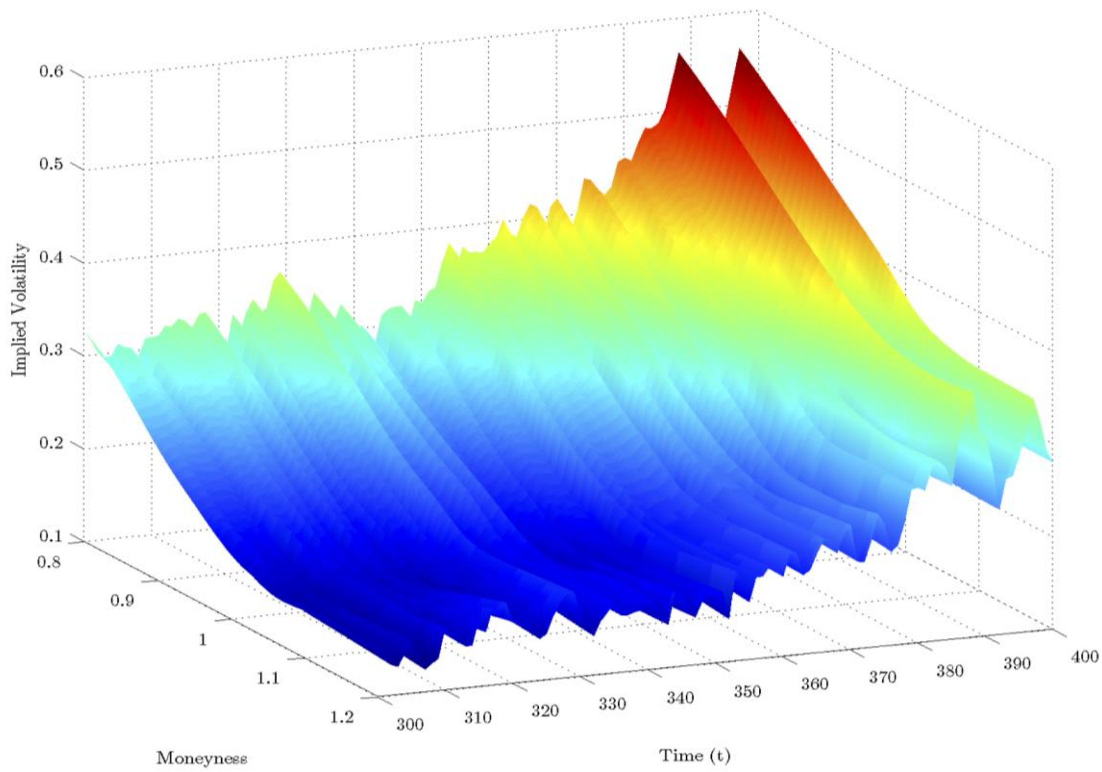


Figure 9. Dynamics of the volatility smiles

These smiles transform into the RNDs (risk neutral densities) shown in Figure 16 in the Appendix. The height of the peak of the RND corresponds with the level of IV, where a higher peak reflects lower IV. A close-up view at the RNDs is given in Figure 10.

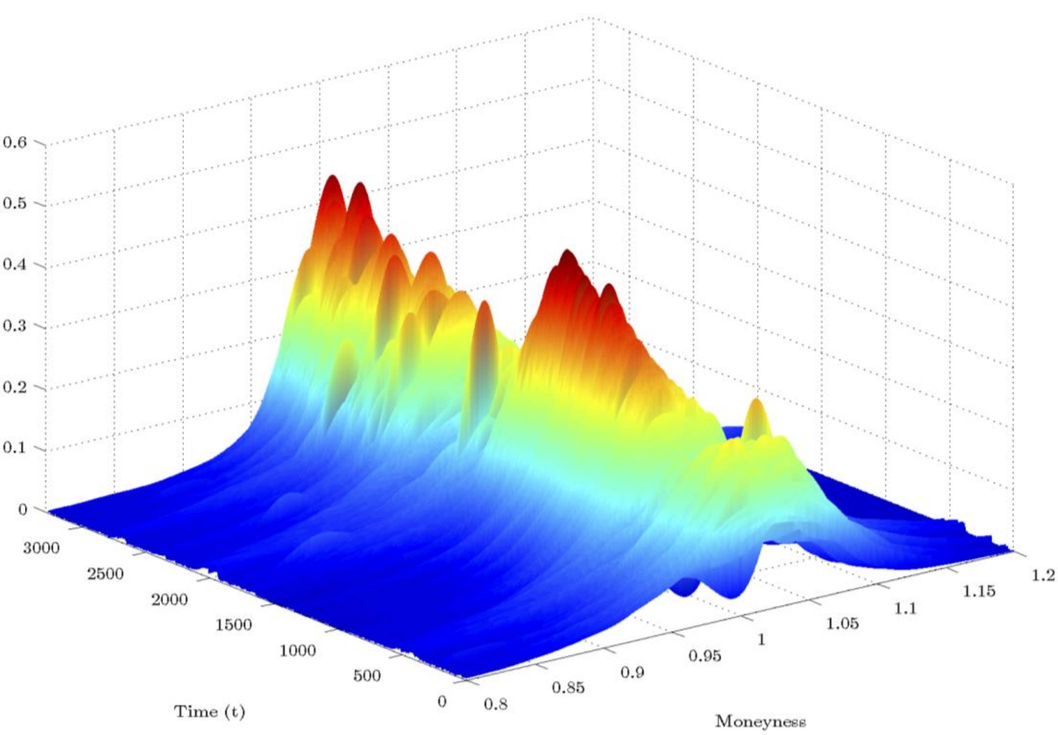


Figure 10. Risk neutral densities for the whole sample period. The densities reflect a 30 days horizon.

Some of the RNDs have more irregular shapes than others. Typically, the more irregular shapes are found in high volatility environment, often caused by the shape of the smile in the transition between using puts and calls. Abrupt changes in the RND for small changes in the moneyness, and sudden large negative areas in the RND could indicate that some strikes are priced unreasonably high compared to others. This could possibly adversely impact the estimations of the risk neutral moments. However, for the most part, the RNDs resemble those of the physical densities close enough to indicate that our derived prices are in general consistent with one another and represent a reasonable market view.

We find that the IV is higher than the RV, which goes along with the previous literature of the existence of a VRP. Furthermore, the average risk neutral skewness is negative and the risk neutral (excess) kurtosis is positive.

While studying the values around the financial crisis it appears to be a tendency that kurtosis is lower in high volatility environments. Since higher volatility typically means a shift in the density from the middle and towards the shoulders, it makes sense that IV and kurtosis are negatively correlated. Similarly, the skewness is slightly less pronounced in very high volatility environments. Further descriptive statistics of the derived risk neutral IV are found in Table 2, and the statistics for skewness and excess kurtosis are found in the Appendix.

Table 2. Descriptive statistics of the implied volatility

	Full Sample	Pre-Crisis	Banking-Crisis	Post-Crisis
Observations	3,245	1636	503	1106
Max	0.873	0.455	0.873	0.513
Min	0.097	0.097	0.157	0.110
Mean	0.211	0.186	0.319	0.198
Median	0.186	0.165	0.261	0.179
Range 5%-95%	0.113-0.396	0.109-0.337	0.182-0.63	0.125-0.336

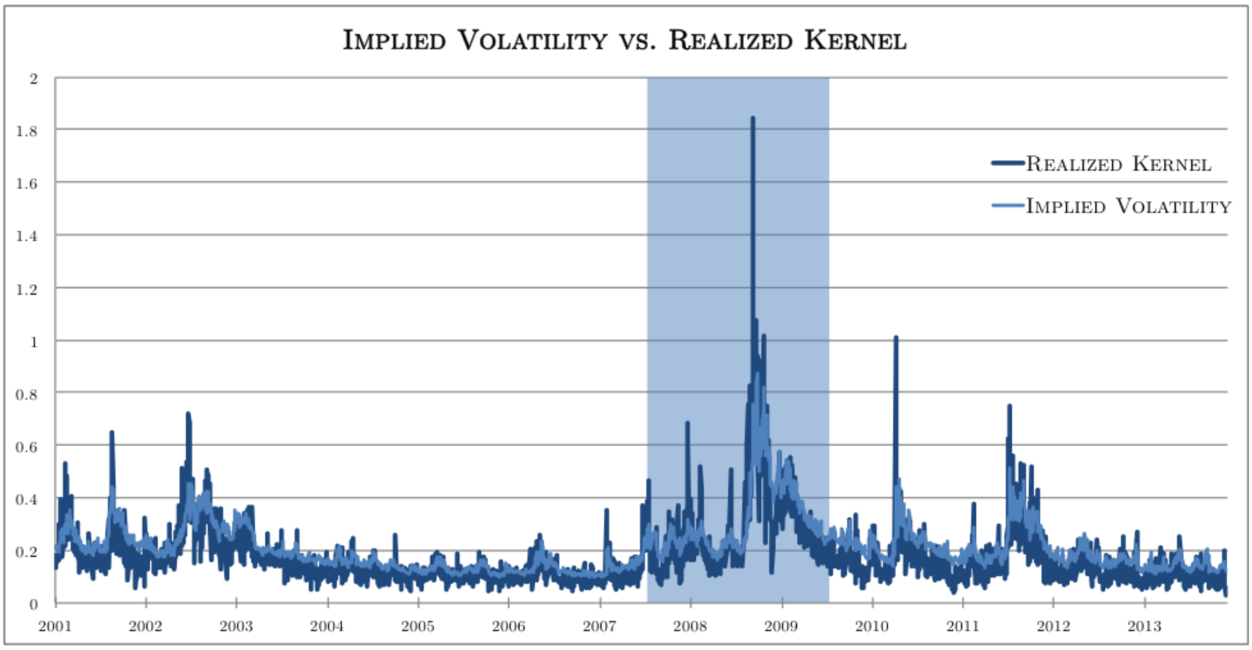


Figure 11. Implied volatility vs. realized kernel

5.2 In-Sample Performance

The Bermuda* model is superior to the other evaluated models for every forecast horizon for the full sample and every sub sample. The *Extended** model outperforms the *Conventional* consistently. This holds for the full sample and all subsamples. For the daily forecast, the adjusted R-squared for *Bermuda** improves relative to *Extended** from approximately 76% to 84%, a non-trivial increase of roughly 8%. For the weekly horizon *Extended** obtains a R-squared of 80 %, whereas Bermuda* has a R-squared of 84%. For the monthly horizon, the R-squares are 69% and 71% respectively. For the daily forecast, in the after crisis subsample, the improvement is astonishing 12%, from approximately 60% to 72%. Looking at the subsamples for the monthly horizons, the performance of the models are quite similar, except for the crisis subsample where the R-squares amount to 61% and 67%, in favor of the *Bermuda** model. The size of the VRP has likely been changing more dramatically in the two last subsamples, which is impacted by the financial crisis (see Figure 12). In this context, it is interesting that the *Bermuda** performs relatively better in the two last subsamples – as it may indicate that the effectiveness is greater when there are large changes in the VRP. For uncertain reasons, the accuracy of all models is substantially lower in the most recent subsample. It is reasonable that the absolute and relative accuracy decline for longer horizons since the models try to capture the short-term dynamics of volatility and they are not designed for long-run predictions. The results are seen in Table 3, 4 and 5.

5.3 Out-of-Sample Performance

The result for the out-of-sample forecasts confirms the impression we already got from the in-sample performance. Looking at MSE, *Bermuda** is performing better for all forecast horizons. The out of sample performance can be seen in Table 6. Concerning criteria measures, a lower value implies a more accurate forecast. For the QLIKE measure *Bermuda** scores best for the daily horizon only. The QLIKE measure penalizes outliers less. Hence, it may look like the *Bermuda** model is relatively better at predicting where the models in general perform badly (in more extreme market conditions when there are large changes in volatility), while the *Extended** models are relatively good in conditions where forecasts in general are accurate (in normal conditions when the volatility is varying less).. Assuming the *Bermuda** model adjusts for the VRP, it could be an indication that the VRP is taking more extreme values during extreme market moves, as this seems to be when the larger adjustments are made by the Bermuda variables. From Figure 14, which shows the VRP and IV plotted together, it is clearly visible that larger changes in VRP is indeed related to larger changes in IV, and hence larger moves on the S&P 500. The general tendency for both in-sample and out-of-sample forecasting is that *Bermuda** is significantly better for the shorter horizons, while the accuracy for the monthly horizon is more similar.

Table 3. In-sample forecast (HAR-RV: *Conventional* model)

Time periode	Full Sample			Pre crisis			Crisis			Post Crisis		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.7371	0.7832	0.6791	0.7180	0.7920	0.6198	0.7068	0.7315	0.5911	0.5570	0.6072	0.5254
$ß_0$	0.0086**	0.0152***	0.0319***	0.0082**	0.0138***	0.031***	0.0193	0.0344**	0.0798***	0.0134***	0.0229***	0.0441***
	(2.11)	(2.69)	(4.23)	(2.24)	(2.7)	(4.91)	(1.63)	(2.14)	(3.14)	(3.27)	(4.08)	(4.07)
$ß_{RV}{}^{D}$	0.392***	0.2813***	0.1825***	0.2892***	0.2239***	0.1552***	0.4309***	0.3312***	0.2083***	0.407***	0.2374***	0.1382***
	(6.1)	(7.47)	(7.6)	(5.47)	(7.01)	(4.62)	(3.18)	(4.59)	(5.31)	(4.75)	(3.69)	(4.09)
$ß_{RV}{}^{W}$	0.4037***	0.4053***	0.3366***	0.5218***	0.5195***	0.3129***	0.4069**	0.3862***	0.3812**	0.2852***	0.3184***	0.2511***
	(4.55)	(6.08)	(3.19)	(5.68)	(5.44)	(3.35)	(2.42)	(3.19)	(2.02)	(2.82)	(5.35)	(4.83)
$ß_{RV}{}^{M}$	0.1549***	0.2263***	0.287***	0.139**	0.1723**	0.3152***	0.0959	0.163	0.1336	0.2161**	0.2875***	0.3093***
	(3)	(3.42)	(3.23)	(2.54)	(2.39)	(3.51)	(1.14)	(1.34)	(0.68)	(2.44)	(3.73)	(2.93)

Table 3 shows the in-sample performance of the HAR-RV model, also referred to as the *Conventional* model.
***Significant on a 10 % level**
****Significant on a 5 % level**
*****Significant on a 1 % level**

Table 4. In-sample forecast (HAR-RV-RNM: *Extended model)**

Time period	Full Sample			Pre crisis			Crisis			Post Crisis		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.7600	0.8037	0.6910	0.7429	0.8212	0.6573	0.7450	0.7651	0.6080	0.6020	0.6410	0.5620
$ß_0$	0.0104	0.0276**	0.0469*	0.0001	0.0122	0.0166	0.0127	0.0696	0.1974	-0.0526***	-0.0381**	-0.0027
	(1.3)	(2.41)	(1.96)	(0.01)	(1.26)	(0.92)	(0.38)	(1.22)	(1.46)	(-3.09)	(-2.02)	(-0.1)
$ß_{RV}{}^{D}$	0.2711***	0.1784***	0.1111***	0.2035***	0.1402***	0.0664**	0.2126	0.1555**	0.1519***	0.2488***	0.124**	0.0806**
	(4.24)	(4.99)	(3.78)	(3.91)	(4.26)	(2.05)	(1.35)	(2.17)	(3.07)	(3.04)	(2.14)	(2.37)
$ß_{RV}{}^{W}$	0.3232***	0.3391***	0.2924**	0.393***	0.3944***	0.1777	0.3562**	0.3566**	0.4015*	0.1383*	0.2126***	0.1969***
	(3.54)	(4.5)	(2.34)	(5.01)	(4.44)	(1.58)	(2.07)	(2.47)	(1.82)	(1.8)	(4.18)	(4.32)
$ß_{RV}{}^{M}$	-0.1964***	-0.0764	0.0731	-0.1143**	-0.0751	0.0646	-0.4084***	-0.2422	-0.0029	-0.1872*	0.0146	0.1945*
	(-2.9)	(-0.88)	(0.72)	(-2.02)	(-1.03)	(0.67)	(-2.59)	(-1.23)	(-0.01)	(-1.7)	(0.15)	(1.76)
$ß_{IV}$	0.6104***	0.5251***	0.3736***	0.4763***	0.4619***	0.4715***	0.9321***	0.7418***	0.2248	0.6938***	0.4821***	0.2579***
	(10.04)	(9.57)	(3.89)	(8.06)	(6.84)	(4.79)	(4.87)	(4.92)	(0.95)	(7.81)	(5.77)	(2.89)
$ß_{SK}$	0.0281**	0.0374***	0.0373	0.0006	0.0061	-0.0134	0.0182	0.065	0.1585	-0.0397**	-0.0355*	-0.0069
	(2.52)	(2.68)	(1.33)	(0.07)	(0.53)	(-0.64)	(0.46)	(1.09)	(1.14)	(-2.12)	(-1.92)	(-0.2)
$ß_{KU}$	0.0022	0.0035**	0.0043	-0.0011	-0.0006	-0.0035	-0.0102*	-0.0007	0.0214	-0.0052**	-0.0035*	0.0026
	(1.51)	(2)	(1.18)	(-0.79)	(-0.37)	(-1.12)	(-1.69)	(-0.08)	(1.09)	(-2.38)	(-1.64)	(0.51)

Table 4 shows the in-sample performance of the HAR-RV-RNM model, also referred to as the *Extended model.**
***Significant on a 10 % level**
****Significant on a 5 % level**
*****Significant on a 1 % level**

Table 5. In-sample forecast (HAR-IVP-RNM: *Bermuda model)**

Time period	Full Sample			Pre crisis			Crisis			Post Crisis		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.8418	0.8381	0.7099	0.8363	0.8558	0.6651	0.8245	0.8076	0.6733	0.7215	0.6794	0.5698
$ß_0$	0.016***	0.039***	0.063**	0.0047	0.0187**	0.0252	0.0782**	0.1441**	0.304*	-0.0007	0.0038	0.034
	(2.59)	(3.63)	(2.47)	(0.79)	(2.12)	(1.45)	(2.44)	(2.27)	(1.8)	(-0.05)	(0.22)	(1.63)
$ß_{IVP}{}^{D}$	0.1417**	0.0443**	0.0433***	0.1785***	0.0361*	0.0169	0.1288	0.0279	0.0113	0.1364*	0.0665**	0.0564***
	(2.45)	(1.98)	(3.01)	(4.41)	(1.69)	(1.19)	(1.06)	(0.64)	(0.48)	(1.76)	(2.1)	(2.79)
$ß_{IVP}{}^{W}$	-1.2102***	-0.6327***	-0.3416**	-1.2583***	-0.6972***	-0.2071	-1.1826***	-0.6331***	-0.377*	-1.2007***	-0.5213***	-0.2971***
	(-10.03)	(-9.12)	(-2.41)	(-13.71)	(-7.87)	(-1.42)	(-5.56)	(-5.59)	(-1.73)	(-5.29)	(-4.46)	(-3.24)
$ß_{IVP}{}^{M}$	0.1221*	-0.2048	-0.3855**	0.1328**	-0.0762	-0.2492	0.0863	-0.3161	-0.7957*	0.1204	-0.2352**	-0.3072**
	(1.73)	(-1.59)	(-2.07)	(2.17)	(-0.76)	(-1.46)	(0.68)	(-1.27)	(-1.78)	(1.31)	(-2.09)	(-2.08)
$ß_{IV}$	0.99***	0.9424***	0.8217***	0.9898***	0.9354***	0.7781***	0.9934***	0.9075***	0.637***	0.9429***	0.8364***	0.7029***
	(53.33)	(32.28)	(18.45)	(50.6)	(28.8)	(12.95)	(28.2)	(15.97)	(5.72)	(25.51)	(23.1)	(16.96)
$ß_{SK}$	0.0162**	0.0347***	0.0424*	0.0036	0.0107	-0.0072	0.0835**	0.1388**	0.2663	-0.009	-0.0084	0.0189
	(2.31)	(3.06)	(1.72)	(0.53)	(1.07)	(-0.35)	(2.24)	(2.19)	(1.62)	(-0.72)	(-0.53)	(0.73)
$ß_{KU}$	0.0018*	0.0038**	0.0052	0.0002	0.0005	-0.0025	0.0101	0.0197*	0.0483*	-0.0011	-0.0002	0.0054
	(1.88)	(2.54)	(1.54)	(0.25)	(0.34)	(-0.86)	(1.46)	(1.95)	(1.77)	(-0.73)	(-0.13)	(1.48)

Table 5 shows the in-sample performance of the HAR-IVP-RNM model, also referred to as the *Bermuda model.**
***Significant on a 10 % level**
****Significant on a 5 % level**
*****Significant on a 1 % level**

Table 6. Out-of-sample forecast

Forecast model	HAR-RV-RNM (Extended*)			HAR-IVP-RNM (Bermuda*)		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly
MSE	4.511	1.991	1.936	2.990	1.644	1.691
QLIKE	-2.708	-2.723	-2.639	-2.747	-2.719	-2.617

Table 6 shows the out-of-sample forecast of the HAR-RV-RNM model and HAR-IVP-RNM model. The lower the values, the more accurate the forecasts.

5.4 Information Content of Risk Neutral Moments

For the *Extended* model the results reveal that skewness has some significance, but it drops once the kurtosis is added, suggesting that the information is already subsumed in the skewness. The results of the regressions are found in Table 12 and 13 in the Appendix. Byun & Kim (2013) have found evidence of the same for S&P 500. The opposite effect occurs for *Bermuda*. When kurtosis is added, the significance of skewness increases, suggesting negative confounding. This would be expected if skewness and kurtosis have an offsetting effect such that the skewness is partially masked by the kurtosis. In general, there is a positive impact on the R-squared when adding the risk neutral moments, even though the effect is not very big. The risk neutral moments are generally more significant in the *Bermuda** model, than in the *Extended** model. In general, the skewness and kurtosis have varying statistical significance, while the economic significance is overall low compared to the other explanatory variables. The IV is overall the most significant among all variables, but the t-stat of the IV is substantially lower in *Extended** than in *Bermuda**. As expected, the lagged RV variables drive down the significance of IV.

Studying the forecasts of semivariance and bipower variation there is particularly one finding of interest (see Table 14-19 in the Appendix). The risk neutral moments (IV, SK and KU) become more significant while predicting the counterpart of semivariance, i.e. the volatility of positive returns. As previously mentioned, research has found skewness to predict future returns. A presumptive thought in line with research is that the skewness and kurtosis may explain the variability in positive returns; while the negative returns as represented by semivariance is more unpredictable from the perspective of using risk neutral moments in forecasts. Concerning the result of the bipower variation, there is a distinctive difference in the significance of the risk neutral moments between the bipower variation (continuous component) and the counterpart (jump component). The former shows substantially higher significance, also higher than in the normal in-sample regressions.

5.5 Interpretation of the IVP Component

It can be inferred from Table 4 and 5 that β_{IV} is different in *Extended** and *Bermuda**. This leads to the conjecture that multicollinearity is present in one of the models. A quick test of running the dependent variable with only IV as explanatory variable (Table 7) shows that the coefficients are very close to the values of β_{IV} in the *Bermuda** regressions (Table 5). The variables change more drastically in *Extended**. One possibility could be that *Extended** better adjusts for the VRP than *Bermuda*,* but this is unlikely as *Bermuda** performs better.

Table 7. Forecast result of implied volatility

Time periode	Full Sample		
Forecast horizon	Daily	Weekly	Monthly
Adj. R^2	0.6986	0.7608	0.6567
$ß_0$	-0.0437***	-0.0318***	-0.0057
	(-5.44)	(-3.26)	(-0.45)
$ß_{IV}$	1.0689***	1.0124***	0.8775***
	(24.5)	(18.86)	(11.6)

Table 7 shows the result when forecasting volatility with only IV. The result is useful in order to see how the coefficient of IV changes for the other models.

As could be presumed, because of the high correlation with past volatility and IV, multicollinearity appears to influence the β_{IV} in *Extended*,* while it seems to have very limited impact on *Bermuda**. This has no implications for the forecasting of volatility per se. Concerning interpretations of the coefficients, and thus the ability to predict the VRP, on the other hand, it is highly relevant. The limitation of the *Bermuda** model is the high correlation among the Bermuda variables. Although this is of lesser concern, it makes it more challenging to interpret the effect from the different Bermuda variables in isolation.

The significance of the Bermuda variables is overall high, as can be seen from Table 5. The β_{IVP}^{M} is the least significant, and for the weekly horizon for the full sample, it is not found significant on a 10 % level. While this indicates that the variable has little incremental information to the other variables, it is worth bearing in mind that the correlations between the Bermuda variables are high.

The signs of the coefficients of the Bermuda variables are consistent over all sample periods. β_{IVP}^{D} is always positive, β_{IVP}^{W} is always negative and the sign of β_{IVP}^{M} changes depending on the forecast horizon. In the following we will mainly interpret the β_{IVP}^{W}, but the insight is transferable to the other Bermuda variables. Assuming the coefficients are true and not biased, the negative β_{IVP}^{W} coefficient tells us that when IVP^{W} is positive (negative), IV continues to overestimate (underestimate) the next period RV. The β_{IVP}^{W} coefficient therefore suggests a momentum or clustering effect in the VRP. This complements the findings of other papers, such as Bollerslev et al. (2009) who find that the VRP is able to explain a nontrivial fraction of the variation in stock market returns, with a high (low) VRP predicting positive (negative) future returns. Our result is consistent with their findings; positive (negative) future returns and low (high) future volatility are in line.

One way to view the results is that low IVP^{W} could be a sign of relatively low hedging demand, thus making portfolios more vulnerable. Investors will then have to take more drastic measures (bigger trades) to hedge their portfolio once volatility increases. Thus, in such a case, an initial shock to volatility can be seen as having a greater impact on the markets than expected, resulting in a continued period of low VRP. In other words, with less protected portfolios the market participants are more likely to be surprised, leading to a clustering effect. Similarly, when investors have protected their portfolios (high IVP^{W}) they are less likely to be caught on the wrong foot.

Another view, perhaps more consistent with an efficient market, would be that the size of the VRP, and hence IVP^{W}, is determined by discount factor variation (and hence the compensation for taking a given amount of variance risk and the amount of variance risk). Intuitively, the amount of variance risk is expected to be related to the uncertainty of the level of the VRP. In turn, this uncertainty can be seen as related to the uncertainty of the future volatility, i.e. the volatility of volatility (see Carr & Wu (2009)). It is imaginable that such an increase in uncertainty is also associated with the compensation for taking a given amount of variance risk, as increased uncertainty may very well be related to the outlook of the economy and marginal utility of wealth. The results presented so far give little foundation for a further discussion at this point. However, we will revisit the topic during our interpretation of the results from the trading strategy.

One could question whether the Bermuda variables are capturing the dynamics of the VRP. In fact, it could be merely an adjustment of IV due to the fact that monthly maturity of IV is used to predict the volatility of the next day. For example, it is possible that the IV underestimates or overestimates tomorrow’s volatility since a mean reversion is incorporated in the monthly IV. Hence, the Bermuda variables might adjust for this shortcoming. It could explain the relatively stronger improvement in accuracy for the daily forecasts, but fails to illuminate the accuracy improvement for the monthly horizon. In the case of a maturity adjustment, the Bermuda variables would not be very useful in creating a trading strategy, as it would suggest that they are not related to the VRP.

Since β_{IV} is below one and decreases with the forecast horizon, and so does its significance (see Table 7), it suggests that an increase in monthly IV is, to greater extent, a prediction of higher RV the next day rather than the next month. Thus, the decreasing β_{IV} is a sign that the monthly IV (including the variance risk component) is more connected with the short-term RV than the monthly RV. This fuels the theory that the IVP variable is a good proxy for how the market is pricing variance risk, notwithstanding that they are only taking the next day RV into consideration. It fuels the notion that the price investors are willing to pay today is hugely affected by the prevailing volatility conditions. Predominantly, it can be considered a vindication of the aforementioned belief that the forecast improvement is due to a maturity adjustment.

Table 8 displays the results from regressing the Bermuda variables on the (volatility) risk premium component.

Table 8. Forecasting the variance risk premium

Independent variable	VRP			
Explanatory variable	Daily IVP	Weekly IVP	Monthly IVP	All IVP
Adj. R^2	0.0356	0.1051	0.0885	0.1154
β_0	0.0299***	0.0205***	0.0172*	0.0167*
	(4.71)	(2.59)	(1.83)	(1.9)
β_{IVP}^{D}	0.1803**	0	0	-0.0634***
	(2.43)	(0)	(0)	(-3.78)
β_{IVP}^{W}	0	0.4592***	0	0.3921***
	(0)	(3.2)	(0)	(2.94)
β_{IVP}^{M}	0	0	0.557***	0.2442*
	(0)	(0)	(2.99)	(1.84)

Table 8 shows the result from forecasting the (square root of) the VRP, with the Bermuda variables as explanatory variables.

All the coefficients are significant on a 10 % level in the multiple regression, but the only the daily and weekly IVP variable is significant on the 5 % and 1 % level. Further examination shows that the Bermuda variables are highly correlated, which explains the lower significance of the monthly variable. In fact, in the single regression the β_{IVP}^{M} is highly significant (on 1 % level). In general, the outcome of the single regressions shows highly significant coefficients. The results indicate that the Bermuda variables have predictive power over VRP. Turning to the economic significance, it is not trivial to interpret the R-squares, which amount to around 12 % for the multiple regression. While this may seem low, the implications of being able to predict any variation of future risk premium can be substantial in economic terms. Recall that the forecasting results show a substantial improvement prompted by the Bermuda variables. Fortunately, Figure 12 can offer some enlightenment.

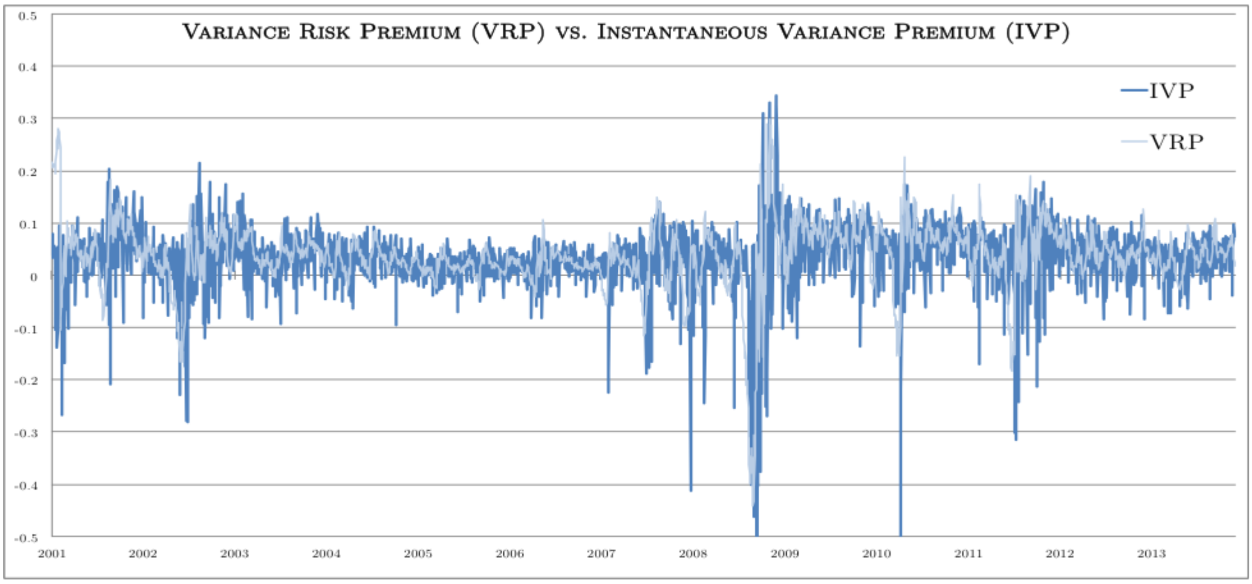


Figure 12. The plot shows the variance risk premium vs. instantaneous variance premium. Two outliers of the variance risk premium (10/10/08; -1.184 and 5/6/10; -0.759) are excluded.

The plot shows the IVP variable plotted against the VRP. Indeed, visually there seems to be a clear relationship. The curves follow each other closely for the highs and lows, indicating that the IVP variable captures the big movements in the VRP. Presumably, the unexplained variance is then mainly related to the smaller variations in the VRP. Recalling the discussion of VRP in the methodology, it is possible that the unexplained variance can be attributed to noise in the estimation. If so, this unexplained variance would not be related to a change in the discount factor, or in the pricing of variance risk. The trading strategy will provide more explicit information on the subject of economic magnitude regarding the predictability of the VRP.

Let us briefly interpret the coefficients of the Bermuda variables conditional on the others. β_{IR}^{W} suggests a momentum effect in the VRP, with a strong statistical and economical significance (Table 7). The high correlation between the weekly and monthly variable of 0.7, indicates that they capture much of the same effect when predicting the VRP. Therefore, we can conclude that there is a momentum effect in the VRP as predicted by both β_{IVP}^{M} and β_{IVP}^{W}. In an environment of high IVP^{M} (high VRP), a day with low IVP^{D} signals, according to the positive β_{IR}^{D}, that the VRP will increase back towards its current moving average (as measured by IVP^{M} and IVP^{W}). The interpretation is that if traders took a low price for variance risk yesterday, compared to average conditions, they will take a higher price going forward - one closer to the average.

The performance of *Bermuda** relative to *Extended** drops slightly when forecasting bipower variation (Table 15 and 16). While the effect is small, it may indicate that the *Bermuda** model are better at taking into account jump activity. A possibility is that the VRP is somewhat related to jump activity, and the Bermuda variables are thus able to adjust for the change in IV coming from the VRP due to jump activity. Whereas the lagged RV variables are insignificant for the jump component, the Bermuda variables are more significant, pointing in the direction that the above relationship is true. This would be in line with the findings of Todorov (2010): that VRP is related to jump activity.

Another observation is that the Bermuda variables are more significant when forecasting the semivariance than when forecasting the counterpart. This fits well with our previous observations, which suggested that *Bermuda** is particularly good at predicting volatility in times where the volatility is difficult to predict (when accuracy is low), i.e. when the market is falling.

5.6 Concluding Remarks on Volatility Forecasting

According to our results, *Bermuda** is a better specified model and superior at forecasting volatility relative to the other models. The outperformance of the forward-looking *Bermuda** model is interesting in the context of market efficiency, but also in the sense that it gives an indication of the effectiveness of the method for adjusting for the VRP. Apart from that, the relative performance of the models is of lesser interest, since the purpose of the analysis is not the select the better model. If optimizing for forecast accuracy, the natural step would be to test a model including both the RV variables, IV and the Bermuda variables. However, we deem the improved analysis of the VRP a more valuable feature than forecast accuracy. The ability to interpret the coefficients of the *Bermuda** model brings us to one of the key differences between the *Bermuda** model and other volatility models. Many volatility models are designed to achieve the highest possible accuracy, with little regard to the economic interpretations of results. The value in *Bermuda** lies not in the excellent accuracy alone; the economic intuition and interpretation should be equally appraised.

Regarding the risk neutral moments, we find that skewness and kurtosis in some cases are significant even on a 1 % level, but the economic impact is tiny as measured by the relative low impact on IV the R-squared.

Altogether we find that the Bermuda variables are linked to the VRP because (1) the Bermuda variables are significant even for monthly horizon, (2) *Bermuda** generally performs better than other models for the monthly horizon, (3) the Bermuda variables have significant explanatory power over the risk premium, (4) our empirical evidence indicates that the monthly IV is more connected to the short-term volatility than the monthly volatility and (5), as will be explained later, the results from the trading strategy provide further support. We find that the Bermuda variables are able to predict VRP, while further investigation is necessary to determine the economic magnitude more accurately. Naturally, it is possible that the three time scales used (1, 5 and 22) not necessarily reflect optimal time scales for predicting the VRP.

The predictability of VRP is useful both in the context of volatility forecasting and in the context of predicting returns of variance risk strategies. The former has been demonstrated in the forecasting session, while the latter will be gauged in the following.

5.7 Trading Strategy

Active Harvester outperforms the *Benchmark*, confirming that the decision rule generates positive return and that the Bermuda variables predict the VRP (see Table 9). Similarly, *Tactician* generates a positive return. At first glance, the improvement of *Active Harvester* compared to *Benchmark* appears modest, bearing in mind it is over 13 years. The excess return over *Benchmark* is only around 10%, which is less than 1% annualized. Still, the erroneous inference from judging too quickly is severe; the predictive ability of IVP^W is in fact impeccable. As we will see, the small improvement is merely because IVP^W has little impact on the exposure in its current format. The portfolio value of *Tactician* seen in Figure 24 in the Appendix, which has zero exposure on average, illustrates the point well: almost every major decision made by IVP^W is correct. The strategy barely makes any losses of magnitude[8]. The strong performance can also be observed from the Sortino ratio in Table 9, which is higher than the benchmark.

[8] While IVP^W is used to demonstrate a strategy in this case, running the same strategies with $\mathrm{IVP}^{\mathrm{M}}$ gives a confirming picture.

Table 9. Performance of trading strategies

	Benchmark	Active Harvester	Tactician	Steamroller	S&P 500
Annualized return	6.3%	7.0%	0.6%	13.8%	2.4 % (4.6 %)
Variance (daily)*	0.10	0.07	0.01	0.49	
Semivariance (daily)*	35.83	18.48	0.17	33.78	
Sharpe ratio	0.24	0.33	0.10	0.23	
Sortino ratio*	4.06	6.30	5.81	8.89	
Correlation S&P 500	0.24	0.23	-0.16	-0.09	1.00
Skewness	-4.57	-2.82	9.70	8.91	
Mean allocation of capital to VRP	44%	44%	0%	44%	
Correlation between Benchmark and Tactician			-0.68		

Table 9 shows various statistics for the trading strategies. The calculations are based on daily log returns, while annualized returns are annualized geometric returns. The return for S&P 500 is the price return adjusted for dividends and splits, and is used to calculate the correlations. With dividends reinvested, the return increases to 4.6 %. *Mean allocation of capital to VRP* is calculated as the total notional exposure to variance swaps in relation to portfolio value (V)

5.8.1 Performance

The get a deeper grasp of the abilities of IVP^W, an evaluation of the performance of the strategies is conducted. The dollar value of the *Benchmark* is increased to $220, corresponding to an annualized return slightly above 6% over the period. *Active Harvester* achieves a portfolio value of $240 and an annualized return of close to 7%. Tactician ends up with a more modest value around $108, an annualized return of 0.6%. *Steamroller* lives up to its name with a final portfolio value of $535, over a fivefold increase in portfolio value. The annualized return amounts to almost 14%. The biggest loss arises from not being able to predict the reversal in the momentum effect of VRP during the financial crisis.

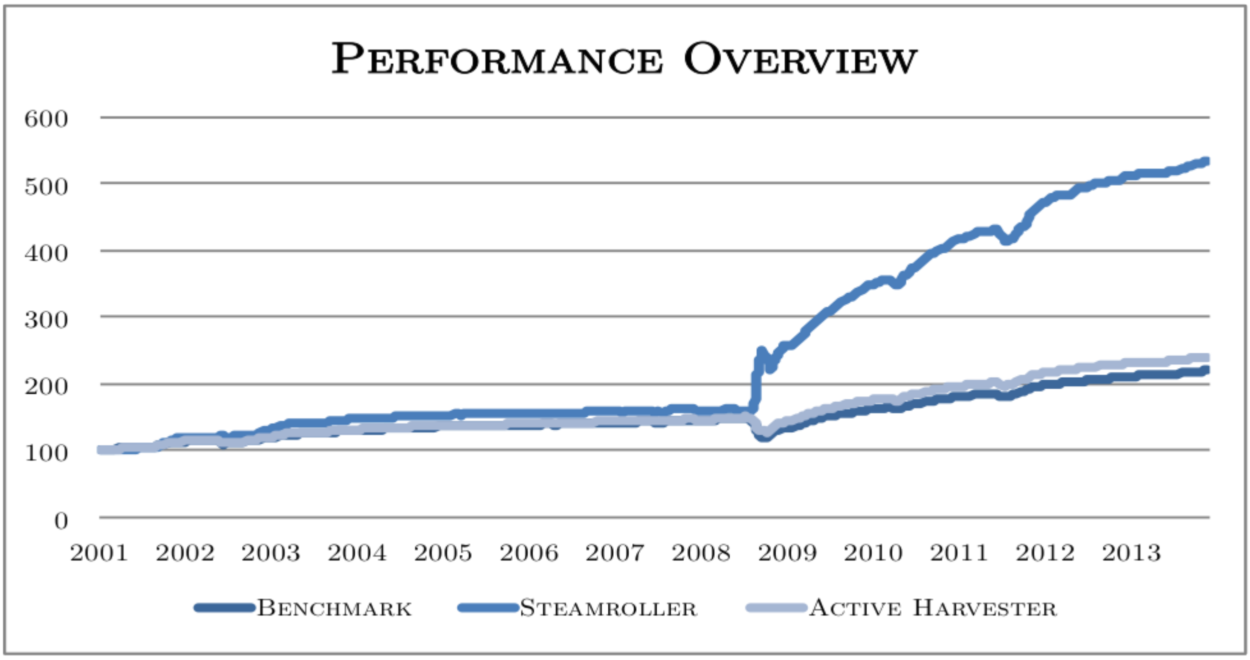


Figure 13. Performance overview of the trading strategies

Turning to risk-adjusted performance, the *Steamroller* does not seem to do very well according to the Sharpe ratio, which shows no improvement over the *Benchmark*. However, the Sharpe ratio is deluding; bearing in mind that the strategy has the same semivariance as the *Benchmark*, yet the return is improved drastically. This illustrates the fallacy if using Sharpe ratios in this analysis, as most of the volatility comes from positive returns. A problem of using Sharpe ratios typically emerges whenever there is a degree of skewness in the returns. Indeed, the skewness of the *Benchmark* is highly negative of -5, while the active bets made by IVP^W gives rise to positive skewness. *Active Harvester* has reduced the negative skewness to -3, while *Tactician* has a positive skewness of close to 10. The skewness of *Steamroller* is 9, reflecting highly positively skewed returns. Hence, the Sortino ratios give a much better indication of the performance among the strategies. The ratios are 4.1, 6.3, 5.8 and 8.9 for *Benchmark*, *Active Harvester, Tactician* and *Steamroller* respectively. The risk-adjusted performance of the Steamroller is nothing short of phenomenal relative to the other strategies, for the sample period used in this study.

The *Steamroller* has valuable characteristics beyond its superior risk-adjusted returns, such as positive skewness and negative correlation to the S&P 500. Simply put, in the current sample period it appears that the decision rule provides us with the means to construct a strategy almost without parallel. A higher allocation to the strategic part would serve to demonstrate the point. If the strategic allocation in *Steamroller* is doubled (from 2% to 4%), the annualized returns jumps to over 27%, yielding a final portfolio value of $2350. The Sortino ratio is still at 8.3, suggesting only a very small decline in the risk-adjusted performance. The practical relevance of the absolute returns will be addressed later.

5.8.2 Interpretations

The IVP^W decision rule is a remedy to the highly negatively skewed returns of the *Benchmark* strategy – a natural consequence of taking on variance risk. The decision rule accurately predicts the VRP so that major blows are avoided; the correlation between *Benchmark* and Tactician is -0.7. This fact explains why the tactical allocation compliments the strategic allocation so well. Moreover, it is a sign that the major changes in the VRP are predicted accurately – when the premium is negative, the strategy takes exposure accordingly. The IVP^W makes big bets when it takes extreme values, which happens to be in volatile markets. The decision-rule gives rise to an intuitively smart strategy, one that search for returns during extreme conditions in the market.

Our results shed some light on the idea that the VRP is related to the level of IV. Indeed, since the IVP^W makes big bets during periods of high IV it may insinuate a connection. The financial crisis leads to huge losses for *Benchmark*, but the subsequent return is also high. The period of high IV can be divided in two. The first period appears to be associated with big losses for those taking variance risk, while the second is associated with big returns, as can be seen in Figure 14.

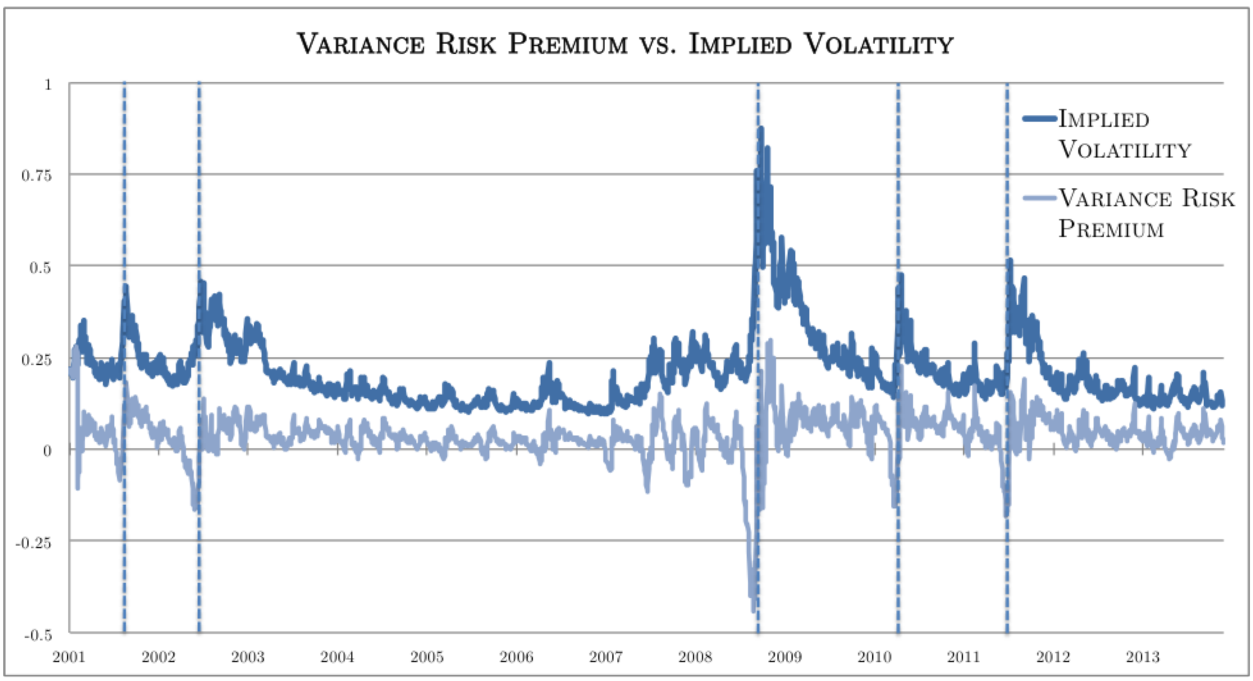


Figure 14. Variance risk premium vs. implied volatility

The transition between these two periods coincides with the turning point of the IV (or RV). Hence, the two periods can roughly be separated by the derivative of IV, giving rise to one period of rising IV and one of falling IV. Being exposed to VRP implies being short option vega. The vega is the option value's sensitivity to volatility. A fall in IV, all else equal, will have a positive impact on a portfolio that is short vega. In practice, this fall in IV has to coincide with a sufficient fall in the volatility to be profitable on average. Evidently, this appears to be the case.

Our findings suggest that the VRP is not necessarily related to the level of IV per se, but to the sign of its derivative - whether it is a period of increasing or decreasing levels of IV. Falling periods of IV is associated with a high VRP, and vice versa. The correlation between the VRP and the first differences of IV is around -0.3, confirming that a decrease in IV is associated with an increase in VRP, and vice versa. This is in line with our previous interpretations of a clustering effect in VRP; a rise in volatility surprises investors and induce higher volatility than expected from the initial shock. This leads to a period of increasing volatility and IV, but the volatility changes more rapidly than the IV. Once the increase in IV comes to an end, it is a sign that investors have repositioned their portfolios. This marks the beginning of a period of positive returns.

Looking at Figure 25 in the Appendix it is evident that the Benchmark and the S&P 500 has higher correlation during extreme market moves, i.e. the financial crisis. This raises the question of whether the equity risk premium and VRP could be correlated during extreme market moves. This could make sense for example if the price investors are willing to pay for taking variance risk are somewhat proportionate to the alternative investments, and hence the returns in the stock market. The point where IV or RV turns from increasing to decreasing reflects the point where the volatility of S&P 500 and presumably discount rates of the economy switches from increasing to decreasing. Hence, with this line of reasoning it makes sense that the transition between negative and positive returns occurs at this point. This could still be in line with the findings of Carr & Wu (2009), who find that the VRP is somewhat related to the volatility of volatility, as long as the volatility of volatility is related to declining markets. Visually, it appears to be the case: the volatility is changing most during larger market falls. However, we have little basis for concluding whether the change in VRP is related to an increased pricing of risk (compensation for taking a given amount of risk) or increased variance risk. The VRP is essentially the returns of a variance swap, which does not necessarily have a constant risk. Hence, changes in the VRP do not have to be connected with changes in market risk aversion or compensation for taking risk. Still, the fact that the *Tactician* has a higher Sortino ratio than the *Benchmark* points in the direction that the uncovered dynamics of the VRP is at least partly related to risk aversion or increased compensation for taking risk, rather than increased risk. On the other hand, given the big impact of events such as the financial crisis, the high Sortino ratio could be a result of "luck" by coincidentally predicting the financial crisis, as the Sharpe ratio is lower for the *Tactician*. This is naturally an uncertainty with our results in general. On the other hand, one can argue it is likely not a coincidence on the basis that the return generated in this period is a result of a number of trades done over many days. Furthermore, the perceived dependency of the few extreme events might be somewhat artificial as it could be due to the fact the decision rule mainly makes big trades on certain occasions, so the other trades are not necessarily wrong, just too small to make an impact. The fact that our forecasts remain superior in all subsamples tested is supportive of this.

5.8.3 Limitations and Relevance

A perhaps striking element of the trading strategies is that the result of the forecast, which may be regarded as "future information" at start of the trading exercise, is used to determine the sign of the impact of IVP^W on the exposure. However, to our defense the same conclusion would be reached in whatever time period the forecasts had been run (and then creating a strategy without "future information"). For the authors, it is even more comforting that such criticism implies that our Bermuda variables in fact are predicting the VRP, which the trading strategy essentially is a test of.

Even though it is not the intention, nor our wish, to create a strategy that is optimized for real life application, it is worthwhile to briefly comment on the practical relevance given the performance. When entering into a variance swap, the two parties will commonly agree to sustain a margin to cover potential losses. In this regard, let us first take a look at the size of the buffers, which ideally should reflect a sufficient amount of capital to cover the margins and losses. Starting with *Benchmark*, a quite large drawdown is visible in Figure 22 in the Appendix. More detailed calculations show a maximum drawdown[9] of around 23% of the portfolio value. The maximum drawdown for *Steamroller* is 14%, giving potential for even more steam. It could be interesting to evaluate the impact if one of the major tactical bets went wrong. Let us assume that the huge increase in portfolio value around the financial crisis were a decline of a similar percentage change. This would give a drawdown of around 60% for the *Steamroller*. A substantial loss, but the buffer would still be sufficient to cover the losses.

Another important aspect concerns the calculation of variance. The returns of the strategies are only calculated from realized returns, which in effect is the return generated over the subsequent month after entering a swap contract. In other words, it does not take into account daily fluctuations such as cash payments that may occur to maintain the agreed upon margin. Hence, it can be questioned whether it is fair to compare the variances of the strategies with daily returns calculated on the S&P 500. For this reason, we have not included variance, Sharpe ratio and Sortino ratio of S&P 500 in the comparison.

[9] A drawdown is the peak-to-trough decline during a specific period of an asset or investment.

In general, we disregard all practical implications related to that variance swaps are traded OTC (for example, many dealers have put a cap on the maximum payout after the financial crisis). The contracts in question are synthetic, and there is no guarantee that our synthetic variance swap accurately reflects quoted swap strikes, even though our variance swap is replicated from a portfolio constructed from market prices of traded products. Firstly, there would be transactions costs associated with the replication. Secondly, there would be a bid-ask spread. Thirdly, jumps and discontinuity of trading play a role. Lastly, our variance swap may be considered a theoretical version and differs slightly from the industry standard in the calculation of the realized leg, as we do not use daily log returns, but instead use a more robust estimate for the realized variance. These factors could undoubtedly affect the returns, yet hardly enough to alter the main interpretations of our analysis. Moreover, the strong performance of strategies exposed to variance risk is in line with previous literature. In their study of using real variance swap data from dealers, Leippold et al. (2007) find that a variance swap strategy can provide smooth returns with higher Sharpe ratios than traditional asset classes. Nevertheless, the main concern of our analysis is not measuring absolute performance of the strategies (although we do devote attention to it as the results appear exceptional).

We advocate that the practical relevance is not dependent on the trading strategy being employable, or that our swap rates are extremely accurate. The purpose of the trading strategy is to test the economic foundation of the IVP variable and the predictability of the Bermuda variables. In this regard, it is only the relative performance of the strategies that matters. Most of the limitations we herein discuss, practically impact all strategies equally.

The verification of our results, which the trading strategy represents, is relevant for anyone who engages in hedging, is exposed to variance risk or cares about risk or risk premium.

6. Conclusion

The *Bermuda** model (HAR-IVP-RNM) outshines conventional models (HAR-RV and HAR-RV-RNM) by providing substantially more accurate volatility forecast both in-sample and out-of-sample. Furthermore, risk neutral skewness and kurtosis show some evidence of explanatory power in forecasting volatility, albeit the impact is limited. We conclude that the Bermuda variables, which the *Bermuda** (HAR-IVP-RNM) model is based on, are able to amend the distortion in IV caused by the embedded VRP in option prices. The excellent forecast performance of *Bermuda** is thus attributed to a refinement of the VRP in IV. The analysis suggests that option implied information is superior to using historical values of RV when forecasting volatility once the VRP is accounted for. However, while these results possibly point in favor of an efficient market view, it is not in any way a rejection of volatility models based on historical data, nor suggesting it is optimal to exclude the RV variables in the model of Corsi (2009) (as is done here for the purpose of the analysis). Above all, this study highlights the importance of adjusting for the VRP when using the information content of IV. The proposed method proves effective in this regard. The IVP variable provides a potent way of approximating the market's recent pricing of variance risk, which in turn is cardinal in capturing the dynamic properties of the VRP. The results are consistent with the very recent paper of Prokopczuk and Simen (2014), who with a different approach are able to improve volatility forecasts in commodity markets by adjusting for the VRP. However, their method is based on capturing the dynamic properties of the VRP.

The constructed trading strategies based on the weekly IVP variable demonstrate convincing precision when predicting the VRP. The risk-adjusted performance is significantly higher than the benchmark. Our results find evidence of a momentum effect in the VRP, despite this effect being rarely observable in the sense that it seems to be tied to large market moves. The inflection point of the momentum effect in VRP appears to be connected with the transition between falling and rising IV and RV – where rising volatility is associated with negative returns for variance risk strategies (a lower VRP). Considering this, in combination with the leverage effect in volatility – that negative returns increases volatility more than positive returns – it makes sense that strategies with exposure to the VRP is positively correlated with the stock market during extreme market moves. This fuels the conjecture that the equity premium and VRP both change in the same direction during extreme market moves, which in turn could indicate that they to some degree are governed by a (systematic) market risk aversion or the marginal utility of wealth in the economy. However, our analysis does not uncover whether the change in VRP is prompted by a change in the pricing of variance risk (related to market risk aversion or marginal utility of wealth) or increased variance risk. Furthermore, the low correlation between VRP and IV over time evinces that the high correlation during extreme market moves, are offset by low or negative correlation during normal conditions. Insofar as this holds true, it could make sense to take this time varying correlation into account when dealing with the VRP. This could explain why known market risk factors are unable to explain the VRP over time, such as in Carr and Wu (2009). That is interesting in the context of the strand of literature that interprets the VRP as some kind of measure of aggregate risk aversion or economic uncertainty. A theory worthwhile exploring in greater detail is thus if the VRP behaves differently than the equity premium (and known market risk factors) under normal conditions, but similar under more volatile markets, when the impact a system wide market risk aversion or marginal utility is more substantial.

Our results further indicate that our method of capturing the dynamic properties of the VRP is particularly effective during volatile markets.

A key feature of the IVP variable is its ability to approximate a monthly VRP that is not yet observable in market data, which could help explain its effectiveness in future predictions of both volatility and VRP. The lack of an up to date estimate of VRP may perhaps also help explain the absence of successful attempts on using VRP in volatility forecasting, despite decades of documentation of its presence and distortions. Analogous to the Bermuda Triangle, where the magnetic field distorts the compass, our approach may provide a convenient way of improving navigation in exposed waters. Hence, the *Bermuda** model provides insight into the long-lived mystery of how to make a magnetic refinement to the volatility compass of financial markets: how to adjust for the risk premium component in IV. Furthermore, the findings concerning the behavior of the VRP are useful in the context of asset allocation and portfolio optimization.

Further Research

The topic of VRP is one that has recently attracted attention from researchers. The results herein motivate further research in using VRP in volatility forecasting as well as in return predictability. Research on how to adjust for the VRP in different ways and on alternative methods for capturing the dynamics of the VRP would be useful. There are many related topics suitable for further research; portfolio allocation and trading strategies using variance swaps or volatility strategies, the behavior of the VRP (also in relation to other risk premiums), determinants of VRP and research on the link between risk neutral moments and physical moments. As already indicated, an avenue for further research could be to explore the determinants and dynamics of the VRP in more normal market conditions.

At the time of writing, the first paper using the VRP for volatility forecasting was published by Prokopczuk & Simen (2014). Given the contemporary popularity of variance as an asset class, along with increased attention on VRP in research, the paper is likely the first of many to come.

Bibliography

Agnolucci, P. (2009). Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. *Energy Economics*, *31*, 316–321.

Andersen, T. G., & Bollerslev, T. (1997). Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns. *Journal of Finance*, *52*(3), 975–1005. doi:10.1111/j.1540-6261.1997.tb02722.x

Andersen, T. G., & Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review*, *39*(4), 885–905.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility. *Review of Economics and Statistics*.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2010). Parametric and Nonparametric Volatility Measurement: Volume 1 - Tools and Techniques. In Y. Ait-Sahalia & L. P. Hansen (Eds.), *Handbook of Financial Econometrics* (pp. 67–137). North-Holland.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, *71*(2), 579–625.

Baillie, R. T., Chung, C. F., & Tieslau, M. A. (1996). Analysing inflation by the fractionally integrated ARFIMA-GARCH model. *Journal of Applied Econometrics*, *11*(1), 23–40.

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, *16*(1), 101–143.

Bakshi, G., & Madan, D. (2006). A Theory of Volatility Spreads. *Management Science*, *52*(12), 1945–1956.

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, *59*(4), 1481–1509.

Banz, R. W., & Miller, M. H. (1978). Prices for State-contingent Claims: Some Estimates and Applications. *The Journal of Business*, *51*(4), 653–72.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2008). Designing Realised Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise. *Econometrica*, *76*(6), 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2009). Realized kernels in practice: trades and quotes. *Econometrics Journal*, *12*(3), C1–C32.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2011). Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics*, *162*(2), 149–169.

Barndorff-Nielsen, O. E., Kinnebrock, S., & Shephard, N. (2010). Measuring downside risk-realised semivariance. *Volatility and Time Series Econo- Metrics: Essays in Honor of Robert F. Engle*, 117–136.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, *2*(1), 1–37.

Barndorff-Nielsen, O. E., & Shephard, N. (2005). *Variation, jumps, market frictions and high frequency data in financial econometrics*. *Department of* (p. 55).

Bianchetti, M. (2010). Two curves , one price. *Risk*, (August), 74–80.

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *81*, 637.

Bliss, R. R., & Panigirtzoglou, N. (2004). Option-Implied Risk Aversion Estimates. *The Journal of Finance*, *LIX*(1), 407–446.

Boenkost, W., & Schmidt, W. (2005). Cross currency swap valuation. *Available at SSRN 1375540*, *49*(2).

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

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

Bollerslev, T., Marrone, J., Xu, L., & Zhou, H. (2012). Stock Return Predictability and Variance Risk Premia: Statistical Inference and International Evidence. *SSRN Electronic Journal*.

Bollerslev, T., Tauchen, G., & Zhou, H. (2009). Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies*, *22*(11), 4463–4492.

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business*, *51*(4), 621.

Busch, T., Christensen, B. J., & Nielsen, M. Ø. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. In *Journal of Econometrics* (Vol. 160, pp. 48–57).

Byun, S. J., & Kim, J. S. (2013). The information content of risk-neutral skewness for volatility forecasting. *Journal of Empirical Finance*, *23*, 142–161.

Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. *The Review of Financial Studies*, *6*(3), 659–681.

Carr, P., & Wu, L. (2009). Variance Risk Premiums. *Review of Financial Studies*, *22*(3), 1311–1341.

Chang, B.-Y., Christoffersen, P., Jacobs, K., & Vainberg, G. (2011). Option-Implied Measures of Equity Risk. *Review of Finance*, *16*(2), 385–428.

Charoenwong, C., Jenwittayaroje, N., & Low, B. S. (2009). Who knows more about future currency volatility? *Journal of Futures Markets*, *29*(3), 270–295.

Cheng, X., & Fung, J. K. W. (2012). The Information Content of Model-Free Implied Volatility. *Journal of Futures Markets*, *32*(8), 792–806.

Chernov, M. (2007). On the Role of Risk Premia in Volatility Forecasting. *Journal of Business & Economic Statistics*, *25*(4), 411–426.

Christensen, B., & Prabhala, N. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, *50*(2), 125–150.

Conrad, J., Dittmar, R. F., & Ghysels, E. (2013). Ex Ante Skewness and Expected Stock Returns. *The Journal of Finance*, *LXVIII*(1), 85–124.

Cont, R. (1998). Beyond implied volatility. *Options*, (July 1997), 26.

Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, *7*(2), 174–196.

Corsi, F., Pirino, D., & Renò, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, *159*(2), 276–288.

Dennis, P., & Mayhew, S. (2002). Risk-Neutral Skewness: Evidence from Stock Options. *The Journal of Financial and Quantitative Analysis*, *37*, 471.

Diebold, F. X., & Hickman, A. (1998). Converting 1-Day Volatility to h-Day Volatility : Scaling by h is Worse than You Think. *Horizons*, 104–107.

Doran, J. S., Peterson, D. R., & Tarrant, B. C. (2007). Is there information in the volatility skew? *Journal of Futures Markets*, *27*, 921–959.

Drechsler, I., & Yaron, A. (2010). What's Vol Got to Do with It. *Review of Financial Studies*, *24*(1), 1–45.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987–1007.

Engle, R. F., & Ng, V. K. (1993). Time-Varying Volatility and the Dynamic Behavior of the Term Structure. *Juornal of Money, Credit and Banking*, *25*(3), 336–349.

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, *38*(1), 34.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, *25*, 383–417. doi:10.2307/2325486

Figlewski, S. (2010). Estimating the Implied Risk-Neutral Density for the US Market Portfolio. In T. Bollerslev, J. Russell, & M. Watson (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle* (pp. 323–353). Oxford University Press.

Fleming, J. (1998). The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of Empirical Finance*, *5*, 317–345.

Frijns, B., Tallau, C., & Tourani-Rad, A. (2010). The information content of implied volatility: Evidence from Australia. *Journal of Futures Markets*, *30*(2), 134–155.

Fujii, M., Shimada, Y., & Takahashi, A. (2011). A Market Model of Interest Rates With Dynamic Basis Spreads in the Presence of Collateral and Multiple Currencies. *Wilmott*, *2011*(54), 61–73.

Goldstein, D. G., & Taleb, N. N. (2007). We Don't Quite Know What We Are Talking About. *The Journal of Portfolio Management*, *33*(4), 84–86.

Gorton, G. (2009). Slapped in the Face by the Invisible Hand: Banking and the Panic of 2007. *Available at SSRN 1401882*, (2008).

Granger, C. W. J., & Joyeux, R. (1980). An Introduction to Long-Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, *1*(1), 15–29.

Guo, D. (1998). The risk premium of volatility implicit in currency options. *Journal of Business & Economic Statistics*, *16*(4), 498–507.

Han, B. (2007). Investor Sentiment and Option Prices. *Review of Financial Studies*, *21*, 387–414.

Hansen, L. P., Heaton, J. C., & Li, N. (2008). Consumption Strikes Back? Measuring Long- Run Risk. *Journal of Political Economy*, *116*(2), 260–302.

Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, *20*(7), 873–889.

Hansis, A., Schlag, C., & Vilkov, G. (2010). The Dynamics of Risk-Neutral Implied Moments: Evidence from Individual Options. *SSRN Electronic Journal.*

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*(2), 327–343.

Hull, J. (2014). *Options, Futures, and Other Derivatives* (9th editio., p. 896). Prentice Hall.

Hull, J., & White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance*, *42*(2), 281–300.

Hull, J., & White, A. (2013). LIBOR vs . OIS : The Derivatives Discounting Dilemma *. *Journal of Investment Management*, *11*(3), 14–27.

Jackwerth, J. (2004). *Option-implied risk-neutral distributions and risk aversion* (Vol. 1., p. 85). Research Foundation Publications.

Jiang, G. J., & Tian, Y. S. (2005). Gauging the Investor Fear Gauge: Implementation Problems of the CBOE's New Volatility Index. *SSRN Electronic Journal*.

Jorion, P. (1995). Predicting Volatility in the Foreign Exchange Market. *The Journal of Finance*, *50*, 507–528.

Lamoureux, C., & Lastrapes, W. (1993). Forecasting stock-return variance: toward an understanding of stochastic implied volatilities. *Review of Financial Studies*, *6*(2), 293–326.

Latané, H. A., & Rendleman, R. J. (1976). Standard Deviations of Stock Price Ratios Implied in Option Prices. *The Journal of Finance*, *31*(2), 369–381.

Leippold, M., Wu, L., & Egloff, D. (2007). Variance Risk Dynamics, Variance Risk Premia, and Optimal Variance Swap Investments. *SSRN Electronic Journal*.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*, *36*(4), 394.

Martens, M., & Zein, J. (2004). Predicting financial volatility: High-frequency time-series forecasts vis-a-vis implied volatility. *Journal of Futures Markets*, *24*(11), 1005–1028.

McAleer, M., & Medeiros, M. C. (2008). Realized Volatility: A Review. *Econometric Reviews*, *27*(1-3), 10–45.

Mercurio, F. (2009). Interest rates and the credit crunch: new formulas and market models. *Bloomberg Portfolio Research Paper*, (November 2008), 1–37.

Merton, R. C. (1973). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, *4*(1), 141–183.

Neumann, M., & Skiadopoulos, G. (2013). Predictable Dynamics in Higher-Order Risk-Neutral Moments: Evidence from the S&P 500 Options. *Journal of Financial and Quantitative Analysis*, *48*(03), 947–977.

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*(3), 703–08.

Oomen, R. (2001). Using High Frequency Data to Calculate, Model and Forecast Realized Volatility. *Computing in Economics and Finance 2001*, *75*.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. In *Journal of Econometrics* (Vol. 160, pp. 246–256).

Poon, S. H., & Granger, C. (2003). Forecasting financial market volatility: A Review. *Journal of Economic Literature*, *41*, 478–539. doi:10.1257/002205103765762743

Poteshman, A. M. (2000). Forecasting Future Volatility from Option Prices. *SSRN Electronic Journal*.

Prokopczuk, M., & Simen, C. W. (2014). The Importance of the Volatility Risk Premium for Volatility Forecasting. *Journal of Banking & Finance*, *40*, 303–320.

Rehman, Z., & Vilkov, G. (2012). Risk-Neutral Skewness: Return Predictability and Its Sources. *SSRN Electronic Journal*.

Ron, U. (2000). *A practical guide to swap curve construction*. *Bank of Canada Working Paper* (Vol. 17, pp. 1 – 32).

Rosenberg, J. V., & Engle, R. F. (2002). Empirical pricing kernels. *Journal of Financial Economics*, *64*(3), 341–372.

Rubinstein, M. (1994). Implied Binomial Trees. *Journal of Finance*, *49*(3), 771–818.

Satchell, S., & Knight, J. (2007). *Forecasting Volatility in the Financial Markets* (p. 432).

Shephard, N., & Sheppard, K. (2010). Realising the future: forecasting with high-frequency based volatility (HEAVY) models. *Journal of Applied Econometrics*, 1–41.

Shimko, D. (1993). Bounds of Probability. *Risk*, *6*(4), 33–37.

Szakmary, A., Ors, E., Kyoung Kim, J., & Davidson, W. N. (2003). The predictive power of implied volatility: Evidence from 35 futures markets. *Journal of Banking and Finance*, *27*(11), 2151–2175.

Taylor, S. J. (1986). *Modelling Financial Time Series*. *Wiley New York*. Wiley.

Taylor, S. J., Yadav, P. K., & Zhang, Y. (2010). The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks. *Journal of Banking and Finance*, *34*(4), 871–881.

Todorov, V. (2010). Variance Risk-Premium Dynamics: The Role of Jumps. *Review of Financial Studies*, *23*(1), 345–383. doi:10.1093/rfs/hhp035

Vilkov, G. (2008). Variance Risk Premium Demystified. *SSRN Electronic Journal.*

Zhou, H. (2010). *Variance Risk Premia, Asset Predictability Puzzles, and Macroeconomic Uncertainty*. *Federal Reserve Board* (p. 50).

Appendix

Acronyms	Definition
IV	Implied Volatility
VRP	Variance Risk Premium
IVP	Instantaneous Variance Premium
RND	Risk Neutral Density
RV	Realized Variance (Volatility)
RNM	Risk Neutral Moments
ATM	At-The-Money
OTM	Out of-The-Money
ITM	In-The-Money

A. Calculation of Risk Neutral Moments

The risk neutral skewness and risk-neutral kurtosis are based on Bakshi, Kapadia and Madan (2003). For further details, please refer to the original paper.

Given an asset price process $\{S_t\}$, Bakshi, Kapadia, and Madan (2003) show that the risk neutral conditional moments of τ–period log return $R(t,\tau) \equiv \ln[S(t+\tau)] - ln[S(t)]$, given information at time t, can be written as an integral of OTM call and put option prices.

First, define the volatility contract, cubic contract and quartic contract payoffs:

$$V(t,\tau) = E_t^Q(e^{-r\tau}R(t,\tau)^2)$$

$$W(t,\tau) = E_t^Q(e^{-r\tau}R(t,\tau)^3)$$

$$X(t,\tau) = E_t^Q(e^{-r\tau}R(t,\tau)^4)$$

where Q refers to the risk neutral measure and r the risk-free rate. Under all martingale pricing measures, the contract prices can then be recovered from the market prices of OTM European calls and puts:

$$V(t,\tau) = \int_{S(t)}^{\infty} \frac{2\left(1 - log\left[\frac{K}{S(t)}\right]\right)}{K^2} C(t,\tau;K)dK + \int_{\infty}^{S(t)} \frac{2\left(1 + log\left[\frac{K}{S(t)}\right]\right)}{K^2} P(t,\tau;K)dK$$

$$W(t,\tau) = \int_{S(t)}^{\infty} \frac{6log\left[\frac{K}{S(t)}\right] - 3\left(log\left[\frac{K}{S(t)}\right]\right)^2}{K^2} C(t,\tau;K)dK$$

$$- \int_{\infty}^{S(t)} \frac{6log\left[\frac{K}{S(t)}\right] + 3\left(log\left[\frac{K}{S(t)}\right]\right)^2}{K^2} P(t,\tau;K)dK$$

$$X(t,\tau) = \int_{S(t)}^{\infty} \frac{12log\left[\frac{K}{S(t)}\right]^2 - 4\left(log\left[\frac{K}{S(t)}\right]\right)^3}{K^2} C(t,\tau;K)dK$$

$$+ \int_{\infty}^{S(t)} \frac{12log\left[\frac{K}{S(t)}\right]^2 + 4\left(log\left[\frac{K}{S(t)}\right]\right)^3}{K^2} P(t,\tau;K)dK$$

where C(t, τ;K) and P(t, τ;K) are the prices of call and put options respectively, with strike K and expiration τ periods from t, and risk-free rate r.

The time t conditional risk neutral moments are then defined as:

The risk neutral mean as:

$$Mean(t,\tau) = \mu(t,\tau) + ln(S_t)$$

The risk neutral variance as:

$$Var(t,\tau) = e^{rt}V(t,\tau) - \mu^2(t,\tau)$$

The risk-neutral skewness as

$$Skew(t,\tau) = \frac{e^{rt}W(t,\tau) - 3\mu(t,\tau)e^{rt}V(t,\tau) + 2(\mu(t,\tau))^3}{(e^{r\tau}V(t,\tau) - (\mu(t,\tau))^2)^{\frac{3}{2}}}$$

The risk-neutral excess kurtosis as

$$Kurt\,(t,\tau) = \frac{e^{rt}X(t,\tau) - 4\mu(t,\tau)e^{rt}W(t,\tau) + 6e^{rt}(\mu(t,\tau))^2V(t,\tau) - 3(\mu(t,\tau))^4}{(e^{r\tau}V(t,\tau) - (\mu(t,\tau))^2)^2} - 3$$

where

$$\mu(t,\tau) = e^{rt}\left[1 - e^{-rt} - \frac{1}{2}V(t-\tau) - \frac{1}{6}W(t,\tau) - \frac{1}{24}X(t,\tau)\right]$$

The calculation of the contracts (integrals) requires a continuum of prices. In our application, we approximate a continuum by interpolating prices and thus obtaining a spectrum of discrete prices with very small intervals. The integrals are then evaluated by using a trapezoidal approach.

B. Deriving The Risk Neutral Density

The procedure for deriving the risk neutral density builds on the result of Breeden and Litzenberger (1978). They show that the risk neutral density can be derived from taking the second derivative of the option price with respect to the strike price. In the following, we keep the same notation as previously (S=spot price, K=strike price, T=time to maturity etc).

The value of a call option is the expected value of its payoff on the expiration date T, discounted back to the present. Under risk neutrality, the expectation is taken with respect to the risk neutral probabilities and discounting is at the risk-free interest rate.

$$C = \int_{K}^{\infty} e^{-rT}(S_T - K)f(S_T)dS_T$$

Where S_T is the stock price at expiration. Taking the partial derivative with respect to the strike price, one can obtain:

$$\frac{\partial C}{\partial K} = e^{-rT}\left((X - X)f(X) + \int_{K}^{\infty} -f(S_t)dS_t\right)$$

From this expression we note two distinctive effects on the option price from a change in the strike price. The second term in the parentheses reflects the effect that for given increase in the strike price K, the *payoff* is reduced by a similar amount. The first term shows the other effect; when increasing the strike price, the *range* for which the option is ITM is reduced. When the change becomes infinitely small, this effect diminishes relative to the other, and it goes to zero in the limit as the strike price goes to zero (Figlewski, 2010). This leaves us with:

$$\frac{\partial C}{\partial K} = e^{-rT}\int_{K}^{\infty} -f(S_t)dS_t = e^{-rT}(1 - F(X))$$

Solving the above expression for the risk neutral distribution F(X) gives:

$$F(X) = e^{-rT}\frac{\partial C}{\partial K} + 1$$

Taking the partial derivative with respect to K once again yields:

$$\mathrm{f(X)} = \mathrm{e}^{-\mathrm{rT}} \frac{\partial^2 \mathrm{C}}{\partial \mathrm{K}^2}$$

which is an expression for the risk neutral density. Hence, taking the second derivative of the option price with respect to the strike price yields the risk neutral density. As an approximation, the second derivative can be calculated using discrete strike prices with small increments.

We approximate first derivative evaluated at strike price n as:

$$\mathrm{F(X_n)} = \mathrm{e}^{-\mathrm{rT}} \frac{\mathrm{C_n - C_{n-1}}}{\mathrm{K_n - K_{n-1}}} + 1$$

And the second derivative is approximated as:

$$\mathrm{f(X_n)} = \frac{\partial^2 \mathrm{C}}{\partial \mathrm{K}^2} = \frac{\frac{\mathrm{C_{n+1} - C_n}}{\mathrm{K_{n+1} - K_n}} - \frac{\mathrm{C_n - C_{n-1}}}{\mathrm{K_n - K_{n-1}}}}{\mathrm{K_n - K_{n-1}}}$$

C. Variance Swap

A return variance swap has zero net market value at entry. At maturity, the payoff to the long side of the swap is equal to the difference between the realized variance over the life of the contract and a constant called the *variance swap rate*:

$$N(RV_{t,T} - SW_{t,T})$$

where $RV_{t,T}$ denotes the realized annualized return variance between time *t* and *T*, $SW_{t,T}$ denotes the fixed variance swap rate that is determined at time *t* and paid at time *T*, and *N* denotes the notional dollar amount that converts the variance difference into a dollar payoff. According to Carr and Wu (2009), no arbitrage dictates that the variance swap rate equals the risk-neutral expected value of the realized variance:

$$SW_{t,T} = E_t^Q(RV_{t,T})$$

where E_t^Q denotes the time-*t* conditional expectation under some risk-neutral measure Q. The risk neutral variance estimated from Bakshi, Kapadia and Madan (2003) represents a measure of a risk neutral expected value of the realized variance. Hence, it is possible to synthesize variance swap rates from a static portfolio of options. We define the monthly variance swap as:

$$variance\ swap = N\left(RV_{t,t+22}^2 - {IV_t}^2\right)$$

where

$${IV_t}^2 = Var(t,\tau) = e^{rt}V(t,\tau) - \mu^2(t,\tau)$$

and

$$V(t,\tau) = \int_{S(t)}^{\infty} \frac{2\left(1 - log\left[\frac{K}{S(t)}\right]\right)}{K^2} C(t,\tau;K)dK + \int_{\infty}^{S(t)} \frac{2\left(1 + log\left[\frac{K}{S(t)}\right]\right)}{K^2} P(t,\tau;K)dK$$

$$\mu(t,\tau) = e^{rt}\left[1 - e^{-rt}\frac{1}{2}V(t-\tau) - \frac{1}{6}W(t,\tau) - \frac{1}{24}X(t,\tau)\right]$$

$RV_{t,t+22}^2$ is the estimate of the realized variance, i.e. the realized kernel. The industry standard is using daily log returns. Since our strike is only a theoretically derived value, we stick to using the realized kernel for the realized leg, which is more theoretically precise estimate of the realized volatility.

The setup ensures that the portfolio is independent of movements in the underlying and thus only dependent on volatility. However, the method is not completely model independent as it is assumed that the spot follows a general continuous diffusion process with deterministic but undefined drift and diffusion term. This means, that if jumps occur in the underlying, the replication will not be perfect. Still, the measure used here (and tiny variations of it) is commonly used as a benchmark for variance swap rate.

D. Realized Bipower Variation and Realized Semivariance

Instead of basing our analysis on only one dependent variable, the realized variance, we also take the realized semivariance and the realized bipower variation into consideration. These two measures of risk have valuable features, which we will explain in the following. Furthermore, we provide in this section the formula, which we use to calculate these risk measures.

The realized semivariance is entirely based on downward movements. Similar to the realized variance it is measured by using high frequency data. This measure embeds some important predictive qualities for the future volatility of the underlying asset (Barndorff-Nielsen, Kinnebrock, & Shephard, 2010) When talking about risk in the financial market, it normally raises the question about what one can lose. This downside risk has been operationalized by the value at risk, expected shortfall and semivariance, the predecessor of the realized semivariance.

In our calculation we follow the approach of Barndorff-Nielsen et al. (2010). It is important to notice that a higher density in high frequency data shrinks the extra information in the sign of the data. Thus the most elegant way is to see Y as a Brownian semimartingale and allow jumps in the process

$$Y_t = \int_0^t a_s ds + \int_0^t \sigma_s dW_s + J_t$$

where a is a locally bounded predictable drift process, σ is a càdlàg volatility process – both adapted to a common filtration process F_t, which implies that the model can allow for classic leverage effects – and J is a pure jump process. Now, writing jumps in Y as $\Delta Y_t = Y_t - Y_{t-1}$ then we receive:

$$[Y]_t = \int_0^t \sigma_s ds + \sum_{s\leq 1} (\Delta Y_s)^2$$

Keeping this in mind, Barndorff-Nielsen et al. (2010) introduce the downside realized semivariance as

$$RS^- = \sum_{j=1}^{t_j \leq 1} (Y_{t_j} - Y_{t_j-1})^2 1_{Y_{t_j} - Y_{t_{j-1}}} \leq 0$$

where 1_Y is the indicator function taking the value 1 of the argument Y is true and hence they are able to study the behavior of the realized semivariance. In particular, they show that

$$RS^- \xrightarrow{p} \frac{1}{2} \int_0^t \sigma_s^2 ds + \sum_{s\leq 1} (\Delta Y_s)^2 1_{\Delta Y_s \leq 0}$$

under in-fill asymptotics.

The realized bipower variation is another alternative measure to the realized variance. Contrary to the RV and semivariance it is robust to jumps. Thus the difference between the RV and the bipower variation can be expressed as the quadratic variation of the jump component.

Here we follow the approach of Barndorff-Nielsen & Shephard (2004) who define

$$y^* = a^* + m^*$$

where a^* is a drift term and has locally finite variation paths FV_{loc} and m is a local martingale M_{loc}. The bipower variation is then

$$\{y_M^*\}_i^{[r,s]} = \left\{ \left(\frac{h}{M}\right)^{1-(r+s)/2} \right\} \sum_{j=1}^{M-1} |y_{j,i}|^r |y_{j+1,i}|^s, \quad r,s \geq 0.$$

For a detailed derivation of the semivariance and the bipower variation see the article of Barndorff-Nielsen et al. (2010) and Barndorff-Nielsen & Shephard (2004) respectively.

E. Details on various forecast variables

Variable	Definition
$\boldsymbol{RV_t^D}$	$RV_{t-1,t}$
$\boldsymbol{RV_t^W}$	$\frac{1}{5}(RV_t^D + RV_{t-1}^D + RV_{t-2}^D + RV_{t-3}^D + RV_{t-4}^D)$
$\boldsymbol{RV_t^M}$	$\frac{1}{22}(RV_t^D + RV_{t-1}^D + \ldots + RV_{t-21}^D)$
$\boldsymbol{IVP_t^D}$	$IV_{t-1} - RV_{t-1,t}$
$\boldsymbol{IVP_t^W}$	$\frac{1}{5}(IVP_t^D + IVP_{t-1}^D + IVP_{t-2}^D + IVP_{t-3}^D + IVP_{t-4}^D)$
$\boldsymbol{SK_t}$	$Skew(t, \tau)$
$\boldsymbol{KU_t}$	$Kurt(t, \tau)$

The current time, t=0, may be thought of as the time after closing on the given day. Hence, $RV_{t-1,t}$ is the RV measured earlier the same day. IV at t=0 is based on closing prices on the given day. Hence, it essentially succeeds the RV. Our volatility data (RV, bipower variation, semivariance) is originally estimated on a daily scale, while the risk neutral moments are calculated on a 30 days scale. Hence, a scaling from one day to 30 days is required. The scaling assumes that the price process follows a Geometric Brownian motion with lognormally distributed prices and normally distributed returns, which may not be entirely accurate, as detailed in Diebold & Hickman (1998). The variables are scaled with the square root of time. The IV is annualized (like the RV) while skewness and kurtosis are not scaled (monthly). However, once the volatility data is scaled to 30 days, further scaling only has a very limited impact on the regression results since all variables, except skewness and kurtosis, are scaled equally. IV is annualized in order make it easier to relate to the values. There is also a convention for scaling skewness and kurtosis that - like volatility - rests upon assumptions about the returns. Since we do not know anything about the relationship between volatility, and skewness and kurtosis (i.e. is it linear or something else?) there is little basis for having an opinion about whether a scaling will make the relationship between the variables more or less linear. The skewness and kurtosis are therefore left unchanged.

F. The Obtained 30 Days Volatility Smiles and Risk Neutral Densities

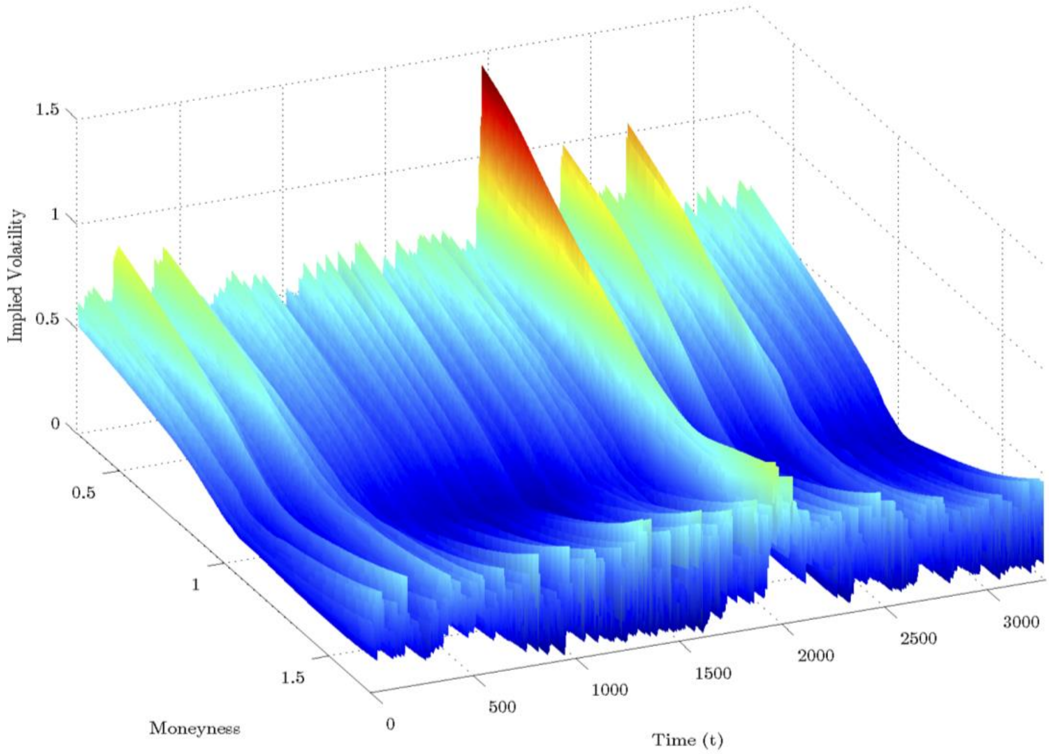


Figure 15. Dynamics of the volatility smiles

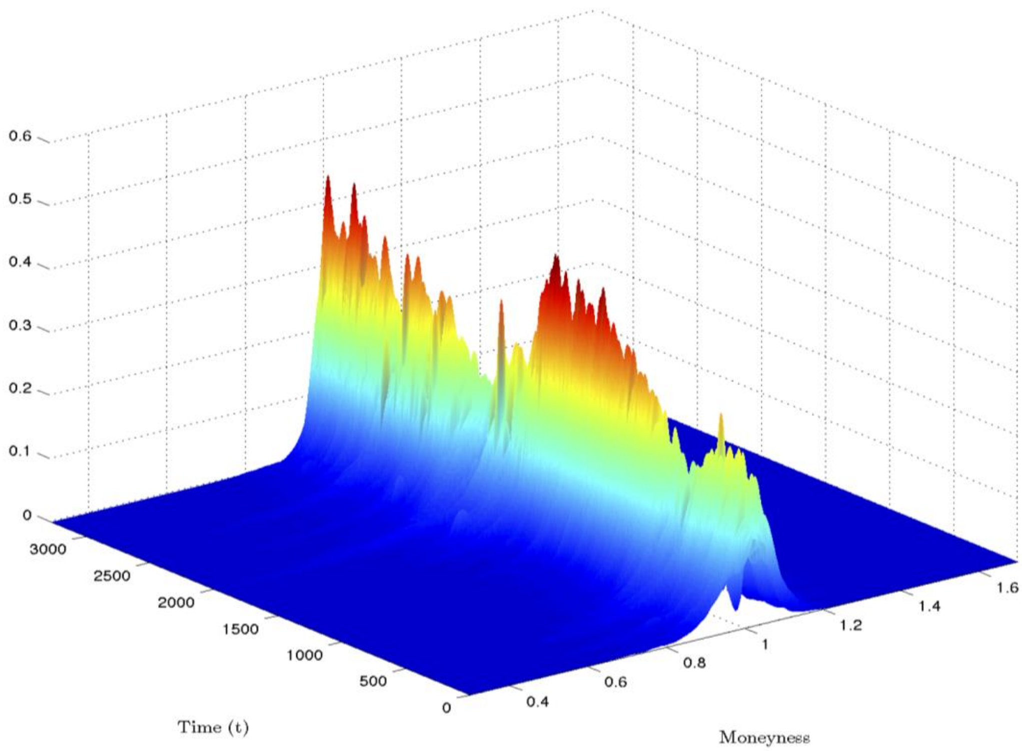


Figure 16. Risk neutral densities

G. Risk Neutral Skewness and Excess Kurtosis

For a better visibility and readability we disclose the plots and descriptive statistics of the risk neutral skewness and kurtosis in this section:

Table 10. Descriptive statistics of the risk neutral skewnes

	Full Sample	Pre-Crisis	Banking-Crisis	Post-Crisis
Observations	3,245	1636	503	1106
Max	-0.584	-0.584	-0.808	-0.933
Min	-3.480	-3.116	-2.322	-3.480
Mean	-1.586	-1.504	-1.357	-1.813
Median	-1.564	-1.442	-1.307	-1.761
Range 5%-95%	-2.276--0.966	-2.195--0.914	-1.893--0.945	-2.453--1.323

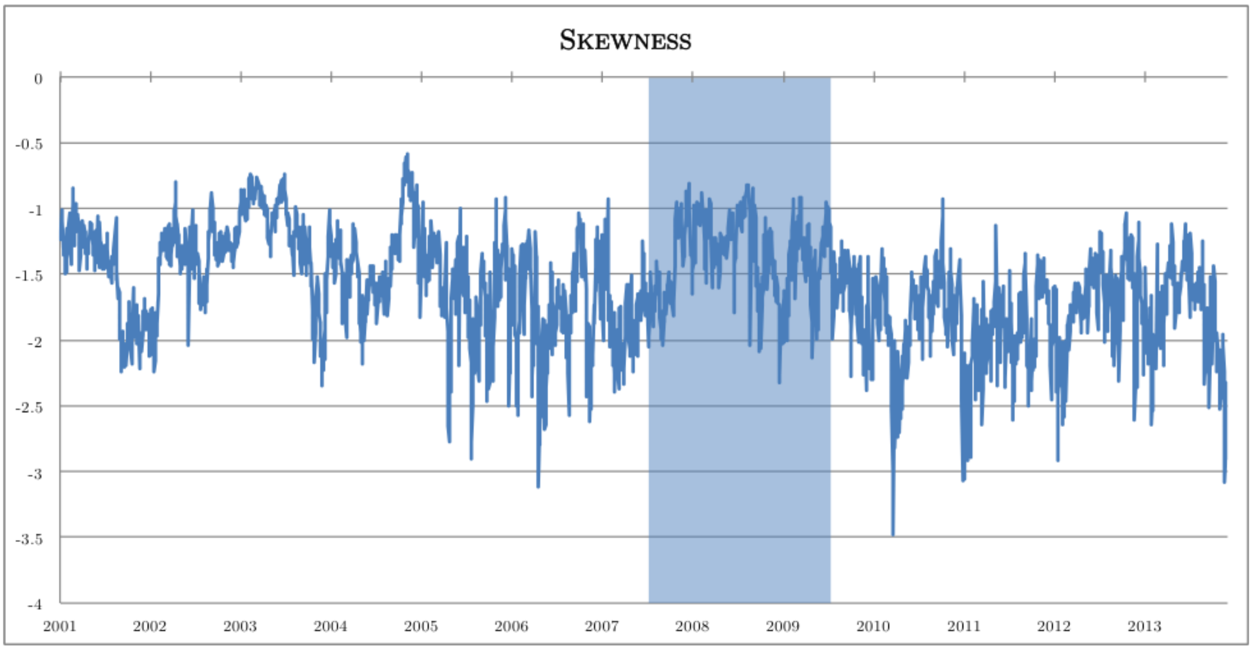


Figure 17. Risk neutral skewness

Table 11. Descriptive statistics of the risk neutral (excess) kurtosis

	Full Sample	Pre-Crisis	Banking-Crisis	Post-Crisis
Observations	3,245	1636	503	1106
Max	25.959	18.524	9.780	25.959
Min	0.814	1.134	0.814	1.876
Mean	5.609	5.392	3.601	6.845
Median	5.034	4.739	3.303	6.028
Range 5%-95%	1.91-10.938	1.94-10.478	1.416-6.677	3.543-12.263

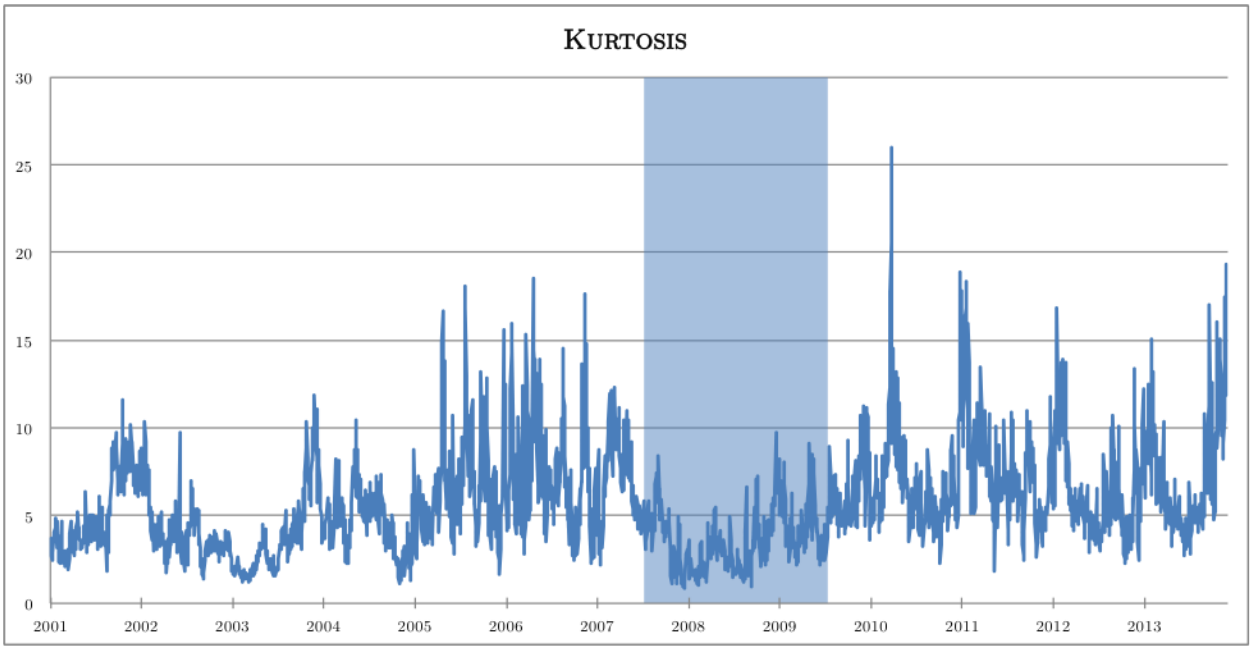


Figure 18. Risk neutral kurtosis

Table 12. Adding risk-neutral moments to the HAR-RV-IV (*Extended model)*

Moments	IV			IV-SK			IV-SK-KU		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.7578	0.8003	0.6883	0.7598	0.8029	0.6896	0.7600	0.8037	0.6910
$ß_0$	-0.0204***	-0.0089	0.0153**	0.0021	0.0143	0.0309*	0.0104	0.0276**	0.0469*
	(-4.16)	(-1.57)	(2.25)	(0.35)	(1.58)	(1.84)	(1.3)	(2.41)	(1.96)
$ß_{RV}{}^{D}$	0.2805***	0.1886***	0.1186***	0.272***	0.1798***	0.1127***	0.2711***	0.1784***	0.1111***
	(4.42)	(5.14)	(3.42)	(4.28)	(5.02)	(3.71)	(4.24)	(4.99)	(3.78)
$ß_{RV}{}^{W}$	0.3247***	0.3396***	0.2913**	0.3216***	0.3364***	0.2891**	0.3232***	0.3391***	0.2924**
	(3.53)	(4.53)	(2.34)	(3.52)	(4.47)	(2.33)	(3.54)	(4.5)	(2.34)
$ß_{RV}{}^{M}$	-0.164***	-0.0388	0.1044	-0.1893***	-0.0649	0.0868	-0.1964***	-0.0764	0.0731
	(-2.59)	(-0.48)	(1.21)	(-2.88)	(-0.77)	(0.91)	(-2.9)	(-0.88)	(0.72)
$ß_{IV}$	0.5664***	0.471***	0.3244***	0.5954***	0.5008***	0.3445***	0.6104***	0.5251***	0.3736***
	(10.24)	(9.01)	(2.99)	(10.53)	(9.49)	(3.38)	(10.04)	(9.57)	(3.89)
$ß_{SK}$	0	0	0	0.0139***	0.0143***	0.0096	0.0281**	0.0374***	0.0373
	(0)	(0)	(0)	(4.67)	(3.22)	(1.01)	(2.52)	(2.68)	(1.33)
$ß_{KU}$	0	0	0	0	0	0	0.0022	0.0035**	0.0043
	(0)	(0)	(0)	(0)	(0)	(0)	(1.51)	(2)	(1.18)

Table 13. Adding risk-neutral moments to the HAR-IVP-IV (*Bermuda model)*

Moments	IV			IV-SK			IV-SK-KU		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.8415	0.8358	0.7067	0.8417	0.8372	0.7079	0.8418	0.8381	0.7099
$ß_0$	0.0016	0.0079*	0.0286***	0.0094**	0.0247***	0.0436**	0.016***	0.039***	0.063**
	(0.62)	(1.79)	(2.79)	(2.05)	(2.98)	(2.24)	(2.59)	(3.63)	(2.47)
$ß_{IVP}{}^{D}$	0.1417**	0.0443**	0.0437***	0.142**	0.0449**	0.0442***	0.1417**	0.0443**	0.0433***
	(2.45)	(1.97)	(3.09)	(2.46)	(2.01)	(3.07)	(2.45)	(1.98)	(3.01)
$ß_{IVP}{}^{W}$	-1.2162***	-0.6455***	-0.3537**	-1.2109***	-0.6341***	-0.3435**	-1.2102***	-0.6327***	-0.3416**
	(-10.08)	(-8.98)	(-2.4)	(-9.99)	(-9.1)	(-2.4)	(-10.03)	(-9.12)	(-2.41)
$ß_{IVP}{}^{M}$	0.1209*	-0.2073	-0.387**	0.1229*	-0.2031	-0.3833**	0.1221*	-0.2048	-0.3855**
	(1.7)	(-1.58)	(-2.09)	(1.74)	(-1.57)	(-2.03)	(1.73)	(-1.59)	(-2.07)
$ß_{IV}$	0.985***	0.9318***	0.8049***	0.9828***	0.927***	0.8006***	0.99***	0.9424***	0.8217***
	(58.34)	(35.31)	(20.64)	(57.15)	(35.07)	(21.1)	(53.33)	(32.28)	(18.45)
$ß_{SK}$	0	0	0	0.0048**	0.0103***	0.0092	0.0162**	0.0347***	0.0424*
	(0)	(0)	(0)	(2.1)	(2.69)	(1.11)	(2.31)	(3.06)	(1.72)
$ß_{KU}$	0	0	0	0	0	0	0.0018*	0.0038**	0.0052
	(0)	(0)	(0)	(0)	(0)	(0)	(1.88)	(2.54)	(1.54)

H. LIBOR-OIS spread

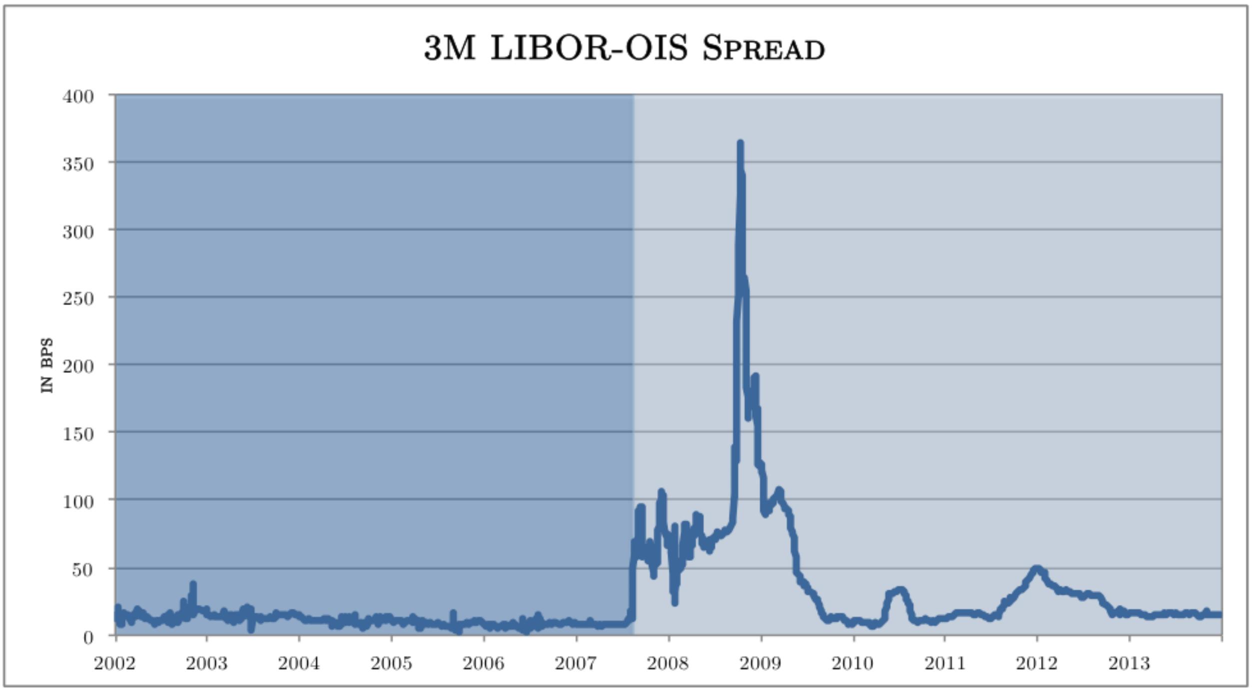


Figure 19. 3M LIBOR-OIS Spread

I. Forecasting Result: Bipower variation and Semivariance

Table 14. In-sample forecast bipower variation (HAR-RV: *Conventional model)*

HAR-RV	Bipower variation			Counterpart (BV)		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.7518	0.7789	0.6722	0.0625	0.2120	0.3697
$ß_0$	0.0028	0.009	0.025***	0.0058***	0.0062***	0.0069***
	(0.67)	(1.59)	(3.3)	(3.68)	(3.74)	(3.6)
$ß_{RV}{}^{D}$	0.3662***	0.2706***	0.1719***	0.0258*	0.0107	0.0106***
	(5.64)	(7.4)	(7.83)	(1.64)	(1.05)	(3.29)
$ß_{RV}{}^{W}$	0.3883***	0.3736***	0.3213***	0.0154	0.0317	0.0153
	(4.61)	(4.99)	(3.07)	(0.68)	(1.55)	(1.59)
$ß_{RV}{}^{M}$	0.1317**	0.2063***	0.2561***	0.0231	0.02	0.0309**
	(2.56)	(2.84)	(2.97)	(1.08)	(0.88)	(2.04)

Table 15. In-sample forecast bipower variation (HAR-RV-RNM: *Extended* model)*

HAR-RV-RNM	Bipower variation			Counterpart (BV)		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.7752	0.7994	0.6843	0.0657	0.2212	0.3803
$ß_0$	0.0091	0.0232**	0.0416*	0.0012	0.0044	0.0053
	(1.14)	(2.06)	(1.89)	(0.36)	(1.23)	(1.08)
$ß_{RV}{}^{D}$	0.2539***	0.1753***	0.1061***	0.0172	0.0031	0.005
	(3.9)	(5.02)	(3.96)	(0.99)	(0.3)	(1.2)
$ß_{RV}{}^{W}$	0.3148***	0.3132***	0.2813**	0.0084	0.0259	0.011
	(3.55)	(3.71)	(2.28)	(0.38)	(1.32)	(1.13)
$ß_{RV}{}^{M}$	-0.1974***	-0.0761	0.057	0.001	-0.0003	0.0161
	(-2.98)	(-0.8)	(0.55)	(0.04)	(-0.01)	(0.82)
$ß_{IV}$	0.5736***	0.4919***	0.3491***	0.0367	0.0332*	0.0245*
	(10.18)	(9.18)	(3.57)	(1.59)	(1.94)	(1.68)
$ß_{SK}$	0.0344***	0.0404***	0.0396	-0.0062	-0.0031	-0.0024
	(3.24)	(3.04)	(1.52)	(-1.56)	(-0.85)	(-0.5)
$ß_{KU}$	0.0033**	0.0043**	0.0048	-0.0011**	-0.0007	-0.0005
	(2.41)	(2.51)	(1.41)	(-2.14)	(-1.61)	(-0.93)

Table 16. In-sample forecast bipower variation (HAR-IVP-RNM: *Bermuda* model)*

HAR-IVP-RNM	Bipower variation			Counterpart (BV)		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.8449	0.8338	0.7047	0.1023	0.2317	0.3809
$ß_0$	0.015**	0.0342***	0.0569**	0.0011	0.0047	0.006
	(2.32)	(3.32)	(2.42)	(0.32)	(1.53)	(1.36)
$ß_{IVP}{}^{D}$	0.081	0.0281	0.0369**	0.0607***	0.0162**	0.0064***
	(1.31)	(1.31)	(2.52)	(2.96)	(2.29)	(3.02)
$ß_{IVP}{}^{W}$	-1.0114***	-0.5707***	-0.3145**	-0.1988***	-0.062***	-0.0271*
	(-9.78)	(-8.31)	(-2.34)	(-4.89)	(-2.58)	(-1.84)
$ß_{IVP}{}^{M}$	0.0527	-0.2124*	-0.3801*	0.0694**	0.0077	-0.0054
	(0.78)	(-1.71)	(-1.87)	(2.28)	(0.23)	(-0.18)
$ß_{IV}$	0.9269***	0.8812***	0.7661***	0.0631***	0.0612***	0.0556***
	(45.83)	(29.54)	(18.54)	(5.86)	(6.99)	(5.66)
$ß_{SK}$	0.0235***	0.0378***	0.044*	-0.0073*	-0.003	-0.0016
	(3.22)	(3.41)	(1.92)	(-1.87)	(-0.91)	(-0.36)
$ß_{KU}$	0.0029***	0.0045***	0.0056*	-0.0011**	-0.0007	-0.0004
	(3.01)	(3.05)	(1.76)	(-2.17)	(-1.57)	(-0.78)

Table 17. In-sample forecast semivariance (HAR-RV: *Conventional model)*

HAR-RV	Semivariance			Counterpart (SV)		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.5974	0.7322	0.6640	0.4147	0.7079	0.6799
$ß_0$	0.0048	0.009**	0.0204***	0.0038**	0.0062***	0.0115***
	(1.32)	(2.05)	(3.52)	(2.26)	(3.4)	(6.07)
$ß_{RV}{}^{D}$	0.242***	0.1735***	0.1218***	0.1499***	0.1078***	0.0607***
	(3.7)	(5.75)	(7.04)	(6.09)	(8.57)	(8.29)
$ß_{RV}{}^{W}$	0.2709***	0.2851***	0.2413***	0.1328***	0.1202***	0.0953***
	(3.6)	(4.98)	(3.09)	(3.74)	(5.63)	(3.17)
$ß_{RV}{}^{M}$	0.162***	0.1926***	0.2147***	-0.0071	0.0337*	0.0723***
	(3.47)	(3.47)	(3.38)	(-0.3)	(1.7)	(2.6)

Table 18. In-sample forecast semivariance (HAR-RV-RNM: *Extended* model)*

HAR-RV-RNM	Semivariance			Counterpart (SV)		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.6102	0.7478	0.6744	0.4393	0.7364	0.6964
$ß_0$	-0.0006	0.0131	0.0257	0.011***	0.0145***	0.0211***
	(-0.08)	(1.36)	(1.43)	(2.69)	(4.03)	(3.24)
$ß_{RV}{}^{D}$	0.1708**	0.1071***	0.0724***	0.1003***	0.0714***	0.0387***
	(2.53)	(3.73)	(3.41)	(3.94)	(5.75)	(4.44)
$ß_{RV}{}^{W}$	0.2221***	0.2417***	0.2099**	0.1011***	0.0974***	0.0825**
	(2.87)	(3.69)	(2.24)	(2.89)	(4.97)	(2.45)
$ß_{RV}{}^{M}$	-0.0427	-0.0022	0.0675	-0.1536***	-0.0741***	0.0056
	(-0.74)	(-0.03)	(0.95)	(-4.87)	(-3.39)	(0.17)
$ß_{IV}$	0.3562***	0.3393***	0.2592***	0.2542***	0.1859***	0.1144***
	(6.58)	(6.78)	(3.44)	(6.68)	(8.14)	(4.77)
$ß_{SK}$	0.0086	0.0205*	0.0216	0.0195***	0.0169***	0.0157**
	(0.92)	(1.82)	(1.02)	(3.37)	(3.52)	(2.09)
$ß_{KU}$	0.0003	0.002	0.0026	0.0019**	0.0016***	0.0016*
	(0.24)	(1.34)	(0.96)	(2.5)	(2.62)	(1.69)

Table 19. In-sample forecast semivariance (HAR-IVP-RNM: *Bermuda* model)*

Model 3	Semivariance			Counterpart (SV)		
Forecast horizon	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Adj. R^2	0.6856	0.7770	0.6909	0.4672	0.7710	0.7203
$ß_0$	0.0051	0.0223**	0.0374*	0.0109***	0.0167***	0.0255***
	(0.77)	(2.46)	(1.95)	(2.91)	(4.88)	(3.74)
$ß_{IVP}{}^{D}$	0.1616***	0.0541***	0.037***	-0.0199	-0.0098	0.0063*
	(2.94)	(2.84)	(3.35)	(-0.76)	(-1.32)	(1.72)
$ß_{IVP}{}^{W}$	-0.9181***	-0.4365***	-0.2394**	-0.2921***	-0.1963***	-0.1023**
	(-9.71)	(-7.63)	(-2.37)	(-4.88)	(-7.06)	(-2.36)
$ß_{IVP}{}^{M}$	0.0956	-0.1639	-0.2769*	0.0264	-0.0408	-0.1086**
	(1.25)	(-1.59)	(-1.94)	(0.68)	(-1)	(-2.31)
$ß_{IV}$	0.6922***	0.6678***	0.5886***	0.2978***	0.2746***	0.2331***
	(38.57)	(28.57)	(17.22)	(26.37)	(25.44)	(20.74)
$ß_{SK}$	0.0042	0.0211**	0.026	0.012**	0.0136***	0.0164**
	(0.58)	(2.19)	(1.39)	(2.57)	(3.47)	(2.52)
$ß_{KU}$	0.0004	0.0024*	0.0033	0.0014**	0.0014***	0.0018**
	(0.43)	(1.83)	(1.31)	(2.13)	(2.79)	(2.05)

J. Trading Strategy

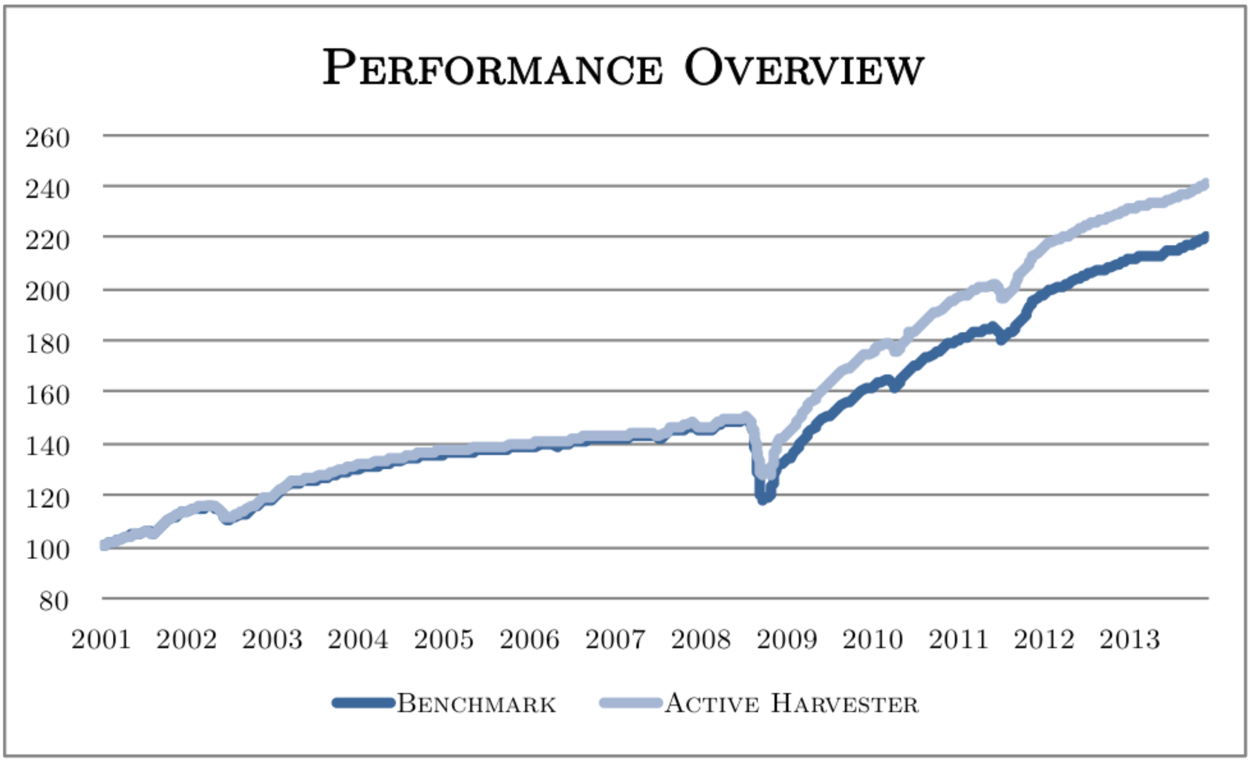


Figure 20. Performance Benchmark vs. Active Harvester

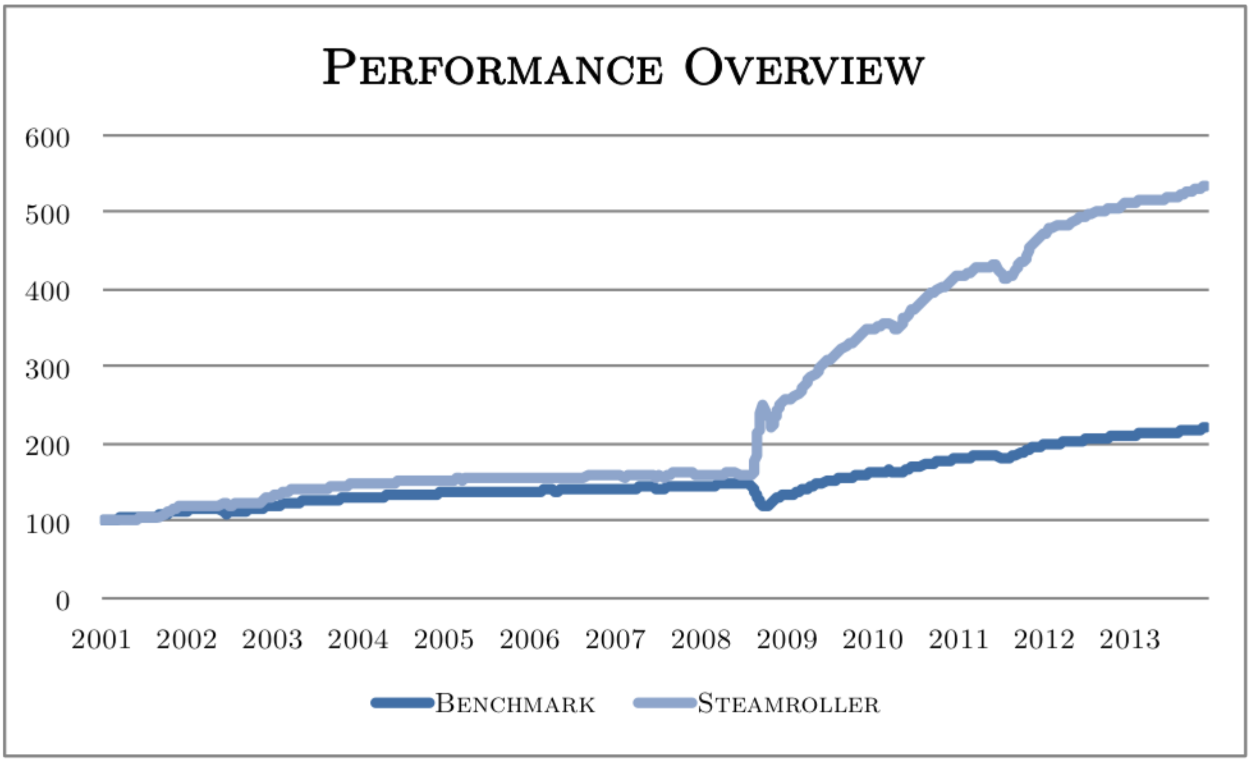


Figure 21. Performance Benchmark vs. Steamroller

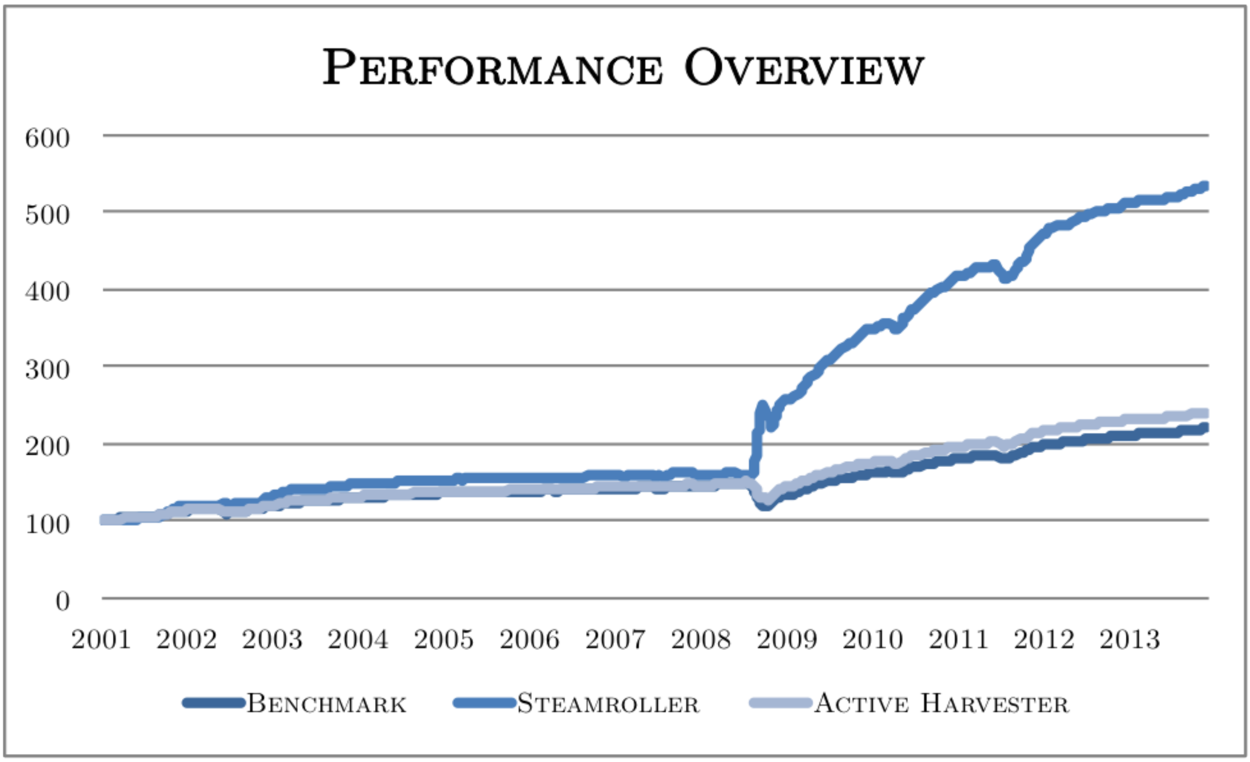


Figure 22. Performance Benchermark vs. Steamroller vs. Active Harvester

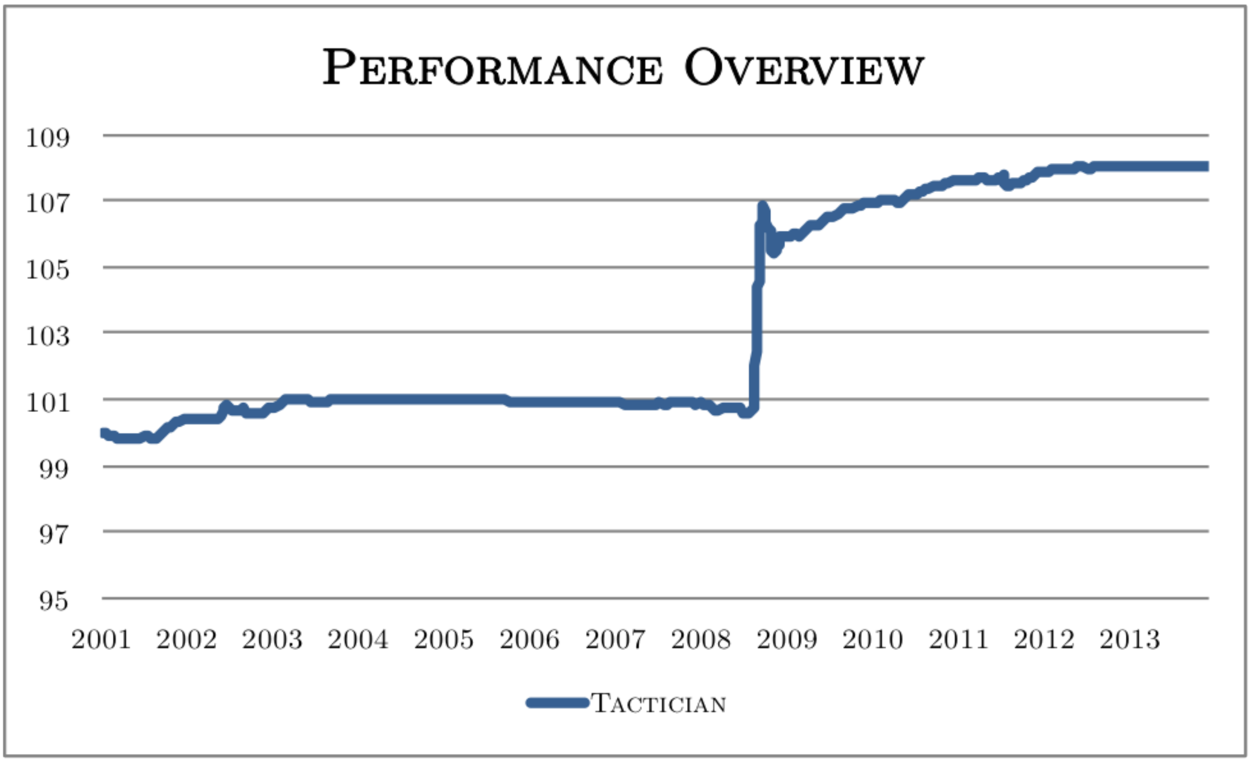


Figure 23. Performance Tactician

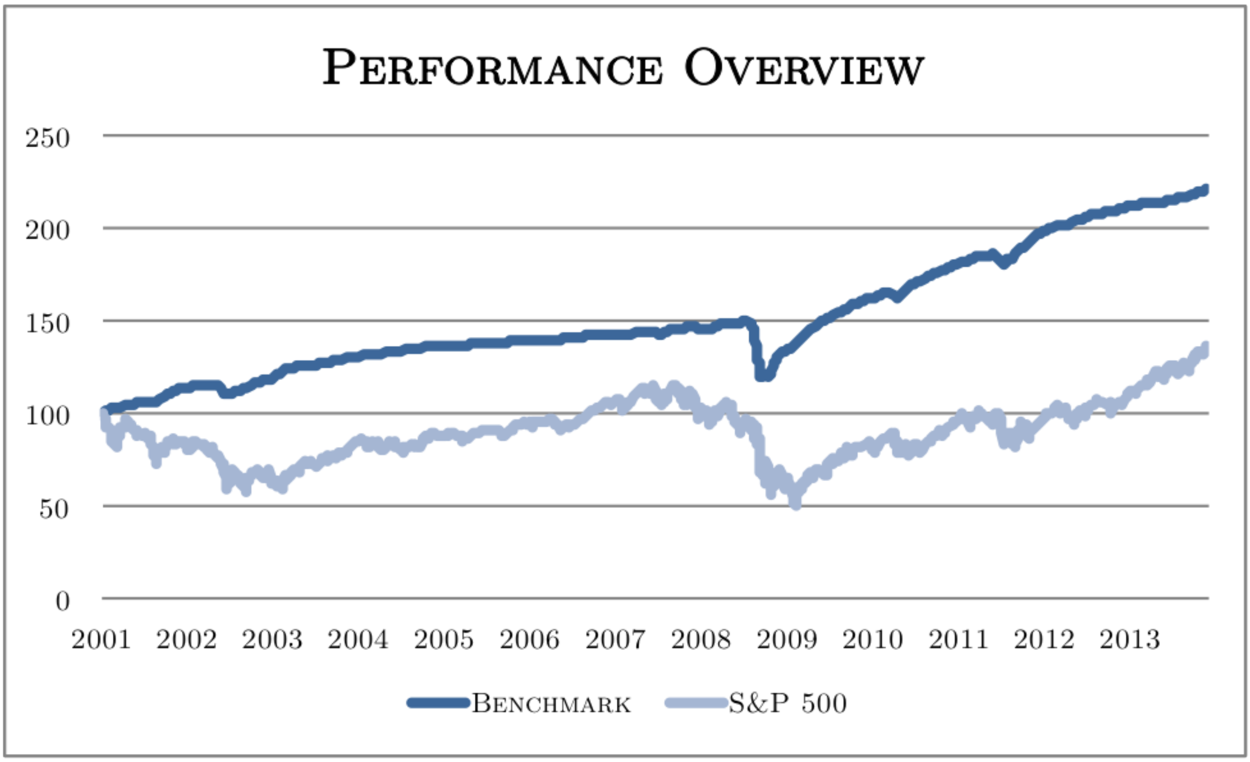


Figure 24. Performance Benchmark vs. S&P 500