GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Moment Risk Premiums in Option Markets:

On Measurement, Structure, and Investment Implications

Dissertation

zur Erlangung des Doktorgrades
der Wirtschaftswissenschaftlichen Fakultät
der Georg-August-Universität Göttingen

vorgelegt von

Julian Dörries, M. Sc.
aus Einbeck

Göttingen, 2021

Thesis Committee:

Prof. Dr. Olaf Korn
Prof. Dr. Jan Muntermann
Prof. Dr. Tino Berger

Contents

List of Figures

List of Tables

1 Introduction

Rational investors with reasonable utility functions are in general not risk-neutral. Indeed, there is a broad strand of literature in psychology, decision theory, and economics that describes and analyzes risk aversion in human behavior.[1] An important implication of risk aversion is that investors demand compensation for certain risks they take—risk premiums. To analyze such risk premiums, they have to be estimated from market data because a risk premium is an ex-ante concept and not directly observable. A market that provides valuable information for the estimation of risk premiums is the options market. Options contain a rich set of information about the market's risk-neutral expectation of the underlying's conditional return density, which is due to their forward-looking nature and their non-linear payoff structure (e.g., Breeden and Litzenberger 1978; Figlewski 2018). The fact that rational investors are in general not risk-neutral means that the option-implied return density as well as its characteristics do not necessarily coincide with their physically realized counterparts. That is to say, any adjustments due to investors' risk preferences translate into physical return distributions that deviate from their risk-neutral counterparts (Harrison and Kreps 1979; Bakshi and Madan 2006).

This deviation allows for the estimation and analysis of risk premiums that are observable only in option markets. By determining the difference between a physical and a risk-neutral (option-implied) quantity, one can draw conclusions on different types of risk premiums. A prominent example are moment risk premiums. They are defined as the difference between a particular statistical moment of the risk-neutral return distribution and the corresponding moment of the physical return distribution. Carr and Wu (2009), for example, provide an economic rationale for the existence of a variance risk premium, that is, the difference between physical and risk-neutral variance. They argue that investors do not only face the risk of uncertain returns but at least the risk of uncertain return variance, for which they might require a compensation. This argument can likewise be extended for skewness, kurtosis, and potentially other higher moments of the

[1] A starting point for this wide topic are, for example, seminal papers by Pratt (1964), Arrow (1971), and Kahneman and Tversky (1979).

return distribution, as the return distribution itself is uncertain.[2] Put differently, one can argue about an investor's hedging demand and say that an investor might be willing to pay a premium to hedge the corresponding moment risk.

There is a broad strand of literature on moment risk premiums. In the recent literature, they are typically determined as payoffs from synthetically derived swap contracts. These contracts exchange a fixed swap rate (i.e., the risk-neutral expectation of a specific moment for a predetermined period of time) for a floating rate (i.e., the realized moment over a predetermined period of time) and are constructed such that their initial value is zero. Realized payoffs of these contracts can be interpreted as realized premiums, while time series averages of realized payoffs serve as proxies for the respective average moment risk premium. Carr and Wu (2009) were the first to propose this methodology for analyzing the variance risk premium in the U.S. equity index market. To date, research has tackled questions of the existence, magnitude, term structure, and other characteristics of moment risk premiums in a variety of settings. There is extensive literature on variance and skewness risk premiums in currency markets (e.g., Broll 2016; Da Fonseca and Dawui 2021), in commodity markets (e.g., Da Fonseca and Xu 2017; Ruan and Zhang 2019), in credit markets (e.g., Ammann and Moerke 2021), in volatility markets (e.g., Da Fonseca and Xu 2019), and in equity markets (e.g., Kozhan, Neuberger, and Schneider 2013; Elyasiani, Gambarelli, and Muzzioli 2020).

For this dissertation, the literature on moment risk premiums for the U.S. equity index option market, particularly for the Standard & Poor's 500 (S&P 500) index, is most relevant. There are a number of stylized facts about these premiums that form the foundation of the research questions this dissertation addresses[3]: (i) There is ample evidence for the existence of variance, skewness and kurtosis risk premiums. Thus, investors are willing to pay a premium to hedge the risk associated with that respective uncertainty. (ii) Moment risk premiums are time varying and highly correlated with each other. There is evidence that they can be traced back to one common source. (iii) The variance risk premium can be decomposed into a downside and an upside related component, each of which has different characteristics. (iv) Payoffs of moment swaps are severely affected by extreme events, such as the 2008 financial crisis and the 2020 Covid-19 stock market crash. These events are reflected as very pronounced peaks in the time

[2] In the course of this thesis, (higher-) moment risk premiums refer to risk premiums associated with variance, skewness and kurtosis.

[3] References for these stylized facts are, among others, Bollerslev, Tauchen, and Zhou (2009); Carr and Wu (2009); Kozhan, Neuberger, and Schneider (2013); Kilic and Shaliastovich (2019); Khrashchevskyi (2020); Fan, Xiao, and Zhou (2020).

series of realized premiums. It stands to reason that the effect of extreme events is stronger the higher the moment is. (v) Moment risk premiums can be utilized as predictors for subsequent market excess returns.

At a first glance, the aforementioned stylized facts present a detailed picture of moment risk premiums. However, the premiums have a very special structure in that they fluctuate at low levels, only show significant peaks during extreme economic events, and in being highly correlated with each other. It is this structure that makes the premiums particularly interesting in research and practice and that raises a variety of issues that have not yet been answered.

First, this structure calls into question the robustness of premiums. Since moments, especially higher ones, are quite sensitive to outliers, one can ask: How can premiums be measured in a robust way? Do more robust moment risk premium definitions lead to the same results about moment risk premiums? How do robustly formulated premiums compare to traditionally defined premiums? Second, the evidence of decomposed variance risk premiums raises the need of a more detailed description of the fine structure of moments higher than variance. Natural questions in this respect are: How can higher-moment risk premiums be decomposed? What are characteristics of decomposed higher-moment risk premiums? Moreover, given stylized fact number (v), can a better knowledge of the premiums' structure and their decomposition improve market return predictions? What drives moment risk premiums' market return predictability? Third, the question arises to what extent long-term investors can benefit at all from the existence of moment risk premiums. With respect to the variance risk premium, for example, it is conceivable that rare extreme events will wipe out any gains made during normal market phases in a long-term strategy. Is it possible for a long-term investor to accumulate capital with a strategy that earns the variance risk premium? What are risk and return characteristics of such strategies? Is the variance factor an attractive factor compared to other investment factors? On a high-level, all these questions can be condensed to three main research questions:

(i) How can moment risk premiums be measured and quantified in a robust and flexible manner and what are these robust premiums' characteristics?

(ii) What are the characteristics of decomposed higher-moment risk premiums and how can a better knowledge of their structure be utilized to predict market returns?

(iii) Is it possible for long-term investors to harvest the variance risk premium, and if so, what are feasible trading strategies to achieve this goal?

This dissertation aims to provide answers to these questions. In doing so, it significantly contributes to the literature on moment risk premiums in the U.S. equity index option market and enhances the understanding of the premiums, which is relevant for both researchers and practitioners in (empirical) asset pricing as well as asset and risk management. In three main chapters—each of them presents an individual study that can be read independently—it one by one addresses the aforementioned questions. The first study (Chapter 2) is concerned with an alternative methodology of quantifying moment risk premiums in a robust and flexible manner. The second study (Chapter 3) focuses on the structure and a decomposition of traditional moment risk premiums and analyzes the premiums' ability to predict future market excess returns. Lastly, the third study (Chapter 4) assesses various strategies for harvesting the variance risk premium with an explicit focus on the suitability to accumulate capital for long-term investors. In the following, I summarize the individual chapters of this thesis.

The first study, *Quantile Risk Premiums*[4], examines quantile-based moment risk premiums, which provide robust and flexible alternatives to variance, skewness, and kurtosis risk premiums. Central moments, which are so far the standard measure to quantify moment risk premiums, have some drawbacks. One particular shortcoming is their sensitivity to outliers (Kim and White 2004) which highly influences moment estimates. Additionally, there is work that argues about potential non-existence of moments for specific distributions (Mandelbrot 1963; Fama 1965). Needless to say, that in such a case moment risk premiums are not well defined. Thus, it is reasonable to shift the perspective from central moments to an alternative perspective based on quantiles of a return distribution. Quantiles are a robust and flexible alternative to characterize a distribution and they exist for every distribution, irrespective of the existence of moments.

In the study, we introduce a new class of synthetic derivatives, the so-called quantile swaps, as a tool to quantify risk premiums. A basic q-quantile swap pays the difference between the realized and risk-neutral probabilities of a return below an arbitrary quantile q of the risk-neutral distribution. By combining quantile swaps with different quantile specifications, we can mimic some well-known location, dispersion, asymmetry, and steepness measures from robust statistics that serve as counterparts to mean, variance, skewness, and kurtosis. We estimate and analyze these quantile-based premiums in an empirical study for the S&P 500 index option market.

The empirical study for the S&P 500 index detects significant tail-associated total dispersion premiums and a downside dispersion premium that is not confined to the tail area. These

[4]This study is joint work with Felix Brinkmann and Olaf Korn.

premiums are truly option-specific as they remain almost unchanged when hedging out any linear exposure. Moreover, we find a significant asymmetry risk premium that is also truly option-specific and driven by both downside and upside components of the swap contract. Lastly, we do not find any evidence for a steepness risk premium. When controlling for traditional moment risk premiums, we find that the total dispersion risk premium loses its significance, whereas both a tail and a center asymmetry premium remain significant. Thus, our approach detects a novel asymmetry-related premium that traditional moment risk premiums fail to measure and that is robust to various quantile specifications.

In the second study, entitled *Decomposed Higher-Moment Risk Premiums and Market Return Predictability*, I turn to the structure of moment risk premiums and their predictive power for future market excess returns. In general, moment risk premiums are known to be suitable market excess return predictors. Within the literature, the variance risk premium, decomposed variance risk premiums, and higher-moment risk premiums up to the kurtosis have been assessed with respect to their predictive power for the subsequent market risk premium. The variance risk premium is found to be a good short-term predictor for horizons up to six months, while considering decomposed variance risk premiums or higher-moment risk premiums improves the long-term predictability (Bollerslev, Tauchen, and Zhou 2009; Kilic and Shaliastovich 2019; Fan, Xiao, and Zhou 2020). In the course of this chapter, the dimensions *decomposition* and *higher-moment risk premiums* are synthesized and the predictive power for future market excess returns by means of *decomposed higher-moment risk premiums* is analyzed. In doing so, the study provides novel insights into the fine structure of moment risk premiums as well as how and to what extent these particular premiums can be used as information signals for return predictions.

The chapter makes both important theoretical and empirical contributions to the literature. First, I propose a simple decomposition of total higher-moment risk premiums that can be applied to all higher-moment risk premiums. Second, I derive a new measure for a kurtosis risk premium that can be determined through a trading strategy according to the methodology from Kozhan, Neuberger, and Schneider (2013). Third, decomposed higher-moment risk premiums are analyzed descriptively. In an empirical study of the S&P 500 index option market I find premiums for variance, skewness, and kurtosis. These premiums are significantly negative when looking at downside and total premiums, while they are significantly positive for upside premiums. Furthermore, the results suggest that the premiums have highly skewed and leptokurtic distributions and that they are strongly correlated with each other. Fourth, turning to the predictive power for subsequent market excess returns, the study provides evidence for increased in-sample predictability through decomposition of total higher-moment risk premiums compared to the known prediction models from the relevant literature. This holds in particular for prediction horizons of six months or longer. Lastly, the out-of-sample predictive power shows to be strongest for the total variance risk premium when considering prediction horizons up to six months. A combination forecast of downside higher-moment risk premiums yields best predictions at horizons larger than six months.

The third paper, *How to Harvest Variance Risk Premiums for the Long-term Investor?*[5], is concerned with an analysis of investment strategies that expose investors to the "variance factor". It is well known that there exists a significantly negative variance risk premium in the S&P 500 index option market (Carr and Wu 2009; Bollerslev et al. 2009). However, the realized premium has a very specific structure as it fluctuates at low negative levels and takes on very high positive values during rare extreme events. So far, it has not yet been answered whether it is possible at all for the long-term investor to harvest the variance risk premium, or whether this structure impedes an effective long-term capital accumulation. The goal of this study is to provide novel evidence into whether and how investors can use the variance risk premium for long-term capital accumulation and it thereby enhances the understanding of the variance risk premium's investment implications.

The construction of variance risk premium strategies involves three major problems that are addressed in the course of this chapter. The main reason for why variance factor-specific issues emerge is that, different from classical factor strategies, the variance factor cannot be earned through a long-short portfolio in stocks but only through short-positions in derivatives. The chapter describes these problems, which we call the *payoff problem*, the *leverage problem*, and the *finite maturity problem*, and depicts trade-offs that have to be considered when constructing this type of strategies. Afterwards, the study introduces seven strategies that are in principle suited to harvesting the variance risk premium.

In an empirical analysis of the S&P 500 the strategies are analyzed and compared both with each other and with a variety of other factor investment strategies. The study finds that the strategies substantially differ from each other in terms of return and risk because of different margin requirements that lead to different factor exposures and (in some cases) excessive risk taking. However, even at identical levels of ex-ante factor exposure, there are differences in the strategies that arise from different instruments which the strategies comprise. Overall, and compared

[5] This study is joint work with Olaf Korn and Gabriel J. Power.

to other factor investment styles, the results suggest that, when properly designed, variance strategies are attractive strategies for long-term investors because they have continuously earned premiums throughout the whole sample period and quickly recovered from large drawdowns. Additionally, variance strategies can be useful complements to a market investment.

The dissertation and its contributions are relevant for both researchers and practitioners in the fields of (empirical) asset pricing as well as asset and risk management. For an audience in academia the thesis provides answers to some important and previously unanswered questions about the robustness and the fine structure of moment risk premiums as well as novel evidence on their predictive power for future market excess returns. Moreover, it presents approaches for harvesting the variance risk premium through trading strategies. Besides that, it shows limitations and drivers of traditionally defined moment risk premiums. Future research can build upon these results and analyze premiums for other asset classes, countries or term structures. Practitioners, such as risk managers, asset managers, and institutional investors, can likewise benefit from the knowledge gained in these studies. As it provides guidance on how to accumulate capital on a long-term basis, it is the nature of Chapter 4 to describe and address practical issues of long-term investors such as pension funds. Nevertheless, the other chapters also provide important practical insights, as they strengthen the intuition about which risks are really priced in option markets. This can help risk managers identify relevant risk factors and to design appropriate strategies to manage these risks.

2 Quantile Risk Premiums

Joint work with Felix Brinkmann and Olaf Korn.

Under review at the Review of Derivatives Research.

Abstract

This paper studies quantile-based moment premiums. The quantile-based approach delivers robust and flexible alternatives to premiums for variance, skewness and kurtosis risk and enhances our understanding of the pricing of risks in derivatives markets. To quantify these premiums, the paper introduces a new class of synthetic derivatives contracts: quantile swaps. Such contracts mimic quantile-based moment measures from robust statistics. An empirical study of index options detects two distinct premiums for dispersion and asymmetry, but no premium for steepness. The premium for dispersion can be explained by traditional moment risk premiums, whereas the asymmetry premium is a novel premium that our approach is able to detect. Moreover, we disaggregate the overall premiums into upside and downside premiums. The downside dispersion premium is not restricted to the lower tail but is also observed in the center of the distribution. However, at the center, this downside premium is partly offset by an upside discount, explaining why the overall premium is mainly a tail effect. Overall, as some of our findings differ markedly from results obtained with traditional moment swaps, they are a warning to interpret moment premiums cautiously.

Acknowledgements: We thank Mobina Shafaati, participants of the 2019 FMA European Conference, Glasgow, the 2019 German Finance Association (DGF) Ph.D. Workshop, Essen, the 2020 virtual FMA Conference, as well as seminar participants at the University of Goettingen for helpful comments and suggestions. Vitus Benson provided excellent research assistance. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Deutsche Bundesbank or its staff.

2.1 Introduction

Derivatives offer unique opportunities to quantify risk premiums. If markets are frictionless, forward prices equal expected future spot prices in a risk-neutral world, and average returns provide direct evidence on the compensation for risk. Other derivatives allow for the measurement of risk premiums with respect to higher moments of an asset's return distribution. Carr and Wu (2009) introduce a synthetic variance swap whose expected payoff equals the difference between the physical and risk-neutral variance. In the same spirit, Kozhan et al. (2013) and Khrashchevskyi (2020) develop synthetic skewness and kurtosis swaps. Moments such as variance and skewness are useful characteristics of a return distribution, but they have some drawbacks; in particular, estimates of higher moments are quite sensitive to outliers (Kim and White 2004).[6] Moreover, some moments may not even exist for certain assets,[7] and the corresponding moment risk premiums may not be well defined. Therefore, a complementary perspective is useful, which shifts the attention from standard moments and the corresponding moment premiums to the quantiles of a return distribution. By taking this perspective, we provide new evidence on a fundamental question in asset pricing: how are risks priced in derivatives markets?

As a tool to quantify risk premiums, this paper uses a new class of synthetic derivatives, which we call quantile swaps. The basic q-quantile swap pays the difference between the realized and risk-neutral probabilities of a return below the q-quantile of the risk-neutral distribution (RND). Specific portfolios of quantile swaps mimic some well-known location, dispersion, asymmetry, and steepness measures from robust statistics, such as the median or the interquartile range. These portfolios allow us to quantify the corresponding quantile-based moment risk premiums through time-series averages of realized payoffs. In an empirical study for the Standard & Poor's 500 (S&P 500) option market, we estimate and analyze these premiums.

Our paper contributes in different ways to understanding risk premiums in derivatives markets. A first important aspect is *robustness.* A fundamental robustness property of our quantile-based method to measure risk premiums is its general applicability, irrespective of the existence of moments. Moreover, some robustness properties of quantile-based measures of location, dispersion,

[6] Conrad, Dittmar, and Ghysels (2013) address the severe problems to estimate higher moments of stock returns by applying a quantile-based measure of skewness. However, such problems do not just occur when estimates use historical return series. As Ammann and Feser (2019b) show, quantile-based moment measures may also be a valuable alternative to traditional moment measures for estimating risk-neutral moments from a cross section of option prices.

[7] Classical papers that cast some doubt on the existence of moments are Mandelbrot (1963) and Fama (1965).

asymmetry, and steepness carry over to the corresponding quantile swap portfolios.[8] As quantile swaps' payoffs are capped from above and below, they are less sensitive to outliers. Given the advantages in terms of robustness, one important question in this study is whether quantile swaps lead to the same conclusions about moment premiums in options markets as traditional moment swaps.

A second important aspect is the *flexibility* of our quantile-based approach. Quantile-based higher-moment swaps are not just single instruments: they are classes of instruments. For example, the interquantile range, as a classical measure of dispersion, can be implemented using different quantiles. In particular, this can be done through quantiles reaching out to the tail area of the distribution. The same idea can be applied to third and fourth moments as well. Starting with portfolios of quantile swaps that mimic the quantile-based skewness and kurtosis measures by Hinkley (1975) and Moors (1988), respectively, we use the flexibility of our approach to construct and analyze the corresponding tail contracts. Another idea is to split the return distribution into its positive and negative parts, leading to upside and downside quantile-based moment swaps.[9] Again, upside and downside quantile-based moment swaps are classes of instruments that allow us to focus either on the tails or on the center of the distribution. In general, our approach provides an easy way to investigate whether moment-risk premiums can be traced back to specific areas of the return distribution. In particular, there is no need to model the underlying's price dynamics to disentangle premiums referring to, for example, a diffusion component or a jump component.

Our empirical analysis for the S&P 500 options market provides the following main results: (i) We find a significant overall dispersion premium, which is associated with the tail areas of the return distribution. This premium remains almost unchanged when hedging out any linear exposure from the underlying. In this sense, it is a truly option-specific premium. After controlling for traditional moment risk premiums, this premiums loses its significance. (ii) Disaggregation of the overall dispersion premium into upside and downside premiums shows that the overall premium is mainly due to the downside part. Importantly, the downside dispersion premium itself is not restricted to the lower tail but is also observed in the center of the distribution. However, at the center, this downside premium is partly offset by an upside discount, which explains why we find an overall premium only in the tails. (iii) We also find a significant asymmetry premium,

[8] Ammann and Feser (2019a), p. 449, also stress the robustness of a quantile-based approach to risk measurement.

[9] In the context of classical moment swaps, the idea of upside and downside variance swaps, as introduced by Carr and Lewis (2004), has recently received much attention. For example, see Da Fonseca and Xu (2017), Feunou et al. (2018), Huang and Li (2019), Kilic and Shaliastovich (2019), and Held et al. (2020).

which remains strong after hedging out linear exposure. This premium is driven both by the downside and the upside part of the return distribution. When controlling for traditional moment risk premiums, the asymmetry premium maintains significance; thus, our analysis documents a distinct premium that is not captured by traditional moment risk premiums. (iv) There is no indication for a steepness premium, even if we restrict our attention to the tail areas or the downside part of the distribution. (v) The signs of the premiums found in this study correspond with reasonable moment preferences: that is, investors favor higher first and third moments, and they dislike higher second moments. Therefore, the premiums are consistent with the economic rationale that moment premiums are a compensation for moment risk. (vi) Finally, there is a clear positive relation between the payoffs of dispersion swaps and traditional variance swaps, particularly if we consider the tail contract. In contrast, the returns of an asymmetry swap and a traditional skewness swap are negatively related and only significant for the tail contract, while the steepness swap and the kurtosis swap are only weakly positively correlated. Quantile-based moment swaps and traditional moment swaps clearly measure different things if we go to third and fourth moments.

Overall, this paper provides new evidence on the pricing of risks in options markets. The quantile-based approach confirms the existence of a positive market risk premium and a negative variance risk premium (Bakshi and Kapadia 2003; Carr and Wu 2009). A new finding with respect to the second-moment premium is that the overall premium is confined to the tails but the downside premium is not. Other new findings refer to skewness and kurtosis risk premiums. We find an asymmetry premium that cannot be fully explained by other moment premiums, in contrast to the traditional skewness premium (Kozhan et al. 2013). Moreover, there is no evidence for a significant steepness premium, whereas the kurtosis premium is large and significant (Khrashchevskyi 2020). These results show that it clearly matters how one measures risk premiums for higher moments. They further warn us to be cautious when interpreting skewness and kurtosis risk premiums. The question remains as to what is actually priced if theses premiums strongly deviate from those obtained through a simple alternative approach that uses an intuitive partitioning of the return distribution.

The remainder of this paper is organized as follows. In Section 2.2, we introduce the quantile swap, which is the basic building block for all our investigations. Beyond that, we suggest specific portfolios of quantile swaps, which are designed to quantify moment risk premiums up to the fourth moment. Section 2.3 then describes the data and the methodology of our empirical study, while the main empirical results on quantile-based moment premiums are presented in

Sections 2.4 and 2.5. Section 2.6 provides a comparison with premiums derived from traditional moment swaps. Finally, Section 2.7 concludes the paper.

2.2 Quantile-Based Moment Swaps

The deviation between an asset's physical and risk-neutral return distributions contains information on the compensation for risk. A way to capture this information is through moment premiums, that is, differences between physical and risk-neutral moments. The market risk premium is a classic example of such a moment premium, providing the difference between the expected return of a market index under the physical measure and its risk-neutral counterpart, the risk-free rate. The same idea has been applied to higher moments, leading to the variance risk premium, the skewness risk premium, and the kurtosis risk premium. However, the use of traditional moments is only one way to characterize a return distribution. Alternatively, it can be described through quantiles. This section develops the idea of quantile-based moment premiums and introduces financial contracts to quantify them.

2.2.1 Quantile Swaps: The Building Blocks

Consider an asset with random return r. A basic partition of the asset's return distribution distinguishes between the areas to the left and to the right of the q-quantile. Denote the q-quantile under the risk-neutral measure $\mathbb{Q}$ by $Q^{\mathbb{Q}}(q)$. By construction, the risk-neutral probability of a return below $Q^{\mathbb{Q}}(q)$ is q. However, the probability under the physical measure $\mathbb{P}$ may be different. Denote this latter probability as $p := F^{\mathbb{P}}(Q^{\mathbb{Q}}(q))$, where $F^{\mathbb{P}}$ is the cumulative distribution function under the physical measure. The difference $p - q$ reflects risk adjustments associated with the area of the return distribution below $Q^{\mathbb{Q}}(q)$ and we suggest to use this difference as a premium measure. Both the sign and magnitude of this measure are easy to interpret. A negative difference arises if the probability of observing returns below the risk-neutral q-quantile is lower under the physical measure than under the risk-neutral measure. If investors dislike the risk of low returns (below the q-quantile) and seek to avoid it, such a negative difference is expected.

To quantify the suggested measure empirically, we follow the same idea as Carr and Wu (2009) to quantify the variance risk premium. We suggest a derivative contract whose expected payoff equals $p - q$, which we call the q-quantile swap. The holder of the contract receives the "realized" probability that the return is below the risk-neutral q-quantile. This is the floating leg of the swap. In exchange for this realized probability, the holder pays the risk-neutral probability q, the fixed leg of the swap. The payoff of the q-quantile swap, $S(q)$, is defined as:

$$S(q) := \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q)\right] - q\right) L, \tag{2.1}$$

where $\mathbb{1}[\cdot]$ is an indicator function that takes a value of one if the condition is fulfilled and is zero otherwise. The notional amount of the swap is denoted by L. Equation (2.1) shows what the “realized" probability of a return below $Q^{\mathbb{Q}}(q)$ means. If the realized return is lower than $Q^{\mathbb{Q}}(q)$, the realized probability takes a value of one; otherwise, the realized probability is zero. By construction, the q-quantile swap has an initial value of zero. It can also be interpreted as a binary option that pays L dollars if the realized return is below $Q^{\mathbb{Q}}(q)$ and nothing otherwise. The payment of the corresponding option premium, qL, would be deferred to date T. For low values of q, the q-quantile swap is also similar to a bear spread (Lu and Murray 2019).

At the beginning of the return period, the payoff of the q-quantile swap is a random variable. Taking expectations on both sides of Equation (2.1) delivers:

$$E^{\mathbb{P}}\left(S(q)\right) = \left(F^{\mathbb{P}}(Q^{\mathbb{Q}}(q)) - q\right) L = (p - q)L, \tag{2.2}$$

that is, the expected payoff of the swap is the desired risk premium, scaled by the notional L. This property allows us to estimate risk premiums through time-series averages of realized payoffs. That said, the calculation of realized payoffs requires the knowledge of the quantile $Q^{\mathbb{Q}}(q)$ of the risk-neutral return distribution. In theory, under the assumption of complete markets, this risk-neutral quantile is known because the risk-neutral distribution is known. In practice, the quantile has to be estimated from available option prices. The estimation method used in our empirical study is explained in Section 2.3.

As an illustration, Part (a) of Figure 2.1 depicts the payoff diagram of the 0.5-quantile swap, which we call the median swap. The notional is set to $L = 1$ dollar. The holder of the median swap either receives an amount of 50 cents or pays an amount of 50 cents, depending on whether the realized return of the underlying is below or above the risk-neutral median return. The payoff diagram highlights the difference between the median swap and a short position in a forward contract, which can be interpreted as the corresponding "mean swap".[10] In statistics, the median serves as a robust alternative to the mean as a location measure. It is less sensitive to outliers than the mean because all that matters is whether the return is below or above the median. Similarly, the payoff of the median swap is less sensitive to extreme returns than the corresponding mean swap because its payoff is capped from above and below. Therefore, the median swap can be seen as a robust alternative to the mean swap when measuring risk premiums associated with the location (first moment) of the underlying's return distribution.

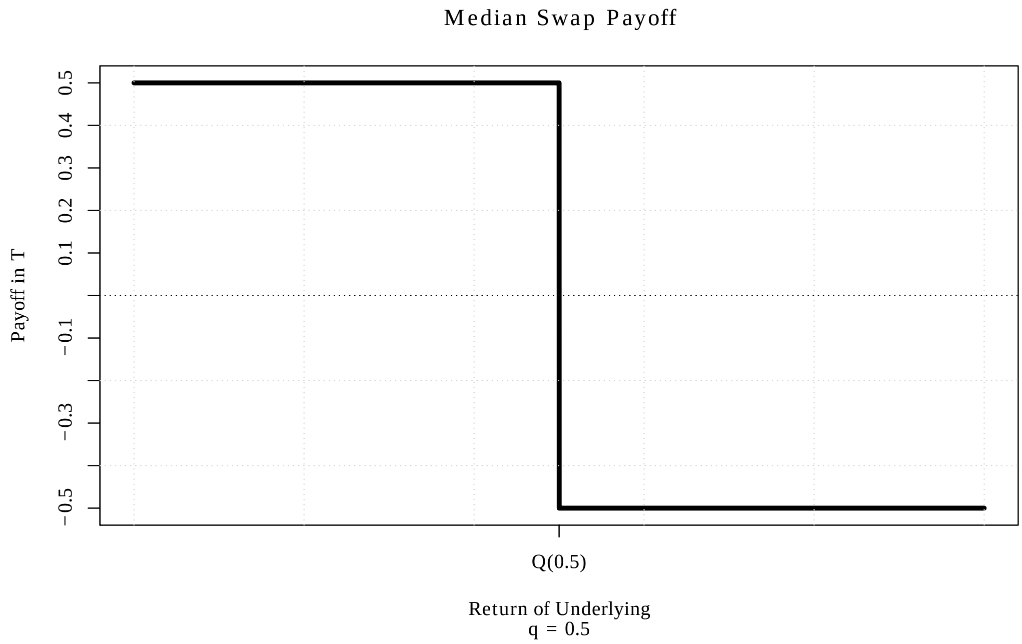
2.2.2 Quantile-Based Moment Swaps for Higher Moments

Both forward contracts and median swaps measure risk premiums with respect to the central tendency of the return distribution. However, the corresponding location measures—the mean and the median—are different. The idea of a quantile-based moment swap can also be applied to higher moments. In statistics, there exist quantile-based measures of dispersion, asymmetry, and steepness of a distribution, serving as robust alternatives to traditional higher moments. This subsection introduces swap contracts that mimic such quantile-based measures. These swap contracts can be replicated by portfolios of basic quantile swaps and allow us to estimate the corresponding risk premiums.

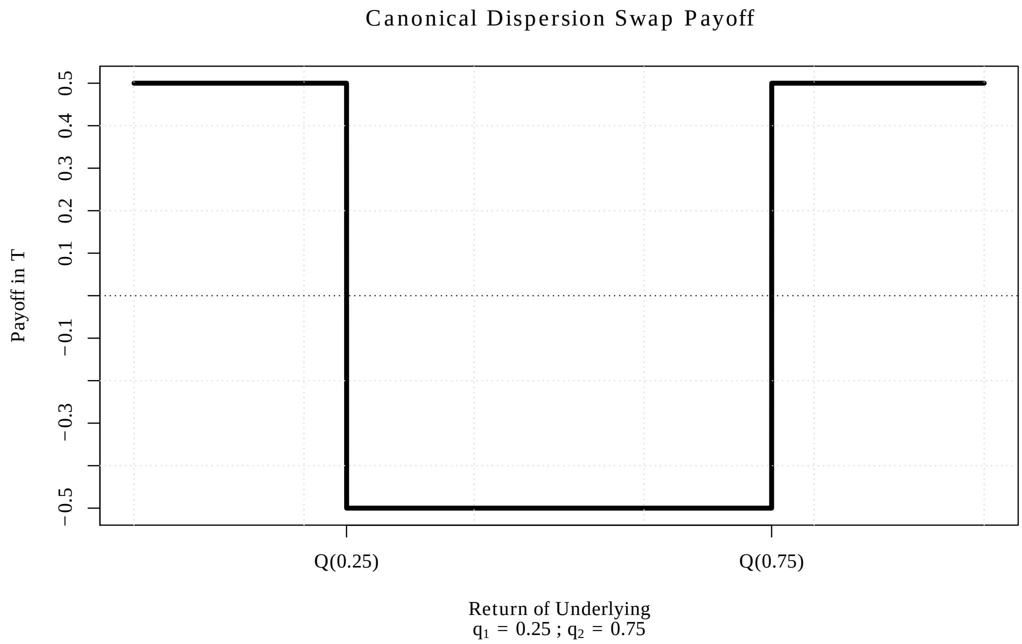
Dispersion Swaps A classic quantile-based dispersion measure is the interquantile range, the difference between a higher quantile $Q(q_2)$ and a lower quantile $Q(q_1)$. The interquantile range, like the median, corresponds to a specific partition of the return distribution. This partition does not simply distinguish between high returns and low returns but discriminates between large returns (high and low) and returns in the center of the distribution. There are three distinct areas of the distribution: returns below $Q(q_1)$, returns between $Q(q_1)$ and $Q(q_2)$, and returns above $Q(q_2)$. Under the risk-neutral measure, the probability that a return falls within the interquantile range is $q_2 - q_1$. Thus, the probability of a return outside that range is $1 - (q_2 - q_1)$. However, under the physical measure these probabilities can be different. In particular, the probability of a return being outside the (risk-neutral) interquantile range is $1 - (F^{\mathbb{P}}(Q^{\mathbb{Q}}(q_2)) - F^{\mathbb{P}}(Q^{\mathbb{Q}}(q_1))) = 1 - (p_2 - p_1)$. Risk premiums associated with the area of the distribution outside the interquantile range could, therefore, be represented by the difference $[1 - (p_2 - p_1)] - [1 - (q_2 - q_1)] = (p_1 - p_2) - (q_1 - q_2)$. If this difference is negative, the risk-neutral

[10]The interpretation of a forward contract as a first-moment swap is suggested by Schneider (2015).

Figure 2.1: Payoffs of Quantile-Based Canonical Moment Swaps

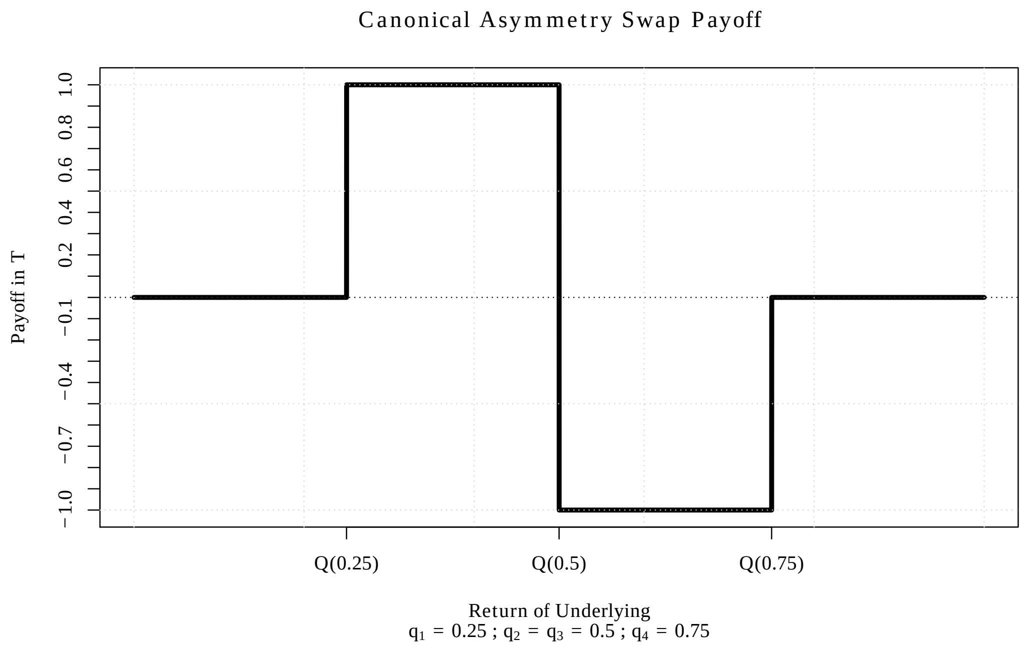


(a) Median Swap Payoff

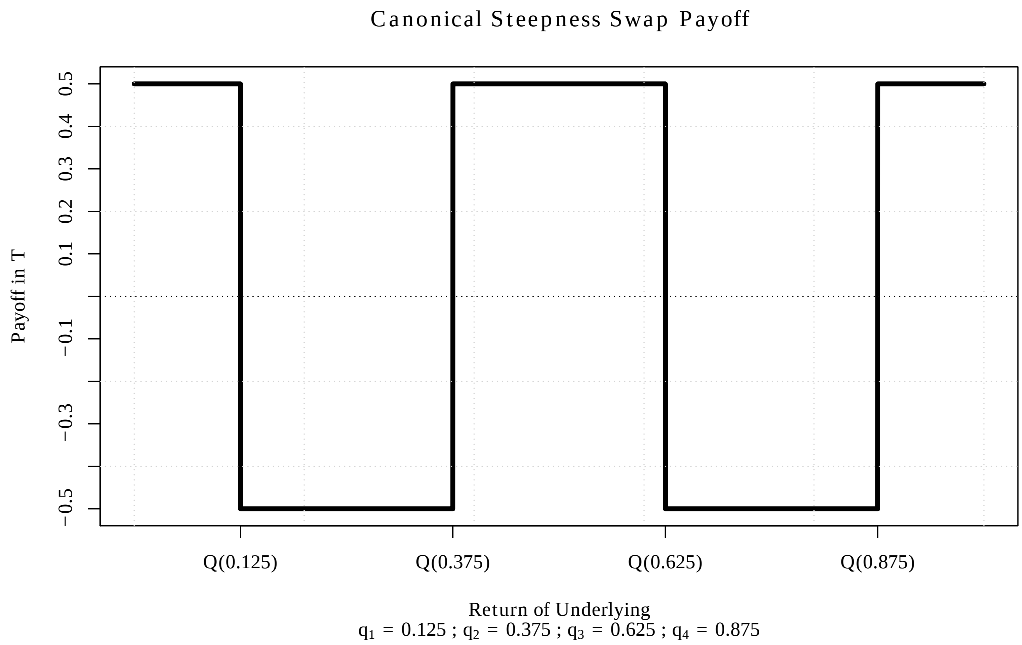


(b) Dispersion Swap Payoff

Figure 2.1: continued



(c) Asymmetry Swap Payoff



(d) Steepness Swap Payoff

Note: This figure shows payoff diagrams for the median swap ($q = 0.5$) (Part (a)), the canonical dispersion swap with $q_1 = 0.25$, $q_2 = 0.75$ (Part (b)), the canonical asymmetry swap with $q_1 = 0.25$, $q_2 = q_3 = 0.5$, $q_4 = 0.75$ (Part (c)), and the canonical steepness swap with $q_1 = 0.125$, $q_2 = 0.375$, $q_3 = 0.625$, $q_4 = 0.875$ (Part (d)). The x-axis shows the return of the underlying given as a quantile of its risk-neutral return distribution and the y-axis shows the swaps' payoffs at maturity in T.

probability of large (positive or negative) returns is higher than the physical probability, an indication that investors are willing to pay a premium to protect themselves against these events.

As a financial instrument to measure such premiums, we suggest a dispersion swap (DS) with the following payoff:

$$
\begin{aligned}
DS(q_1, q_2) := &\left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_1)\right] - q_1\right) L \\
&- \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_2)\right] - q_2\right) L.
\end{aligned}
\tag{2.3}
$$

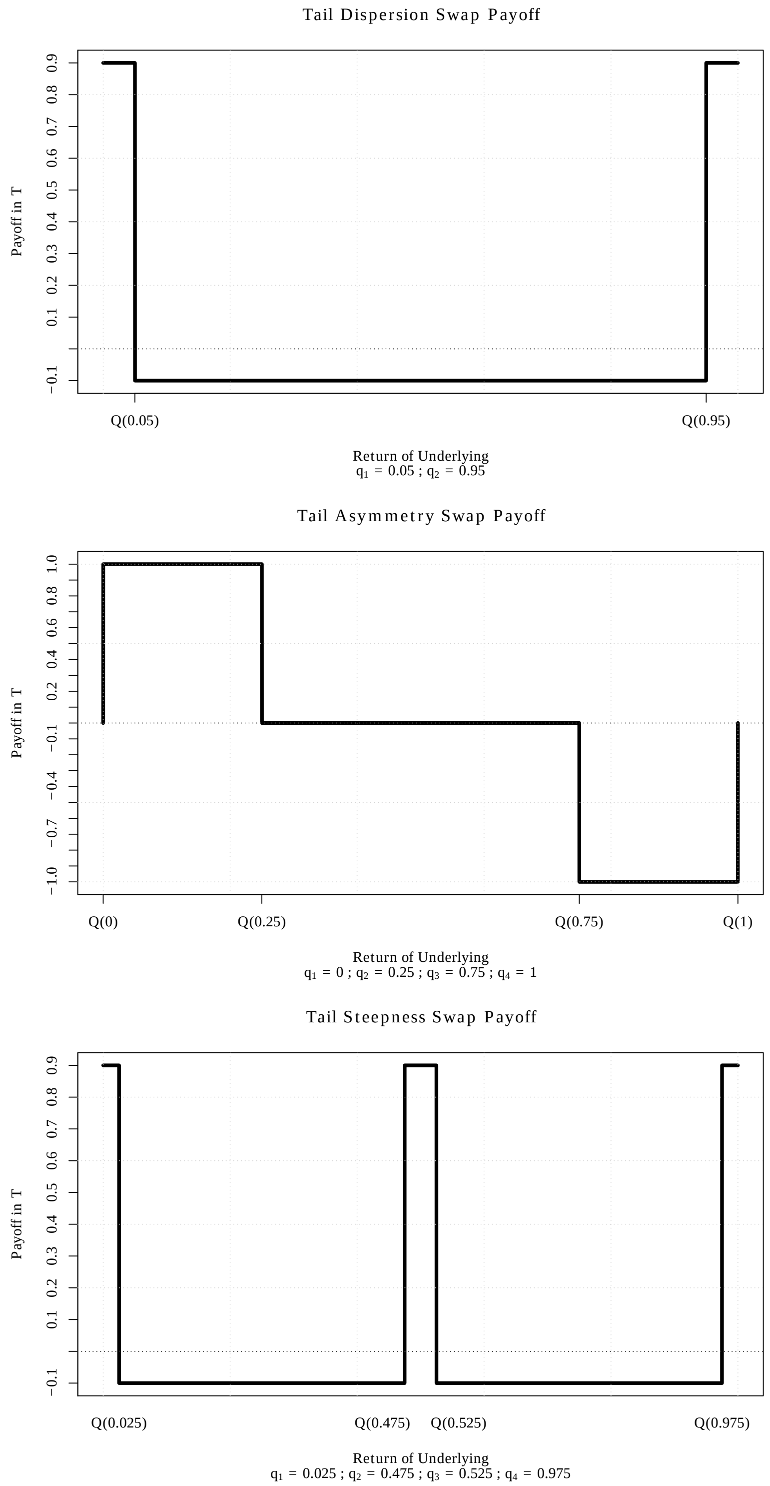
The expected payoff of the dispersion swap is just the desired premium measure, scaled by the notional. Equation (2.3) highlights that the dispersion swap is a simple portfolio of quantile swaps. It consists of a long position in the q_1-quantile swap and a short position in the q_2-quantile swap. Part (b) of Figure 2.1 depicts the payoff of the "canonical" dispersion swap, the interquartile range swap, that is, a swap with $q_1 = 0.25$ and $q_2 = 0.75$. That said, we are not restricted to the interquartile range because q_1 and q_2 can be freely chosen. In particular, we can concentrate on the tails of the distribution. As an illustration, the upper part of Figure 2.2 shows the payoff of a "tail" dispersion swap, where $q_1 = 0.05$ and $q_2 = 0.95$.

Part (b) of Figure 2.1 suggests an intuitive decomposition of the dispersion premium into an upside premium and a downside premium through up- and downside dispersion swaps.[11] A downside dispersion swap has the same payoff than the dispersion swap if the return of the underlying is below the risk-neutral median and makes zero payments otherwise. Conversely, an upside dispersion swap's payoff is identical to the dispersion swap's payoff, conditional on above median returns. If realized returns are below the risk-neutral median, then no payment occurs. By construction, a dispersion swap is simply a portfolio of one upside dispersion swap and one downside dispersion swap, and the overall dispersion premium is the sum of the up- and downside premiums. Such decomposition helps to better understand the origins of an overall premium. Also note that different dispersion swaps have their own up- and downside versions; in particular, there are "canonical" as well as "tail" up- and downside dispersion swaps.

Asymmetry Swaps To capture differences in asymmetry between the risk-neutral and physical distributions, we suggest another partition of the return distribution. This partition compares

[11] Analogous decompositions of the variance risk premium into an up- and downside premium have recently been suggested and studied by Da Fonseca and Xu (2017), Feunou et al. (2018), Huang and Li (2019), Kilic and Shaliastovich (2019), and Held et al. (2020).

Figure 2.2: Payoffs of Quantile-Based Tail Moment Swaps



Note: This figure shows payoff diagrams for the tail dispersion swap with $q_1 = 0.05$, $q_2 = 0.95$, the tail asymmetry swap with $q_1 = 0$, $q_2 = 0.25$, $q_3 = 0.75$, $q_4 = 1$, and the tail steepness swap with $q_1 = 0.025$, $q_2 = 0.475$, $q_3 = 0.525$, $q_4 = 0.975$. The x-axis shows the return of the underlying given as a quantile of its risk-neutral return distribution and the y-axis shows the swaps' payoffs at maturity in T.

some area to the left of the median with a corresponding area to the right of the median. Specifically, consider the quantiles $Q(q_1)$, $Q(q_2)$, $Q(q_3)$, and $Q(q_4)$, which define the partition. The values q_1 and q_2 are below the median, and q_3 and q_4 are above the median under the risk-neutral distribution; meanwhile, q_1 and q_2 determine the left-side area that we want to compare with the right-side area, as determined by q_3 and q_4. We require $q_{0.5} - q_2 = q_3 - q_{0.5}$ to ensure an equal probability mass between the median and the nearest points of the left- and right-side areas. We further require that $q_4 - q_3 = q_2 - q_1$. This condition ensures that the probabilities for a return falling into the left-side and right-side areas are exactly the same under the risk-neutral measure. However, these probabilities can be different under the physical measure, and any differences provide information on differences in symmetry between the risk-neutral and physical distributions. Specifically, the probability for returns to fall in the left-side and right-side areas under the risk-neutral measure are $q_2 - q_1$ and $q_4 - q_3$, respectively. Under the physical measure, they are $p_2 - p_1$ and $p_4 - p_3$. Therefore, a natural measure of asymmetry premiums is $[(p_2 - p_1) - (q_2 - q_1)] - [(p_4 - p_3) - (q_4 - q_3)]$. If the physical distribution has less probability mass in the area to the left of the median and more probability mass in the area to the right of the median than the risk-neutral one, the physical distribution is more skewed to the right and we obtain a negative premium.

The corresponding quantile-based moment swap, which we call an asymmetry swap (AS), has the following payoff function:

$$
\begin{aligned}
AS(q_1, q_2, q_3, q_4) := & \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_2)\right] - q_2\right) L \\
& - \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_1)\right] - q_1\right) L \\
& - \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_4)\right] - q_4\right) L \\
& + \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_3)\right] - q_3\right) L.
\end{aligned}
\tag{2.4}
$$

Equation (2.4) shows that the asymmetry swap is simply a portfolio of long positions in the q_2- and q_3-quantile swaps and short positions in the q_1- and q_4-quantile swaps. Part (c) of Figure 2.1 depicts the payoff profile of our "canonical" asymmetry swap, which resembles the classical quantile-based skewness measure of Bowley (1920). This measure is based on quartiles and was generalized by Hinkley (1975). For the corresponding asymmetry swap, $q_1 = 0.25$, $q_2 = q_3 = 0.5$, and $q_4 = 0.75$. As shown in Figure 2.1, the canonical asymmetry swap compares two areas near the center of the distribution. The flexibility of our approach also allows us to construct a corresponding "tail" contract, depicted in the middle part of Figure 2.2. To define this contract, we leave the distances $q_2 - q_1$ and $q_4 - q_3$ unchanged but shift the left- and right-side areas as far as possible to the tails of the distribution. Concretely, we use $q_1 = 0$, $q_2 = 0.25$, $q_3 = 0.75$, and $q_4 = 1$ to construct the tail asymmetry swap.

We also consider an up- and downside decomposition and split the asymmetry swap's payoff in an upper part above the median and a lower part below the median.[12] This results in a downside asymmetry swap component that comprises short and long positions in quantile swaps with quantiles q_1 and q_2, respectively, and an upside asymmetry swap component that comprises long and short positions in quantile swaps with quantiles q_3 and q_4, respectively. The sum of these components yields the asymmetry swap. Studying the components is informative as it sheds light on whether the total asymmetry swap payoff is driven be the downside or the upside part of the distribution.

Steepness Swaps A return distribution exhibits high steepness if there is much probability mass both in the tails and in the center of the distribution but little mass in the intermediate parts. The idea of steepness translates well into a partition of the return distribution that distinguishes the tail areas below the quantile $Q(q_1)$ and above the quantile $Q(q_4)$, a central area between the quantiles $Q(q_2)$ and $Q(q_3)$, and the intermediate areas between the quantiles $Q(q_1)$ and $Q(q_2)$, as well as $Q(q_3)$ and $Q(q_4)$. With respect to potential risk premiums associated with steepness, we ask whether the physical probabilities of observing a return in these specific areas differ from the risk-neutral ones. In particular, the question is whether physical probabilities are higher or lower than risk-neutral probabilities in the tail plus center areas. If they are lower, there is a negative steepness premium, which may be seen as an indication of steepness risk aversion.

[12] Analogously, Da Fonseca and Xu (2017) suggest an upside- and downside version of a classical skewness swap. Schneider et al. (2020) use the decomposition of the risk-neutral skewness into upper and lower skewness to predict the co-skewness between individual stocks and the market.

To quantify a quantile-based steepness risk premium, we suggest the steepness swap (SS) with the following payoff:

$$
\begin{aligned}
SS(q_1, q_2, q_3, q_4) := & \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_1)\right] - q_1\right) L \\
& - \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_2)\right] - q_2\right) L \\
& + \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_3)\right] - q_3\right) L \\
& - \left(\mathbb{1}\left[r^{\mathbb{P}} < Q^{\mathbb{Q}}(q_4)\right] - q_4\right) L.
\end{aligned}
\tag{2.5}
$$

Part (d) of Figure 2.1 depicts this payoff function for the "canonical" steepness swap, resembling the quantile-based steepness measure by Moors (1988). This measure is based on the octiles of the distribution—that is, $q_1 = 0.125$, $q_2 = 0.375$, $q_3 = 0.625$, and $q_4 = 0.875$—and provides a robust alternative to the traditional kurtosis measure. The realized payoff of the swap delivers an estimate of the differences in probability for the tail plus center areas of the distribution under the physical versus risk-neutral measures. As shown from Equation (2.5), the steepness swap is just a portfolio of four quantile swaps. The q_1- and q_3-quantile swaps enter as long positions, whereas short positions are held in the q_2- and q_4-quantile swaps. A steepness swap can also be modified with respect to the quantiles used. In particular, we can construct "tail" steepness swaps that consider more extreme tails and a narrower center. The payoff of such a tail steepness swap with $q_1 = 0.025$, $q_2 = 0.475$, $q_3 = 0.525$, and $q_4 = 0.975$ is shown in the lower part of Figure 2.2. Moreover, it is useful to consider up- and downside steepness swaps. These are constructed in an analogy to up- and downside dispersion and asymmetry swaps by conditioning on above and below median returns, respectively. The performance of these contracts provides evidence on whether downside steepness risk is priced differently from upside steepness risk.

Two general points about quantile-based higher-moment swaps are worth noting. First, these swaps represent whole classes of instruments because they depend on the specific choice of quantiles $(q_1, q_2, \ldots)$. Therefore, they offer some flexibility for the study of risk premiums, particularly with respect to the tails of the distribution. Consider the "canonical" dispersion swap with $q_1 = 0.25$ and $q_2 = 0.75$ again. In this specification, the tails represent 50% of the probability mass. For the "tail" dispersion swap with $q_1 = 0.05$ and $q_2 = 0.95$, the tail area covers only 10% of the distribution. It is conceivable that risk premiums associated with these two contracts may be quite different.

Second, note that all the presented quantile-based moment swaps are constructed in such a way that standard moment preferences lead to negative premiums. As shown by Rubinstein (1973), standard utility functions imply that investors favor higher odd moments and dislike higher even moments. Therefore, hedging instruments against lower odd moments and higher even moments should sell at a premium and earn negative expected returns. Since all our quantile-based moments swaps (median swap, dispersion swap, asymmetry swap, and steepness swap) are designed as such hedging instruments—which can be seen from Figure 2.1—negative expected returns are the natural hypothesis for our empirical investigation.

2.3 Data and Methodology of the Empirical Study

We use S&P 500 index options data from the OptionMetrics data base, which provides historical closing quotes from the Chicago Board Options Exchange (CBOE). Our sample period starts in January 1996 and ends in December 2020. Interest rates, spot prices of the underlying index as well as implied dividend yields were also obtained from OptionMetrics.

For our analysis of quantile swaps, we compute estimates of quantiles of the S&P 500 risk-neutral distribution in a model-free fashion. We follow a standard procedure along the lines of Jiang and Tian (2005) and Chang et al. (2012). For every month in our data period, we use data of the first trading day after the third Friday of that respective month, which is the standard expiration date of options traded at the CBOE. This choice ensures the availability of option prices with a time to expiration close to our considered time horizon of one month. For our analysis, we concentrate on European options written on the S&P 500 spot index without special settlement. We use all out-of-the-money put and call options and require that the bid price is positive and not higher than the ask price.

For all remaining data points we compute Black and Scholes (1973) implied volatilities using the mid quotes.[13] Then, we perform a cubic spline fitting to obtain a smooth volatility curve for every considered point in time. Since the interpolation technique is only applicable between the lowest and highest available strike prices, we use a horizontal extrapolation above and below the observable strike levels, as suggested by, for example, Shimko (1993).[14] Ammann and Feser (2019b) show that such a cubic spline interpolation of the volatility curve with horizontal

[13] As Chang et al. (2012) note, use of the Black–Scholes formula does not imply that we assume the valuation model to hold. Rather, the Black-Scholes formula provides a tool to convert option prices into implied volatilities, and vice versa.

[14] Only about 5% of the realized returns of the underlying in our sample fall in the regions of the risk-neutral distributions (RNDs) that have been constructed by extrapolation.

extrapolation works well for the estimation of risk-neutral moments because it achieves a good trade-off between bias and variance of the resulting estimates.

Next, we compute a fine grid of 12,000 equally spaced strike prices on the interval $[0.001, 3 \cdot X_t]$, where X_t is the current index level. The corresponding implied volatilities are then converted back into prices. For our purposes, it is most convenient to convert the implied volatilities directly into the prices of digital options according to a Black–Scholes type formula.[15] Finally, we obtain the RND by compounding the prices of these digital options, as their prices represent discounted probabilities of the respective options being in the money at the maturity date. Given the RND, we compute the corresponding risk-neutral quantiles, covering a range between 2.5% and 97.5%, with a step width of 2.5 percentage points. Since we know the current index level, knowledge of the quantiles of the price distribution finally allows us to calculate the quantiles of the index return distribution.

In the next step, we determine the quantile swaps' realized payoffs. As discussed in Section 2.2, the payoff of the fixed leg simply equals the risk-neutral probability of index returns at maturity being below the respective quantile of the risk-neutral distribution, that is, the payoff of the fixed leg equals q. The payoff of the floating leg is the realized probability of index returns being below the level of the RND that corresponds to the chosen quantile. Thus, it is equal to one if the index return is below the respective quantile of the RND, and zero otherwise. Finally, the difference between the floating and fixed legs' payoffs yields the realized payoff of the quantile swap. To obtain higher-moment swaps' payoffs, we employ the payoffs of the corresponding quantile swaps. We build the portfolios introduced in Section 2.2 and combine the various long and short positions of quantile swaps with different quantile specifications q to obtain the realized payoffs.

We analyze quantile-based moment swaps' payoffs for long positions in the four respective moment swaps with a return horizon of one month and notional values of $L = 1$. This yields a sample with 299 monthly return periods. In addition, we consider upside and downside dispersion, asymmetry, and steepness swaps. As the swaps do not require any capital at initiation, their payoffs are stated in monetary units (gross returns $[GR]$). The average payoffs of the swaps over

[15] Alternatively, we could convert the implied volatilities back into call and put prices using the Black-Scholes formula and obtain the prices of digital options through no-arbitrage arguments. Given the smooth volatility curve, both approaches finally lead to the same prices.

time serve as estimates of the average premiums.[16] Since all swap contracts are constructed such that standard utility theory suggests negative premiums, we will speak of higher premiums if average swap payoffs take more negative values.

We also determine deltas for every swap specification to obtain an indication of how sensitive the value of the swap is to changes in the underlying. As Black-Scholes deltas depend on the unrealistic assumptions of the Black-Scholes model, we compute deltas in a model-free fashion by applying the following comparative static analysis: First, we use the RND obtained from observed option prices and shift the whole distribution to the right by a small amount (0.01). Then, we reprice the quantile swap of interest based on this shifted RND. Since the original value of the swap is zero, the new value under the shifted RND, divided by 0.01, provides an estimate of delta. To consider both price increases and decreases of the underlying, we repeat the calculation for a small price decrease and obtain a second delta estimate. Finally, we use the average of the two delta estimates. Essentially, this model-free method to estimate delta assumes that a small price change of the underlying just changes the location of the RND by this amount, leaving all other moments of the risk-neutral price distribution unchanged.

To compare quantile-based moment premiums with traditional moment premiums, we compute the market risk premium as well as the variance and skewness risk premiums following Kozhan et al. (2013). We consider the volatility contract and the skewness contract from Neuberger (2012) to determine realized and implied second and third moments. Moreover, we use the kurtosis contract that was recently introduced by Khrashchevskyi (2020) along the lines of Kozhan et al. (2013) to calculate realized and implied fourth moments.

2.4 Premiums of Quantile-Based Moment Swaps

2.4.1 Premiums of Basic Quantile Swaps

As a starting point of the empirical analysis, Table 2.1 provides evidence on the premiums of basic quantile swaps with $L = 1$ for varying quantile specifications q. Mean gross returns of the swap contracts—that is, the swaps' average payoffs with averages taken over time—are all negative up to $q = 0.950$ and only turn slightly positive for $q = 0.975$. Further, mean returns exhibit a U-shape. They are less negative (or even slightly positive) for extreme quantiles (close

[16] Our approach does not put any restrictions on the potential time variation of risk premiums. However, because we observe only one realized premium per month, our estimates are to be interpreted as average premiums over time.

to zero and close to one) and more negative for quantiles closer to the median, with the lowest average return of -16.92 cents for $q = 0.4$. Negative premiums are statistically significant (at least at the 5% level) for all quantile specifications up to $q = 0.825$. Thus, Table 2.1 provides robust evidence for a significant market risk premium, even if the quantile specification somehow moves away from the median swap, the canonical choice for the quantile-based measurement of a location premium.

Table 2.1: Premiums of Basic Quantile Swaps

q	Mean GR	t-statistic	Mean Delta
0.025	−0.0217	−6.4968	−0.0004
0.050	−0.0366	−5.6689	−0.0007
0.075	−0.0482	−5.0886	−0.0011
0.100	−0.0632	−6.2014	−0.0016
0.125	−0.0681	−4.7837	−0.0020
0.150	−0.0731	−4.6931	−0.0024
0.175	−0.0847	−5.2060	−0.0029
0.200	−0.0997	−5.8477	−0.0033
0.250	−0.1129	−5.7840	−0.0041
0.300	−0.1261	−6.4205	−0.0048
0.400	−0.1692	−7.0904	−0.0060
0.500	**−0.1622**	**−5.9601**	**−0.0069**
0.600	−0.1485	−5.3621	−0.0073
0.700	−0.1381	−5.2047	−0.0071
0.750	−0.1045	−3.8186	−0.0067
0.800	−0.0809	−3.2770	−0.0061
0.825	−0.0625	−2.5380	−0.0057
0.850	−0.0373	−1.5865	−0.0051
0.875	−0.0322	−1.4667	−0.0045
0.900	−0.0237	−1.2035	−0.0038
0.925	−0.0120	−0.6790	−0.0029
0.950	−0.0002	−0.0133	−0.0020
0.975	0.0083	1.1482	−0.0008

Note: This table provides evidence on the premiums of basic quantile swaps with notional values of one on the S&P 500 index for varying quantiles. The first column shows the swaps' quantile specifications (q), ranging from 0.025 to 0.975, with some entries omitted for brevity. The results for the median swap ($q = 0.5$) are shown in bold face. The second column presents the mean monthly gross returns, calculated from 299 observations over the sample period from January 1996 to December 2020. The corresponding t-statistics, as given in the third column, are adjusted for serial dependence according to Newey and West (1987). The last column presents the quantile swaps' mean model-free deltas.

2.4.2 Dispersion Swap Premiums

Table 2.2 provides evidence on dispersion premiums by presenting mean gross returns of different dispersion swaps with varying quantile specifications. Mean gross returns are generally negative with only two exceptions. There is also some tendency that longer distances between q_2 and q_1 yield higher premiums. A longer distance refers to a more extreme definition of dispersion, covering the tail areas of the distribution. For example, the tail dispersion swap ($q_2 = 0.95$ and $q_1 = 0.05$) has an average return of -3.65 cents per dollar of notional value, whereas the canonical dispersion swap ($q_2 = 0.75$ and $q_1 = 0.25$) loses on average only about 0.84 cents. However, even the tail dispersion swap's premium is considerably smaller than the location premium, as measured by the median swap. Dispersion premiums are only statistically significant ($\alpha = 5\%$) for the four most extreme swap specifications. Thus, the results do suggest the existence of a negative dispersion risk premium, yet it seems that only extreme dispersion (a tail effect) receives compensation. Such a finding is well in line with evidence suggesting that the variance risk premium is mainly a compensation for jump tail risk (Bollerslev and Todorov 2011).

The results for up- and downside dispersion swaps show that such an interpretation may not provide the full picture. Mean gross returns of downside dispersion swaps, which are given in the fifth column of Table 2.2, are all negative and there is no tendency that more extreme quantiles (low values of q_1) go along with higher premiums. For example, the tail downside dispersion swap ($q_1 = 0.05$) loses about 2 cents on average, whereas the canonical downside dispersion swap ($q_1 = 0.25$) loses about 3.2 cents. Both of these premiums are statistically significant. For the left part of the distribution, the dispersion premium is not just a tail phenomenon, which is a new result of our study. Mean returns of upside dispersion swaps, as given in the seventh column of Table 2.2, show very different behavior. The six most extreme upside dispersion contracts (highest values of q_2) have mean returns between -0.84 and -1.64 cents. Although only the most extreme one has a premium that is marginally statistically significant, in terms of magnitude these mean returns contribute substantially to the total premium of dispersion swaps (third column). Upside dispersion swaps for less extreme quantiles either have positive or negligible mean returns and do not contribute to the generally negative overall premiums of dispersion swaps. These results on upside dispersion swaps correspond with the findings by Bakshi et al. (2010) for other claims on the upside and the findings by Cuesdeanu and Jackwerth (2018) on a U-shaped pricing kernel. In general, our results are in line with a rationale based on the distinction between good and bad dispersion (Kilic and Shaliastovich 2019). Dispersion below the median is seen as bad dispersion and investors not only seek protection against it in the tails, but are also willing to pay a premium for such protection. Further, dispersion above the median is rather seen as good dispersion. Investors seek the chance for very high returns in the upper tail of the distribution due to lottery preferences[17] and are willing to pay a corresponding premium. When combining these two effects, one obtains an overall dispersion premium that appears to be a tail premium only.

Table 2.2: Dispersion Swap Premiums

q_1	q_2	Mean GR	t-statistic	Downside Mean GR	t-statistic	Upside Mean GR	t-statistic	Mean Delta
0.025	0.975	−0.0299	−3.8610	−0.0135	−3.6006	−0.0164	−2.1592	0.0005
0.050	**0.950**	**−0.0365**	**−2.8591**	**−0.0204**	**−3.2106**	**−0.0161**	**−1.2226**	**0.0012**
0.075	0.925	−0.0363	−1.9016	−0.0239	−2.4806	−0.0124	−0.7114	0.0018
0.100	0.900	−0.0395	−1.7845	−0.0308	−3.0812	−0.0087	−0.4739	0.0022
0.125	0.875	−0.0360	−1.4395	−0.0276	−2.1364	−0.0084	−0.3995	0.0025
0.150	0.850	−0.0358	−1.3212	−0.0244	−1.6954	−0.0114	−0.5331	0.0027
0.175	0.825	−0.0222	−0.8272	−0.0279	−1.8371	0.0057	0.2554	0.0028
0.200	0.800	−0.0187	−0.6882	−0.0348	−2.3775	0.0161	0.7211	0.0028
0.225	0.775	−0.0186	−0.6634	−0.0383	−2.5813	0.0197	0.8523	0.0028
0.250	**0.750**	**−0.0084**	**−0.2969**	**−0.0318**	**−2.0788**	**0.0234**	**1.0168**	**0.0027**
0.275	0.725	0.0052	0.1867	−0.0353	−2.3563	0.0405	1.7637	0.0025
0.300	0.700	0.0120	0.4572	−0.0288	−2.0064	0.0408	2.0124	0.0023
0.325	0.675	−0.0079	−0.2847	−0.0256	−1.4908	0.0177	0.9134	0.0021
0.350	0.650	−0.0110	−0.4035	−0.0324	−1.8583	0.0214	1.1887	0.0019
0.375	0.625	−0.0142	−0.5367	−0.0326	−2.0389	0.0184	1.0625	0.0016
0.400	0.600	−0.0207	−0.8154	−0.0395	−2.4283	0.0187	1.1851	0.0013
0.425	0.575	−0.0273	−1.1691	−0.0363	−2.4002	0.0090	0.6055	0.0010
0.450	0.550	−0.0204	−1.0840	−0.0130	−1.1271	−0.0074	−0.4837	0.0006
0.475	0.525	−0.0169	−1.0699	−0.0099	−1.0526	−0.0070	−0.5974	0.0003

Note: This table provides evidence on the premiums of dispersion swaps with notional values of one on the S&P 500 index for varying quantile specifications. The first two columns show the swaps' quantile specifications (q_1 and q_2), ranging from 0.025 to 0.475 for q_1 and from 0.975 to 0.525 for q_2. The results for the canonical dispersion swap ($q_1 = 0.25$ and $q_2 = 0.75$) and the tail dispersion swap ($q_1 = 0.05$ and $q_2 = 0.95$) are shown in bold face. The third column presents the mean monthly gross returns, calculated from 299 observations over the sample period from January 1996 to December 2020. The corresponding t-statistics, as given in the fourth column, are adjusted for serial dependence according to Newey and West (1987). Columns five and six deliver the mean gross returns and t-statistics, respectively, of the downside dispersion swaps and columns seven and eight deliver the corresponding values for the upside dispersion swaps. The last column presents the dispersion swaps' mean model-free deltas.

2.4.3 Asymmetry Swap Premiums

First evidence on asymmetry premiums is given in Table 2.3, which presents the mean gross returns of asymmetry swaps for different quantile specifications. Mean gross returns are negative and statistically significant without exception. There is also a slight tendency of higher premiums if asymmetry is measured in the tails compared to asymmetry close to the center of the distribution. For example, the tail asymmetry swap ($q_1 = 0$, $q_2 = 0.25$, $q_3 = 0.75$, and $q_4 = 1$) has a mean

[17] Kumar (2009), Bali et al. (2011), Doran et al. (2011), and Bali et al. (2017) provide evidence on lottery preferences and their impact on prices.

return of about −22 cents per dollar of notional, whereas the canonical swap ($q_1 = 0.25$, $q_2 = 0.5$, $q_3 = 0.5$, and $q_4 = 0.75$) has only a return of about −11 cents.

Looking at up- and downside asymmetry swaps sheds light on the components that make up the total asymmetry swap premium. In contrast to the decomposed dispersion swap, both downside and upside asymmetry swap premiums contribute significantly to the total premium, with the highest average gross returns for tail contracts and the lowest for contracts specified with quantiles around the median. Interestingly, in most cases, the upside premiums take even higher values than the downside premiums, suggesting that the upside portion of the return distribution is also relevant for investors.

One interpretation of our findings is that Table 2.3 provides evidence for a strong compensation of asymmetry risk, especially in the tails of the distribution. The existence of such an asymmetry premium would be in line with previous results on skewness risk premiums, as in Bali and Murray (2013) and Kozhan et al. (2013). That said, the relatively large negative deltas of the asymmetry swaps warn us to treat this interpretation with caution. If asymmetry swaps have significant market risk, the asymmetry premium may be (at least partly) just a market risk premium. Section 2.5 revisits this point.

Table 2.3: Asymmetry Swap Premiums

q_1	q_2	q_3	q_4	Mean GR	t-statistic	Downside Mean GR	t-statistic	Upside Mean GR	t-statistic	Mean Delta
0.000	**0.250**	**0.750**	**1.000**	**−0.2174**	**−5.6824**	**−0.1129**	**−5.7840**	**−0.1045**	**−3.8186**	**−0.0108**
0.025	0.275	0.725	0.975	−0.2408	−5.6306	−0.1028	−5.0533	−0.1380	−4.7612	−0.0102
0.050	0.300	0.700	0.950	−0.2274	−5.6423	−0.0895	−5.2362	−0.1380	−4.7878	−0.0092
0.075	0.325	0.675	0.925	−0.1940	−4.7156	−0.0828	−3.8229	−0.1112	−3.8242	−0.0083
0.100	0.350	0.650	0.900	−0.1940	−5.4397	−0.0828	−3.9422	−0.1112	−4.1328	−0.0074
0.125	0.375	0.625	0.875	−0.1940	−5.4885	−0.0861	−4.0055	−0.1079	−3.9088	−0.0066
0.150	0.400	0.600	0.850	−0.2074	−5.6629	−0.0962	−4.4011	−0.1112	−4.4447	−0.0058
0.175	0.425	0.575	0.825	−0.1739	−4.7011	−0.0895	−4.3439	−0.0844	−3.1338	−0.0051
0.200	0.450	0.550	0.800	−0.1171	−3.3191	−0.0594	−2.7731	−0.0577	−2.3384	−0.0044
0.225	0.475	0.525	0.775	−0.1070	−2.6431	−0.0527	−2.6108	−0.0543	−2.1239	−0.0038
0.250	**0.500**	**0.500**	**0.750**	**−0.1070**	**−2.6572**	**−0.0493**	**−2.1299**	**−0.0577**	**−2.2403**	**−0.0031**

Note: This table provides evidence on the premiums of asymmetry swaps with notional values of one on the S&P 500 index for varying quantile specifications. The first four column entries report the swaps' quantile specifications q_1, q_2, q_3, and q_4. The results for the canonical asymmetry swap ($q_1 = 0.25$, $q_2 = 0.5$, $q_3 = 0.5$, and $q_4 = 0.75$) and the tail asymmetry swap ($q_1 = 0$, $q_2 = 0.25$, $q_3 = 0.75$, and $q_4 = 1$) are shown in bold face. The fifth column presents the mean monthly gross returns, calculated from 299 observations over the sample period from January 1996 to December 2020. The corresponding t-statistics, as given in the sixth column, are adjusted for serial dependence according to Newey and West (1987). Columns seven and eight deliver the mean gross returns and t-statistics, respectively, of the downside asymmetry swaps and columns nine and ten deliver the corresponding values for the upside asymmetry swaps. The last column presents the asymmetry swaps' mean model-free deltas.

2.4.4 Steepness Swap Premiums

Results for various steepness contracts are given in Table 2.4. Steepness swaps have generally negative average returns, which are similar but smaller in magnitude than those of dispersion swaps. However, none of these average returns is statistically significant at the 5% level and our results do not suggest the existence of a steepness risk premium. The same holds true for up- and downside steepness premiums. Even if we decompose the steepness premiums into an up- and downside part, there is only one single significant component. Our findings on steepness premiums are in contrast to previous results in the literature on kurtosis premiums. Notably, Chang et al. (2013) and Harris and Qiao (2018) quantify a fourth moment risk premium; however, their realized kurtosis measures have the disadvantage of not fulfilling the aggregation property of Neuberger (2012). Recently, Khrashchevskyi (2020) introduced a kurtosis measure with this property, analogous to the skewness measure applied by Kozhan et al. (2013) to quantify skewness premiums. Essentially, he finds a significantly negative kurtosis risk premium that is strongly correlated with variance and skewness risk premiums. The apparent differences between the kurtosis premium by Khrashchevskyi (2020) and the mean returns of our steepness swaps are further investigated in Section 2.6.

Table 2.4: Steepness Swap Premiums

q_1	q_2	q_3	q_4	Mean GR	t-statistic	Downside Mean GR	t-statistic	Upside Mean GR	t-statistic	Mean Delta
0.025	**0.475**	**0.525**	**0.975**	**−0.0130**	**−0.7118**	**−0.0037**	**−0.3782**	**−0.0094**	**−0.6830**	**0.0001**
0.050	0.450	0.550	0.950	−0.0161	−0.7136	−0.0074	−0.5905	−0.0087	−0.4807	0.0006
0.075	0.425	0.575	0.925	−0.0090	−0.3335	0.0124	0.7718	−0.0214	−1.0087	0.0008
0.100	0.400	0.600	0.900	−0.0187	−0.6348	0.0087	0.4775	−0.0274	−1.3723	0.0009
0.125	**0.375**	**0.625**	**0.875**	**−0.0217**	**−0.6792**	**0.0050**	**0.2672**	**−0.0268**	**−1.1444**	**0.0009**
0.150	0.350	0.650	0.850	−0.0247	−0.7999	0.0080	0.4202	−0.0328	−1.6754	0.0008
0.175	0.325	0.675	0.825	−0.0144	−0.4892	−0.0023	−0.1462	−0.0120	−0.5302	0.0007
0.200	0.300	0.700	0.800	−0.0308	−1.2570	−0.0060	−0.5192	−0.0247	−1.2344	0.0005
0.225	0.275	0.725	0.775	−0.0237	−1.4853	−0.0030	−0.3759	−0.0207	−1.4045	0.0002

Note: This table provides evidence on the premiums of steepness swaps with notional values of one on the S&P 500 index for varying quantile specifications. The first four columns show the swaps' quantile specifications q_1, q_2, q_3, and q_4. The results for the canonical steepness swap ($q_1 = 0.125$, $q_2 = 0.375$, $q_3 = 0.625$, and $q_4 = 0.875$) and the tail steepness swap ($q_1 = 0.025$, $q_2 = 0.475$, $q_3 = 0.525$, and $q_4 = 0.975$) are shown in bold face. The fifth column presents the mean monthly gross returns, calculated from 299 observations over the sample period from January 1996 to December 2020. The corresponding t-statistics, as given in the sixth column, are adjusted for serial dependence according to Newey and West (1987). Columns seven and eight deliver the mean gross returns and t-statistics, respectively, of the downside steepness swaps and columns nine and ten deliver the corresponding values for the upside steepness swaps. The last column presents the steepness swaps' mean model-free deltas.

2.5 Delta-Hedged Moment Swap Premiums

As options are derivatives, options returns partly reflect premiums earned by the underlying. However, our main interest are premiums related to the non-linearity of the payoff functions. Dispersion swaps, asymmetry swaps, and steepness swaps do not fully disentangle such option-specific premiums from the market risk premium, as they have non-zero deltas. Therefore, we construct "zero-beta" versions of higher-moment swaps by delta-hedging the swaps with positions in the spot instrument. Hedges are set up at the beginning of each month using model-free deltas. Any outflows (inflows) from buying (selling) the index are financed (invested) at the risk-free interest rate. This hedging approach follows Goyal and Saretto (2009), who point out the conservatism of the buy-and-hold strategy compared to a frequently rebalanced strategy.

2.5.1 Delta-Hedged Dispersion Swap Premiums

Table 2.5 provides information on premiums of zero-beta dispersion swaps. Mean gross returns are always negative and larger in absolute terms than the respective returns of dispersion swaps, as shown in Table 2.2. This is due to the positive deltas of dispersion swaps, leading to delta hedges that take short positions in the index and deliver negative returns on average. Higher t-statistics, as compared to Table 2.2, also reflect this effect, and a wider range of delta-hedged dispersion swaps shows statistically significant premiums. Most importantly, zero-beta dispersion swaps confirm the main findings on dispersion risk premiums: that is, we still find negative overall premiums that are significant if the quantile specification moves toward the tails of the distribution. When looking at zero-beta downside and upside dispersion premiums[18] separately, we confirm that downside dispersion swaps have significant negative premiums for a wide range of different quantiles, whereas upside dispersion swaps contribute significant negative premiums only for high quantiles.

2.5.2 Delta-Hedged Asymmetry Swap Premiums

Results for zero-beta asymmetry swaps are reported in Table 2.6. The magnitude of mean gross returns is substantially reduced compared to the mean gross returns of the asymmetry swaps, as shown in Table 2.3. Due to the negative deltas of asymmetry swaps, the S&P 500 index enters the zero-beta contract with a long position, leading to less negative gross returns. Moreover, as deltas monotonically increase with the distance $q_2 - q_3$, delta hedges reduce average premiums more

[18] Downside and upside swaps are delta-hedged individually; that is, every delta-hedged down- and upside swap has an initial delta of zero. Given that the delta of a portfolio equals the sum of the deltas of the portfolio components, the overall premium of zero-beta swaps still equals the sum of the downside and upside premiums.

Table 2.5: Premiums of Delta-Hedged Dispersion Swaps

q_1	q_2	Mean GR	t-statistic	Downside Mean GR	t-statistic	Upside Mean GR	t-statistic
0.025	0.975	−0.0343	−4.2521	−0.0150	−3.7763	−0.0193	−2.4967
0.050	**0.950**	**−0.0418**	**−3.0631**	**−0.0188**	**−2.9428**	**−0.0230**	**−1.7390**
0.075	0.925	−0.0471	−2.3969	−0.0233	−2.4298	−0.0238	−1.3864
0.100	0.900	−0.0543	−2.3615	−0.0288	−2.8767	−0.0255	−1.3611
0.125	0.875	−0.0534	−2.0825	−0.0249	−1.9763	−0.0286	−1.4268
0.150	0.850	−0.0559	−2.0367	−0.0219	−1.5454	−0.0340	−1.6345
0.175	0.825	−0.0419	−1.6680	−0.0240	−1.6849	−0.0179	−0.8664
0.200	0.800	−0.0402	−1.5773	−0.0312	−2.2622	−0.0090	−0.4344
0.225	0.775	−0.0406	−1.5729	−0.0347	−2.5208	−0.0058	−0.2719
0.250	**0.750**	**−0.0295**	**−1.1324**	**−0.0278**	**−1.9426**	**−0.0016**	**−0.0726**
0.275	0.725	−0.0161	−0.5927	−0.0315	−2.2003	0.0154	0.7092
0.300	0.700	−0.0065	−0.2550	−0.0247	−1.7404	0.0183	0.9758
0.325	0.675	−0.0262	−0.9555	−0.0222	−1.3111	−0.0040	−0.2187
0.350	0.650	−0.0271	−1.0116	−0.0283	−1.6202	0.0012	0.0719
0.375	0.625	−0.0273	−1.0342	−0.0290	−1.8070	0.0017	0.0976
0.400	0.600	−0.0335	−1.3148	−0.0360	−2.1976	0.0025	0.1543
0.425	0.575	−0.0366	−1.5302	−0.0322	−2.1077	−0.0044	−0.2750
0.450	0.550	−0.0296	−1.5241	−0.0122	−1.0485	−0.0174	−1.1889
0.475	0.525	−0.0223	−1.3693	−0.0099	−0.9989	−0.0124	−1.0350

Note: This table provides evidence on the premiums of delta-hedge dispersion swaps with notional values of one on the S&P 500 index for varying quantile specifications. Delta-hedges use appropriate short positions (according to the model-free deltas of the swaps) in the index. Proceeds from these short positions are invested at the risk-free rate. The first two columns of the table show the swaps' quantile specifications (q_1 and q_2), ranging from 0.025 to 0.475 for q_1 and from 0.975 to 0.525 for q_2. Results for the delta-hedged canonical dispersion swap ($q_1 = 0.25$ and $q_2 = 0.75$) and the delta-hedged tail dispersion swap ($q_1 = 0.05$ and $q_2 = 0.95$) are shown in bold face. The third column presents the mean monthly gross returns, calculated from 299 observations over the sample period from January 1996 to December 2020. The corresponding t-statistics, as given in the fourth column, are adjusted for serial dependence according to Newey and West (1987). Columns five and six deliver the mean gross returns and t-statistics, respectively, of the delta-hedged downside dispersion swaps and columns seven and eight deliver the corresponding values for the delta-hedged upside dispersion swaps.

severely for the corresponding contracts. This likewise holds for up- and downside asymmetry swap mean gross returns. Nevertheless, mean gross returns are still negative and statistically significant for all contracts and the overall picture from Section 2.4.3 is confirmed: there is a significant premium of asymmetry swaps that is more pronounced when asymmetry is measured in the tails of the distribution compared to the center. Both up- and downside asymmetry swap components contribute to these premiums, with the upside premium accounting for the larger share to the total premium.

Table 2.6: Premiums of Delta-Hedged Asymmetry Swaps

q_1	q_2	q_3	q_4	Mean GR	t-statistic	Downside Mean GR	t-statistic	Upside Mean GR	t-statistic
0.000	**0.250**	**0.750**	**1.000**	**−0.1392**	**−7.3695**	**−0.0843**	**−7.5144**	**−0.0548**	**−2.7057**
0.025	0.275	0.725	0.975	−0.1641	−6.4754	−0.0729	−5.5893	−0.0912	−4.1555
0.050	0.300	0.700	0.950	−0.1601	−5.8309	−0.0624	−4.9408	−0.0977	−4.0612
0.075	0.325	0.675	0.925	−0.1316	−3.9095	−0.0553	−3.0666	−0.0763	−2.8711
0.100	0.350	0.650	0.900	−0.1392	−4.3870	−0.0560	−3.0287	−0.0832	−3.2335
0.125	0.375	0.625	0.875	−0.1474	−4.2425	−0.0606	−2.9747	−0.0868	−3.3318
0.150	0.400	0.600	0.850	−0.1635	−5.0212	−0.0705	−3.2954	−0.0929	−3.6473
0.175	0.425	0.575	0.825	−0.1347	−3.6883	−0.0647	−3.2061	−0.0700	−2.5848
0.200	0.450	0.550	0.800	−0.0856	−2.5506	−0.0375	−1.9603	−0.0481	−1.9733
0.225	0.475	0.525	0.775	−0.0815	−2.1075	−0.0316	−1.5279	−0.0499	−1.9545
0.250	**0.500**	**0.500**	**0.750**	**−0.0867**	**−2.2359**	**−0.0286**	**−1.3237**	**−0.0581**	**−2.2429**

Note: This table provides evidence on the premiums of delta-hedged asymmetry swaps with notional values of one on the S&P 500 index for varying quantile specifications. Delta-hedges use appropriate long-positions in the index (according to the model-free deltas of the swaps), which are financed via risk-free debt. The first four columns of the table report the swaps' quantile specifications q_1, q_2, q_3, and q_4. Results for the delta-hedged canonical asymmetry swap ($q_1 = 0.25$, $q_2 = 0.5$, $q_3 = 0.5$, and $q_4 = 0.75$) and the delta-hedged tail asymmetry swap ($q_1 = 0$, $q_2 = 0.25$, $q_3 = 0.75$, and $q_4 = 1$) are shown in bold face. The fifth column presents the mean monthly gross returns, calculated from 299 observations over the sample period from January 1996 to December 2020. The corresponding t-statistics, as given in the sixth column, are adjusted for serial dependence according to Newey and West (1987). Columns seven and eight deliver the mean gross returns and t-statistics, respectively, of the delta-hedged downside asymmetry swaps and columns nine and ten deliver the corresponding values for the delta-hedged upside asymmetry swaps.

2.5.3 Delta-Hedged Steepness Swap Premiums

Zero-beta steepness swaps yield consistently negative mean gross returns that are very similar to the mean gross returns of the base-case steepness swaps. Table 2.7 shows these results. The close similarity to the base case from Table 2.4 is due to the relatively small deltas of steepness swaps. There is still only one statistically significant mean return and thus no robust evidence for any systematic steepness premium, even if we consider a decomposition of the overall premium into zero-beta up- and downside premiums.

Table 2.7: Premiums of Delta-Hedged Steepness Swaps

q_1	q_2	q_3	q_4	Mean GR	t-statistic	Downside Mean GR	t-statistic	Upside Mean GR	t-statistic
0.025	**0.475**	**0.525**	**0.975**	**−0.0121**	**−0.6442**	**−0.0052**	**−0.5055**	**−0.0069**	**−0.4974**
0.050	0.450	0.550	0.950	−0.0122	−0.5353	−0.0066	−0.5141	−0.0056	−0.3079
0.075	0.425	0.575	0.925	−0.0105	−0.3857	0.0089	0.5512	−0.0194	−0.9085
0.100	0.400	0.600	0.900	−0.0208	−0.6877	0.0072	0.3878	−0.0280	−1.3524
0.125	**0.375**	**0.625**	**0.875**	**−0.0262**	**−0.8068**	**0.0041**	**0.2170**	**−0.0303**	**−1.2983**
0.150	0.350	0.650	0.850	−0.0287	−0.9308	0.0064	0.3336	−0.0351	−1.8342
0.175	0.325	0.675	0.825	−0.0157	−0.5275	−0.0018	−0.1120	−0.0139	−0.6178
0.200	0.300	0.700	0.800	−0.0337	−1.3868	−0.0064	−0.5540	−0.0273	−1.3259
0.225	0.275	0.725	0.775	−0.0244	−1.5187	−0.0032	−0.4003	−0.0212	−1.4069

Note: This table provides evidence on the premiums of delta-hedged steepness swaps with notional values of one on the S&P 500 index for varying quantile specifications. The first fours columns show the swaps' quantile specifications q_1, q_2, q_3, and q_4. Results for the delta-hedged canonical steepness swap ($q_1 = 0.125$, $q_2 = 0.375$, $q_3 = 0.625$, and $q_4 = 0.875$) and the delta-hedged tail steepness swap ($q_1 = 0.025$, $q_2 = 0.475$, $q_3 = 0.525$, and $q_4 = 0.975$) are shown in bold face. The fifth column presents the mean monthly gross returns, calculated from 299 observations over the sample period from January 1996 to December 2020. The corresponding t-statistics, as given in the sixth column, are adjusted for serial dependence according to Newey and West (1987). Columns seven and eight deliver the mean gross returns and t-statistics, respectively, of the delta-hedged downside steepness swaps and columns nine and ten deliver the corresponding values for the delta-hedged upside steepness swaps.

2.6 Quantile-Based and Traditional Moment Swaps

Quantile-based and traditional moment swaps follow a similar idea, yet they build on different concepts. While the quantile-based approach has its roots in robust statistics, traditional moments—higher moments in particular—are known to be non-robust with respect to outliers (Kim and White 2004). A better understanding of the potential effects of this difference on risk premiums is the general goal of this section. A more specific goal is to shed light on the results that seem to contradict findings on traditional moment premiums: The quantile-based approach identifies significant quantile-based moment premiums both for second and third moments but not for the fourth moment. However, the evidence on traditional moment risk premiums suggests the joint existence of variance, skewness, and kurtosis risk premiums (Kozhan et al. 2013; Khrashchevskyi 2020). Another goal is to understand the factors driving quantile-based moment risk premiums. Traditional premiums can be traced back to one common factor (Kozhan et al. 2013; Khrashchevskyi 2020), which is likely to be a factor associated with large downside jumps or tail events, respectively (Hain et al. 2018). This section aims to answer whether quantile-based higher-moment swaps' payoffs are driven by the same common factor as traditional moment premiums.

To start with the comparison, Table 2.8 shows summary statistics of quantile-based and traditional moment swap returns for first to fourth moments. As representatives of quantile-based moment swaps, we use delta-hedged canonical swaps and delta-hedged tail swaps. Traditional moment swaps are the mean swap (short forward), the variance and skewness swaps, as in Kozhan et al. (2013), and the kurtosis swap, as in Khrashchevskyi (2020). As expected, mean gross returns are negative for all contracts. The most striking difference between quantile-based and traditional moment swaps is indeed the non-significant steepness premium, as compared to the significant kurtosis premium. Additionally, we observe that unlike the variance swap and the tail dispersion swap, the canonical dispersion swap does not have a significantly negative mean gross return. Traditional moment swaps also show a specific pattern in the magnitudes of premiums (average returns) and return standard deviations. The higher the moment, the lower (less negative) the premium and the lower the standard deviation. This is not the case for quantile-based moment swaps because asymmetry swap premiums are higher (more negative) than dispersion swap premiums. Asymmetry swaps also have a higher return standard deviation and a higher (more negative) Sharpe ratio.

Table 2.8: Quantile-Based and Traditional Swaps: Summary Statistics

	Mean GR	St. Dev. GR	Sharpe Ratio
S^{Δ}_{can}	-0.1129^{***}	0.2993	-0.3773
DS^{Δ}_{can}	-0.0295	0.5095	-0.0579
AS^{Δ}_{can}	-0.0867^{**}	0.6911	-0.1255
SS^{Δ}_{can}	-0.0262	0.5067	-0.0516
DS^{Δ}_{tail}	-0.0418^{***}	0.2546	-0.1641
AS^{Δ}_{tail}	-0.1392^{***}	0.3345	-0.4160
SS^{Δ}_{tail}	-0.0121	0.2895	-0.0417
MRP×100	-0.5879^{**}	4.6670	-0.1260
VRP×100	-0.0768^{***}	0.4644	-0.1654
SRP×100	-0.0146^{***}	0.1330	-0.1098
KRP×100	-0.0079^{***}	0.0650	-0.1221

Note: This table shows mean gross returns, gross return standard deviations and Sharpe ratios for delta-hedged canonical moment swaps, delta-hedged tail moment swaps and traditional moment swaps. The canonical swaps have the following quantile specifications: Quantile Swap (S^{Δ}_{can}): $q = 0.5$, dispersion swap (DS^{Δ}_{can}): $q_1 = 0.25$, $q_2 = 0.75$, asymmetry swap (AS^{Δ}_{can}): $q_1 = 0.25$, $q_2 = q_3 = 0.5$, $q_4 = 0.75$, and steepness swap (SS^{Δ}_{can}): $q_1 = 0.125$, $q_2 = 0.375$, $q_3 = 0.625$, $q_4 = 0.875$. The tail swaps are defined via the following quantiles: dispersion swap (DS^{Δ}_{tail}): $q_1 = 0.05$, $q_2 = 0.95$, asymmetry swap (AS^{Δ}_{tail}): $q_1 = 0$, $q_2 = 0.25$, $q_3 = 0.75$, $q_4 = 1$, and steepness swap (SS^{Δ}_{tail}): $q_1 = 0.025$, $q_2 = 0.475$, $q_3 = 0.525$, $q_4 = 0.975$. Traditional moment swaps are denoted by MRP (mean swap (short forward)), VRP (variance swap), SRP (skewness swap), and KRP (kurtosis swap). Asterisks indicate significant negative mean gross returns at the 1%(***), 5%(**), and 10%(*) significance levels, respectively.

Return correlations between different swaps are shown in Table 2.9. For traditional moment swaps, we can corroborate the findings of Kozhan et al. (2013) and Khrashchevskyi (2020) that all premiums are strongly positively correlated. This is especially the case for the kurtosis premium and the skewness premium as well as the variance premium and the skewness premium, which show a very high correlation above 0.9.

Table 2.9: Quantile-Based and Traditional Swaps: Correlations

	DS^{Δ}_{can}	AS^{Δ}_{can}	SS^{Δ}_{can}	DS^{Δ}_{tail}	AS^{Δ}_{tail}	SS^{Δ}_{tail}	MRP	VRP	SRP	KRP
S^{Δ}_{can}	−0.0550	0.8751	0.0837	0.0544	−0.0185	0.0145	−0.0172	−0.1807	−0.2202	−0.1989
DS^{Δ}_{can}		0.1757	−0.0553	0.2982	−0.4614	−0.1243	0.1730	0.3039	0.2805	0.2725
AS^{Δ}_{can}			0.0685	0.0891	−0.5000	−0.0119	0.0018	−0.0281	−0.0367	−0.0271
SS^{Δ}_{can}				0.3070	0.0083	0.3283	0.1315	0.1539	0.1419	0.1308
DS^{Δ}_{tail}					−0.0866	0.2160	0.1857	0.4358	0.3829	0.3414
AS^{Δ}_{tail}						0.0504	−0.0345	−0.2654	−0.3181	−0.3000
SS^{Δ}_{tail}							0.0208	0.1681	0.2055	0.2001
MRP								0.6073	0.5609	0.5149
VRP									0.9450	0.8988
SRP										0.9734

Note: This table shows return correlations between delte-hedged canonical moment swaps, delta-hedged tail moment swaps and traditional moment swaps. The canonical swaps have the following quantile specifications: Quantile Swap (S^{Δ}_{can}): $q = 0.5$, dispersion swap (DS^{Δ}_{can}): $q_1 = 0.25$, $q_2 = 0.75$, asymmetry swap (AS^{Δ}_{can}): $q_1 = 0.25$, $q_2 = q_3 = 0.5$, $q_4 = 0.75$, and steepness swap (SS^{Δ}_{can}): $q_1 = 0.125$, $q_2 = 0.375$, $q_3 = 0.625$, $q_4 = 0.875$. The tail swaps are defined via the following quantiles: dispersion swap (DS^{Δ}_{tail}): $q_1 = 0.05$, $q_2 = 0.95$, asymmetry swap (AS^{Δ}_{tail}): $q_1 = 0$, $q_2 = 0.25$, $q_3 = 0.75$, $q_4 = 1$, and steepness swap (SS^{Δ}_{tail}): $q_1 = 0.025$, $q_2 = 0.475$, $q_3 = 0.525$, $q_4 = 0.975$. Traditional moment swaps are denoted by MRP (mean swap (short forward)), VRP (variance swap), SRP (skewness swap), and KRP (kurtosis swap).

Compared to correlations between traditional moment swaps, correlations between quantile-based canonical swaps are weaker and often insignificant. Only S^{Δ}_{can} and AS^{Δ}_{can} have a very high positive correlation, suggesting that they measure similar aspects of the discrepancy between risk-neutral and physical distributions around the center. We also observe rather low correlations between different quantile-based tail swaps.

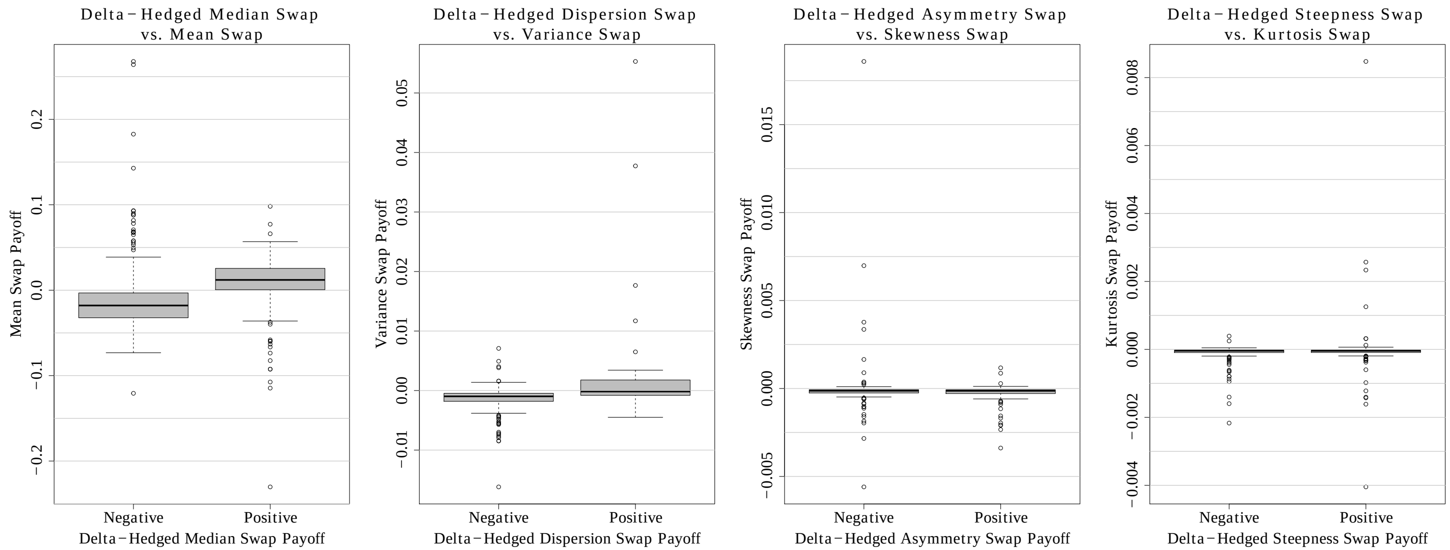
Looking at the return correlations between canonical swaps and tail swaps, we see that the tail variant always shows a significant correlation with the canonical variant of the respective moment. However, correlations are far from perfect. Moreover, AS^{Δ}_{can} and AS^{Δ}_{tail} even have a negative correlation, suggesting that the two underlying concepts of asymmetry around the center versus asymmetry in the tails are economically very different, and realized premiums materialize at different times. Overall, correlations suggest that moment premiums referring to the tails of the distribution have different sources than moment premiums referring to the center.[19]

[19] We can confirm this notion when regressing quantile-based premiums on each other. Such an analysis, which is similar to the analysis in Kozhan et al. (2013), identifies two distinct quantile-based higher-moment risk premiums; a premium related to dispersion and a premium related to asymmetry. This finding likewise confirms the results of Sections 2.4 and 2.5. The results are unreported but available upon request.

We now take a closer look at the return distributions of different moment swaps, paying particular attention to extreme returns or outliers. Figure 2.3 gives a visual impression of the return relation between quantile-based delta-hedged tail swaps and their traditional counterparts. As quantile-based moment swaps have only two (dispersion swaps and steepness swaps) or three (asymmetry swaps) possible payoffs, we just distinguish between positive and negative returns.[20] For the two subgroups of positive and negative returns of quantile-based moment swaps, the figure shows boxplots of the corresponding conditional return distributions of traditional moment swaps.

[20]The delta-hedged versions of the swaps can take a continuum of values. However, the effect of the hedge position in the underlying is quite small.

Figure 2.3: Quantile-Based Versus Traditional Moment Swap Premiums



No. of Observations	209	90	263	36	183	116	267	32
Mean×100	−0.9589	0.2738	−0.1322	0.3278	−0.0060	−0.0282	−0.0110	0.0179
Median×100	−1.7911	1.2047	−0.0972	−0.0161	−0.0130	−0.0134	−0.0042	−0.0021
χ^2	80.8398		25.8975		0.0000		1.0363	
p-value	0.0000		0.0000		1.0000		0.3087	
No. of Outliers	23	17	25	6	27	16	42	3

Note: This figure provides further information on the relation between quantile-based moment swap premiums and traditional moment swap premiums. The x-axes depict the delta-hedged tail swaps' payoffs, categorized by their signs. The y-axes show the payoff distributions of the corresponding traditional moment swaps (mean swap (short forward), variance swap, skewness swap, and kurtosis swap) in terms of boxplots. The boxes depict the interquartile range of the data and whiskers are set at 1.5 times the interquartile range above the upper quartile (upper whisker) and 1.5 times the interquartile range below the lower quartile (lower whisker). Dots indicate outliers that are outside of this range. The table below the boxplots shows the number of observations, the mean and the median of the respective subsample as well as test statistics of chi-square tests. The tests checks whether the signs of payoffs of the two instruments (quantile-based and traditional) are independent. Finally, the table provides the number of outliers for each boxplot.

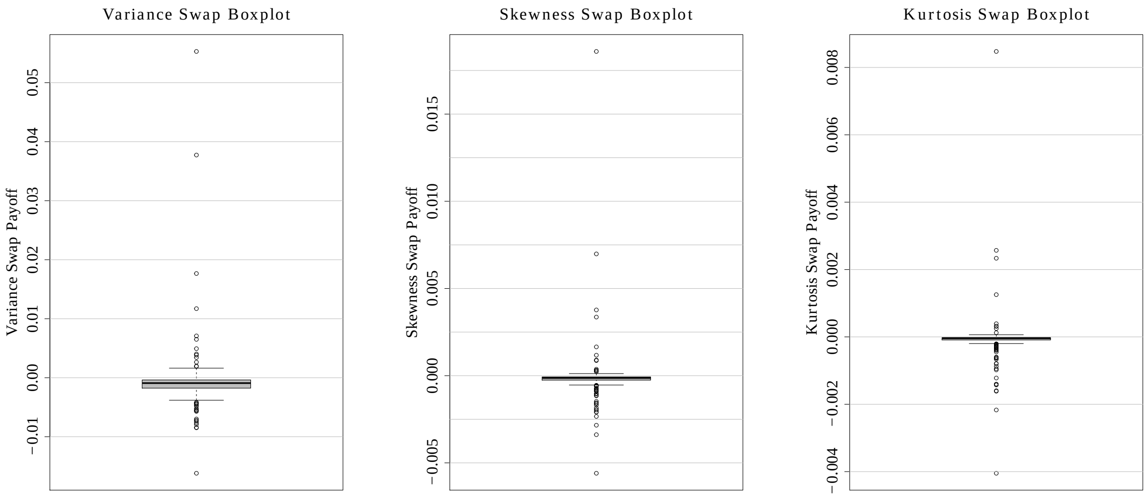
The boxplots illustrate a positive association between quantile-based and traditional moment swaps' returns for first and second moments. Realized returns of the mean swap are higher in months where the median swap has a positive payoff. This is emphasized by a mean return of -0.9589 if the payoffs of median swaps are negative and a mean return of 0.2738 if they are positive. Qualitatively similar findings are obtained for negative and positive payoffs of $\mathrm{DS}^{\Delta}_{tail}$, with mean returns of variance swaps of -0.1322 and 0.3278, respectively. Such an association is not visible for third and fourth moment swaps. The boxes corresponding to negative and positive returns of quantile-based swaps are very similar and there is no indication of one subsample being substantially different from the other. Tests for independence confirm the visual impression. They reject the hypothesis of independence between returns of quantile-based and traditional moment swaps for first and second moments but not for third and fourth moments.

Another observation from Figure 2.3 is the high number of outliers (returns above the upper or below the lower whisker) of traditional moment swaps. For a better understanding of the relation between variance, skewness and kurtosis swaps, it is interesting to observe whether such outliers occur at the same times. If they do, tail extremity of the return distribution would be an important commonality of higher-moment swaps. Figure 2.4 provides evidence on this issue.

The figure shows boxplots of the unconditional return distributions of traditional higher-moment swaps. For a total of 299 monthly returns, we observe 47, 55, and 56 outliers of the variance swap, the skewness swap, and the kurtosis swap, respectively. The tables below the boxplots indicate how often outliers of two different swaps occur in the same month. For example, 38 of the 47 extreme returns (outliers) of a variance swap occur in a month where the skewness swap also has an extreme return, and 49 of the 55 extreme returns of the skewness swap are observed in months with an extreme return of the kurtosis swap. There is clear evidence of a strong commonality of extreme returns for the three traditional higher-moment swaps. This property offers a possible explanation for why higher-moment premiums based on traditional moment swaps differ in important ways from quantile-based higher-moment premiums. If traditional higher-moment swap returns are mainly driven by the same extreme events, there is likely only one dominating higher-moment factor that may be called a tail extremity factor.[21] In contrast,

[21] Empirical evidence, for example from Bollerslev and Todorov (2011) and Neumann et al. (2016), supports this notion. These papers find that risk compensation for extreme events determines a large part of the variance risk premium.

Figure 2.4: Traditional Moment Swap Premiums: Outlier Analysis



	SRP No Outl.	SRP Outl.	Total
VRP No Outl.	235	17	252
VRP Outl.	9	38	47
χ^2	140.02		
p-value	2.635303e-32		

	KRP No Outl.	KRP Outl.	Total
SRP No Outl.	237	7	244
SRP Outl.	6	49	55
χ^2	213.585		
p-value	2.268614e-48		

	VRP No Outl.	VRP Outl.	Total
KRP No Outl.	233	10	243
KRP Outl.	19	37	56
χ^2	127.232		
p-value	1.652745e-29		

Note: This figure provides information on the return distribution of traditional moment swaps. It shows boxplots of the realized returns of the variance swap, the skewness swap, and the kurtosis swap. The boxes depict the interquartile range of the data and whiskers are set at 1.5 times the interquartile range above the upper quartile (upper whisker) and 1.5 times the interquartile range below the lower quartile (lower whisker). Dots indicate outliers that are outside of this range. The table below the boxplots gives information on the simultaneous occurrence of outliers (observations above the upper whisker and below the lower whisker) and non-outliers for different moment swaps. Results of chi-square tests are also provided. The null hypothesis of these tests is that the occurrence of an outlier in one moment swap is independent from the occurrence of an outlier in another moment swap.

the quantile-based approach works explicitly with specific partitions of the distribution, which, at least for third and fourth moments, also include areas closer to the center of the distribution.[22]

To assess whether a "tail extremity" factor can explain premiums of quantile-based contracts, we run a linear regression in which we regress the payoffs of delta-hedged quantile-based canonical and tail swaps on payoffs of traditional swaps. The regression takes the following form:

$$QS_t^{\Delta} = \alpha + \beta_{MRP} \cdot MRP_t + \beta_{VRP} \cdot VRP_t + \beta_{SRP} \cdot SRP_t + \beta_{KRP} \cdot KRP_t + \epsilon_t, \tag{2.6}$$

where QS_t^{Δ} denotes the payoff of a delta-hedged quantile-based moment swap for return period t. If quantile-based moment premiums just compensate for tail extremity, we expect any unexplained returns of quantile-based moment swaps to disappear, as indicated by insignificant regression alphas. In contrast, if there is an additional factor that determines the average payoff of quantile-based moment swaps, we expect the regression alpha to be statistically significant and negative. Table 2.10 provides the regression results.

Canonical delta-hedged dispersion and steepness swaps do not exhibit significant alphas. As these swaps do not show significantly negative average payoffs in the main analysis, we would not expect significant alphas. However, the alphas have decreased compared to the average payoffs in Section 2.5. For the delta-hedged canonical dispersion swap this could be due to the significant loading of the variance risk premium. In contrast, the average canonical delta-hedged asymmetry swap premium cannot be explained by traditional moment risk premiums as there are no significant loadings and an adjusted R^2 of -0.0090. Turning to the tail contracts, we observe that the variance risk premium and the kurtosis risk premium significantly load on the delta-hedged tail dispersion contract, which has no significant alpha anymore. Thus, the significant average tail dispersion premium can be explained by the VRP and the KRP with an adjusted R^2 of 0.2052. After controlling for traditional moment risk premiums the delta-hedged tail asymmetry swap has a remaining significant alpha. Interestingly, all of the loadings are statistically significant and the adjusted R^2 amounts to 0.1315; however, the alpha is still significantly different from zero. Lastly, we do not see a significantly negative alpha for the delta-hedged tail steepness swap.

Thus, Table 2.10 provides evidence that quantile-based moment risk premiums reveal an additional source of risk that is priced in index options and goes beyond tail extremity. As

[22] Even the tail asymmetry swap compares the first quartile with the fourth quartile of the distribution and considers 50% of the probability mass of the risk-neutral distribution.

Table 2.10: Sources of Robust Higher-Moment Risk Premiums

	DS^{Δ}_{can}	AS^{Δ}_{can}	SS^{Δ}_{can}	DS^{Δ}_{tail}	AS^{Δ}_{tail}	SS^{Δ}_{tail}
α	−0.0005	−0.0823**	−0.0126	−0.0237	−0.1389***	−0.0118
	(0.0270)	(0.0403)	(0.0314)	(0.0145)	(0.0185)	(0.0167)
β_{MRP}	−0.1770	0.4951	0.6356	−0.7686	1.4224***	−0.7671*
	(0.7640)	(0.7468)	(0.9062)	(0.7756)	(0.4979)	(0.3978)
β_{VRP}	44.2718***	12.0129	14.4208	39.2766***	19.2935**	−11.4560
	(12.6846)	(19.0636)	(16.8597)	(11.9753)	(8.8923)	(8.2448)
β_{SRP}	−88.4434	−178.2741	13.2259	31.5881	−272.3291***	128.7528***
	(68.2752)	(132.8487)	(89.2668)	(54.1692)	(56.8786)	(44.5047)
β_{KRP}	111.9652	230.6303	−40.4421	−152.8727**	211.3714**	−65.3491
	(106.7576)	(226.9116)	(145.3192)	(70.5154)	(82.0871)	(79.8872)
n	299	299	299	299	299	299
Adj. R^2	0.0817	−0.0090	0.0129	0.2052	0.1315	0.0459
F-Stat	7.6287	0.3326	1.9753	20.2306	12.2844	4.5823

Note: This table shows estimation results for a regression model where we regress payoffs of different quantile-based moment swaps on payoffs of traditional moment swaps according to the following regression:

$$QS^{\Delta}_t = \alpha + \beta_{MRP} \cdot MRP_t + \beta_{VRP} \cdot VRP_t + \beta_{SRP} \cdot SRP_t + \beta_{KRP} \cdot KRP_t + \epsilon_t,$$

where QS^{Δ}_t denotes the payoff of a delta-hedged quantile-based moment swap in time t. The sample period is January 1996 to December 2020. The different columns show estimation results for the canonical dispersion swap (DS^{Δ}_{can}), the canonical asymmetry swap (AS^{Δ}_{can}), the canonical steepness swap (SS^{Δ}_{can}) as well as for the tail dispersion swap (DS^{Δ}_{tail}), tail asymmetry swap (AS^{Δ}_{tail}), and tail steepness swap (SS^{Δ}_{tail}). Traditional moment swap payoffs are denoted by MRP (mean swap (short forward)), VRP (variance swap), SRP (skewness swap), and KRP (kurtosis swap). The table provides the regression alphas, beta coefficients, the corresponding t-statistics, and the regressions' adjusted R^2. Standard errors and covariances that enter t-statistics are adjusted for serial dependence according to Newey and West (1987). Asterisks indicate significant coefficients at the 1%(***), 5%(**), and 10%(*) level, respectively.

significantly negative alphas are only apparent for the asymmetry swap (irrespective of its quantile specifications), this distinct risk factor seems to be associated with the asymmetry in the return distribution. Nonetheless, some quantile-based premiums can be traced back to tail extremity. This is observable for the delta-hedged tail dispersion swap, in particular. In summary, due to their flexibility and robustness, quantile-based moment swaps are a means to identify compensations for both tail extremity and asymmetry in the return distribution. The latter premium is an additional option-specific one that traditional moment swaps are unable to capture.

2.7 Conclusion

This paper investigates quantile-based alternatives to traditional moment swaps, such as variance swaps and skewness swaps. It introduces the q-quantile swap as a basic building block and shows that portfolios of these swaps can be used to quantify risk premiums associated with any partition of the risk-neutral return distribution. In particular, one can choose partitions that correspond to well-known moment measures from robust statistics. In principle, this approach allows us to quantify quantile-based moment premiums for any higher moment.

In an empirical study for the S&P 500 options market, we investigate quantile-based moment premiums for the first to fourth moments. We find significant premiums for location, dispersion, and asymmetry, all of which have the expected signs; however, we do not find any significant steepness premium. The study also shows that it is important in distinguishing between the tails, the center, the upside and the downside parts of the return distribution. After controlling for the effects of traditional moment risk premiums, we are essentially left with another distinct option-specific premium related to the asymmetry in the return distribution. The quantile-based approach is thus able to identify two option-specific premiums—one referring to tail extremity and one referring to the asymmetry in the return distribution.

3 Decomposed Higher-Moment Risk Premiums and Market Return Predictability

Under review at Quantitative Finance.

Abstract

In an empirical study of Standard & Poor's 500 index options, this paper analyses the predictability of future market excess returns by means of decomposed higher-moment risk premiums. The study proposes a new measure of kurtosis risk premium and suggests a decomposition of higher-moment risk premiums up to the fourth moment into downside and upside premiums. Thereby, the paper enhances the understanding of higher-moment risk premiums. The decomposition uncovers valuable information for return forecasts, as decomposed higher-moment risk premiums deliver improved in-sample predictions. In an out-of-sample study, the predictive power of decomposed higher-moment risk premiums is shown to be particularly driven by downside higher-moment risk premiums.

Acknowledgements: I would like to thank Olaf Korn and Alexander Merz for helpful comments and suggestions.

3.1 Introduction

Moment risk premiums are capable of predicting future market excess returns. Bollerslev, Tauchen, and Zhou (2009) provide evidence that the variance risk premium is a reasonably good predictor for subsequent three- to six-month returns in the U.S. market with deteriorating predictive power for longer horizons. Kilic and Shaliastovich (2019) demonstrate that a decomposition of the variance risk premium into downside and upside premiums can improve predictions of market excess returns, particularly at longer horizons. Fan et al. (2020) investigate the complementary predictive power of higher-moment risk premiums and show that longer-term predictive power up to 24 months can be improved through the consideration of additional information embedded in the difference between risk-neutral and physical skewness and kurtosis, respectively.

This paper aims to synthesize these dimensions and analyze the predictive power of decomposed higher-moment risk premiums. As the literature shows that market excess return predictions can be improved by *decomposing* the variance risk premium and by considering *higher moments* than variance, the natural next step is to evaluate whether *decomposing higher-moment risk premiums* can likewise improve subsequent market excess return forecasts. In doing so, this paper enhances our understanding of moment risk premiums and provides novel insights into how and the extent to which they can be exploited as information signals for return predictions. This is particularly interesting and important given that accurate predictions are crucial for investment decisions and in the creation of value for investors.

Several methodological and empirical steps are necessary to conduct the analysis. First, it is important to ensure a consistent measurement of higher-moment risk premiums: I use ex-ante moment risk premiums constructed as profits from trading strategies based on the model developed by Kozhan et al. (2013). I extend their model with a kurtosis risk premium that has a lower approximation error in measuring r^4 than the kurtosis risk premium measure described by Khrashchevskyi (2020). The measures used in this paper have the Neuberger (2012) aggregation property and therefore facilitate a precise estimation of risk premiums that is independent of the data sampling frequency. Moreover, they can be interpreted as true economic risk premiums by virtue of their construction as profits from a trading strategy. Second, I suggest a decomposition of total higher-moment risk premiums that can be consistently applied to all higher-moment premiums. This can mitigate the loss of potentially relevant information ensuing from the offsetting effects of upside and downside premiums. Third, I analyze these premiums' ability to predict future market excess returns. To verify the plausibility of decomposed higher-moment risk premiums as stock return predictors, I additionally evaluate their capability to forecast macroeconomic variables that pose proxies for the business cycle. If return predictor variables are simultaneously able to forecast variations in the business cycle, their predictive power can be more plausibly related to time-varying risk premiums that can be traced back to changes in risk aversion (Cochrane 2007; Lin et al. 2018).

In an empirical study of the Standard & Poor's 500 (S&P 500) index option market, I investigate the characteristics of decomposed higher-moment risk premiums and their predictive power for subsequent market excess returns. The analyses yield the following main findings: (i) On average, downside premiums and total premiums are negative while upside premiums are positive but smaller in absolute value than downside premiums. The premiums have extreme distributions with large negative (positive) skewness for downside and total (upside) premiums as well as considerably high kurtosis. Moreover, the premiums are highly correlated. (ii) Decomposition of higher-moment risk premiums can increase the in-sample predictive power for market excess returns compared to the predictive power of the variance risk premium and the total moment risk premiums. This is particularly true for predictions horizons of six months or longer. (iii) At various prediction horizons, higher-moment risk premiums can predict changes in macroeconomic variables that serve as a proxy for the state of the economy. The decomposition of premiums increases the predictive power in the majority of considered cases. (iv) In an out-of-sample study, (decomposed) variance risk premiums yield the best predictions up to prediction horizons of six months. At forecasting horizons that are larger than six months, a combination forecast that comprises only downside higher-moment risk premiums is the best predictor for subsequent returns, which indicates that longer-term predictive power is driven by downside higher-moment risk premiums.

This paper is related to different strands of literature that study moment risk premiums. There is ample evidence for the existence of the variance risk premium (Carr and Wu 2009; Todorov 2010), the skewness risk premium (Kozhan, Neuberger, and Schneider 2013; Lin, Lehnert, and Wolff 2019), and the kurtosis risk premium (Khrashchevskyi 2020) for the S&P 500 index option market. Beyond that, Bollerslev and Todorov (2011) show that compensation for jump tail risk determines a significant portion of the variance risk premium. In addition, Held et al. (2020) and Londono and Xu (2020) investigate the characteristics of decomposed variance risk premiums in an international setting. They find that both upside and downside premiums are statistically significant but different from each other.

My paper contributes to this literature by offering extended evidence of higher-moment risk premiums and their decomposition. In particular, I introduce a new trading strategy for determining the kurtosis risk premium that involves a smaller approximation error than the established approach in the literature. Moreover, this paper is the first to study a decomposition of the skewness risk premium on the equity index market and the first to propose a decomposition of the kurtosis risk premium. Through its analysis of these premiums, this paper contributes to an enhanced understanding of the pricing of moment risks in the equity index market.

Along with premiums' sizes, several studies have also investigated their predictive power for future equity index returns. Bollerslev et al. (2009) demonstrate that the variance risk premium is a reasonably good predictor for subsequent three- to six-month returns in the U.S. market. While Bollerslev et al. (2014) extend this evidence to include international stock markets, Fan et al. (2020) investigate the complementary predictive power of higher-moment risk premiums. They show that longer-term predictive power up to 24 months can be improved by the additional consideration of skewness and kurtosis risk premiums. Furthermore, Bollerslev et al. (2015) provide evidence that the variance risk premium's predictive power for market returns can be enhanced through the inclusion of a jump tail risk component as a separate predictor variable. Moreover, Feunou et al. (2018) and Kilic and Shaliastovich (2019) study the effects of decomposing the variance risk premiums into downside and upside premiums for market excess return predictions. They find that decomposition of the variance risk premium can increase the predictability of equity index returns, particularly for longer horizons. Da Fonseca and Xu (2017) and Da Fonseca and Dawui (2021) investigate decomposed variance and skewness risk premiums in the crude oil market and currency markets, respectively. The former provide in-sample evidence on the short-term predictive power for crude oil returns and demonstrate that decomposition into downside and upside premiums increases predictive power at a short horizon.

I contribute to this strand of literature by analyzing the predictive power of decomposed higher-moment risk premiums up to the fourth moment. First, I show that the decomposition of total higher-moment risk premiums can improve the prediction of various macroeconomic variables such that the predictive power of these premiums can be more plausibly related to time-varying risk premiums caused by changes in risk or risk aversion. Second, I provide evidence that the decomposition increases the in-sample and out-of-sample predictive power, particularly at horizons of six months or longer. The out-of-sample predictive power in particular appears to be driven by information incorporated in downside higher-moment risk premiums.

3.2 Decomposing Higher-Moment Risk Premiums

3.2.1 Risk Premiums and Their Economic Interpretation

The literature includes a multitude of methods for calculating moment risk premium proxies. Early attempts by Coval and Shumway (2001) and Bakshi et al. (2003) indicate that trading strategies in S&P 500 options in which the market exposure is hedged yield negative average payoffs that can be understood as a variance risk premium. More recently, abundant research estimates proxies for premiums based on spreads between higher realized and implied moments (Da Fonseca and Xu 2017; Harris and Qiao 2018; Ruan and Zhang 2019). For third and fourth moments, however, these measures lack clear interpretation as risk premiums in the economic sense, which would require that the premium can be earned as a profit from a (synthetic) trading strategy.[23] This issue can be overcome using the approach developed by Kozhan et al. (2013), who provide trading strategies for variance and skewness swaps with variance and skewness measures that have the Neuberger (2012) aggregation property. Khrashchevskyi (2020) extends this model with a kurtosis swap and a kurtosis measure that has the aggregation property.

3.2.2 Higher-Moment Risk Premium Construction

Consider a twice-differentiable payoff function with payoff $g\left(r_{t,T}\right)$ at time T. The payoff is dependent on the log-return $r_{t,T}$ of futures prices $F_{t,T}$ between t and T. Following Neuberger and Payne (2021), I define the following g-functions:

$$g^V(r_{t,T}) = 2\left(e^{r_{t,T}} - 1 - r_{t,T}\right), \tag{3.1}$$

$$g^S(r_{t,T}) = 6\left(2 + r_{t,T} - 2e^{r_{t,T}} + r_{t,T}e^{r_{t,T}}\right), \tag{3.2}$$

$$g^K(r_{t,T}) = 12\left(r_{t,T}^2 + 2r_{t,T}e^{r_{t,T}} + 4r_{t,T} - 6e^{r_{t,T}} + 6\right), \tag{3.3}$$

where g^V, g^S, and g^K are return variance, return skewness, and return kurtosis mimicking payoffs, respectively. As Neuberger and Payne (2021) point out, an entire family of potential kurtosis mimicking measures fulfill the aggregation property. Khrashchevskyi (2020), for example, defines

$$g^K_{Khrash}(r_{t,T}) = 6\left(e^{2r_{t,T}} - 2r_{t,T} - 5 + 4e^{r_{t,T}} - 4r_{t,T}e^{r_{t,T}}\right) \tag{3.4}$$

as the kurtosis mimicking g-function. However, straightforward computations show that the fifth-order Taylor series expansion approximation error of $g^K_{Khrash}(r_{t,T})$ around 0 exceeds the

[23] Fan et al. (2020) likewise outline this aspect in a footnote.

approximation error of $g^K(r_{t,T})$ if $2^n > 2r_{t,T} + 2n$. This holds true for $n \geq 3$ and reasonable values for $r_{t,T}$, which is why I deviate from the literature and instead use $g^K(r_{t,T})$ to calculate the kurtosis risk premium.

Furthermore, let $\mathbb{Q}$ denote the risk-neutral measure and the forward price of the payoff g as $G_{t,T} = \mathbb{E}_t^{\mathbb{Q}}\left[g(r_{t,T})\right]$. Application of the spanning theorem of Bakshi and Madan (2000) and Carr and Madan (2001) gives the forward prices of the g-functions from Equations (3.1)–(3.3) that serve as the prices (fixed legs) of the moment swaps in a world with a tradeable continuum of option prices:

$$G_{t,T}^V = \frac{2}{B_{t,T}}\left\{\int_0^{F_{t,T}} \frac{P_{t,T}(K)}{K^2}dK + \int_{F_{t,T}}^{\infty} \frac{C_{t,T}(K)}{K^2}dK\right\}, \tag{3.5}$$

$$G_{t,T}^S = \frac{6}{B_{t,T}}\left\{\int_{F_{t,T}}^{\infty} \frac{(K - F_{t,T})C_{t,T}(K)}{K^2 F_{t,T}}dK - \int_0^{F_{t,T}} \frac{(F_{t,T} - K)P_{t,T}(K)}{K^2 F_{t,T}}dK\right\}, \tag{3.6}$$

$$\begin{aligned} G_{t,T}^K = \frac{24}{B_{t,T}}\Bigg\{ &\int_0^{F_{t,T}} \frac{K - F_{t,T}\left(1 + \ln \frac{K}{F_{t,T}}\right)}{F_{t,T}K^2} P_{t,T}(K)dK \\ &+ \int_{F_{t,T}}^{\infty} \frac{K - F_{t,T}\left(1 + \ln \frac{K}{F_{t,T}}\right)}{F_{t,T}K^2} C_{t,T}(K)dK\Bigg\}, \end{aligned} \tag{3.7}$$

where $B_{t,T}$ is the time t price of a bond paying one dollar at time T, and $C_{t,T}$ and $P_{t,T}$ are the prices of a continuum of out-of-the-money put and call options with strike prices K, respectively. Equations (3.5) and (3.6) are well known from Kozhan et al. (2013), and the derivation of Equation (3.7) can be found in Appendix A.1.

To determine the floating legs, I employ the results from Kozhan et al. (2013) for $g^V(r_{t,T})$ and $g^S(r_{t,T})$. They show that changes in the floating legs' values are given by

$$\delta Y_{t,T}^V = 2\left(e^{\delta f_{t,T}} - 1 - \delta f_{t,T}\right), \tag{3.8}$$

$$\delta Y_{t,T}^S = 3\delta G_{t,T}^E\left(e^{\delta f_{t,T}} - 1\right) + 6\left(2 - 2e^{\delta f_{t,T}} + \delta f_{t,T} + \delta f_{t,T}e^{\delta f_{t,T}}\right), \tag{3.9}$$

when investors discretely rebalance their portfolio positions between t and T, which is defined as the rebalancing period δ. This also holds for jumps in the underlying return process. $\delta Y_{t,T}$ denotes the change in a floating leg's value that is realized between t and T, $\delta f_{t,T}$ is the change of log futures prices within the rebalancing period, and $G_{t,T}^E$ is the forward price of the entropy contract introduced by Neuberger (2012). Moreover, as derived in Appendix A.2, the corresponding floating leg for Equation (3.3) is

$$\begin{aligned}\delta Y_{t,T}^{K} = 12\left(\delta G_{t,T}^{E} - \delta G_{t,T}^{V}\right)\left(e^{\delta f_{t,T}} - 1\right) + 6G_{t+\delta t,T}^{V}\left(e^{\delta f_{t,T}} - 1\right)^{2} \\ + 12\left(\delta f_{t,T}^{2} + 2\delta f_{t,T}e^{\delta f_{t,T}} + 4\delta f_{t,T} - 6e^{\delta f_{t,T}} + 6\right).\end{aligned} \quad (3.10)$$

3.2.3 Higher-Moment Risk Premium Decomposition

It is common knowledge that investors have asymmetric perceptions of risk with respect to losses and gains (Kahneman and Tversky 1979). Thus, it is advisable that more differentiated analyses of moment risks and the related premiums be conducted. Downside and upside decompositions offer a meaningful way of shedding light on the effects of investors' asymmetric risk perceptions with respect to moment risks. These premiums may incorporate different information and different investor preferences that are offset to a certain extent when only total moment risk premiums are considered. The value of decomposing (higher-) moments is also emphasized by Schneider et al. (2020), who argue that a decomposition of option-implied skewness into downside and upside skewness is informative because identical values of total skewness can result from different combinations of downside and upside skewness. This argument also holds for decomposed variance and kurtosis.

For example, evidence suggests that decomposition into downside and upside variance risk premiums increases the informational content for future S&P 500 index returns (Feunou et al. 2018; Kilic and Shaliastovich 2019). Beyond that, Held et al. (2020) study decomposed variance risk premiums in an international setting and find that downside and upside premiums in various countries have different signs and magnitudes. The economic rationale behind these results is that investors dislike bad (downside) uncertainty while favoring good (upside) uncertainty. Hence, they are willing to pay a premium to hedge bad uncertainty while agreeing to pay a price for exposure to variation in good uncertainty (Feunou et al. 2018).[24] What has been shown for variance risk is also conceivable for skewness risk and kurtosis risk. Investors' attitudes toward skewness and kurtosis risk may depend on the direction of risk, which would lead to different premiums for downside and upside risks.

For this reason, this paper suggests a decomposition of the aforementioned moment risk premiums into downside and upside risk premiums. Given the g-functions from Equations (3.1)–(3.3),

[24]Da Fonseca and Xu (2017) and Da Fonseca and Dawui (2021) provide some initial results for decomposed skewness risk premiums for selected commodities and currencies, that are, however, not applicable to the equity market because of very different market characteristics.

their payoffs can be written as the sum of a component associated with negative returns and a component associated with positive returns, respectively:

$$g(r_{t,T}) = g(r_{t,T})\mathbb{1}_{\{r_{t,T}\leq 0\}} + g(r_{t,T})\mathbb{1}_{\{r_{t,T}>0\}}, \tag{3.11}$$

where $\mathbb{1}$ is an indicator variable that takes the value 1 if the condition in the brackets is met and 0 otherwise. The same idea applies to the g-functions' prices, $G_{t,T}$, for which we can say

$$\begin{aligned} G_{t,T} &= \mathbb{E}_t^{\mathbb{Q}}\left[g(r_{t,T})\mathbb{1}_{\{r_{t,T}\leq 0\}}\right] + \mathbb{E}_t^{\mathbb{Q}}\left[g(r_{t,T})\mathbb{1}_{\{r_{t,T}>0\}}\right] \\ &= G_{t,T}^{down} + G_{t,T}^{up}. \end{aligned} \tag{3.12}$$

In their appendix, Kilic and Shaliastovich (2019) show that the downside component incorporates information solely from a continuum of out-of-the-money put options, while the upside component is determined by a continuum of out-of-the-money call options. As the forward prices for all g-functions are derived using the Bakshi and Madan (2000) and Carr and Madan (2001) spanning theorem, this property likewise transfers to skewness and kurtosis.

Moreover, this decomposition can also be applied to the floating legs:

$$\begin{aligned} \delta Y_{t,T} &= \delta Y_{t,T}\mathbb{1}_{\{r_{t,T}\leq 0\}} + \delta Y_{t,T}\mathbb{1}_{\{r_{t,T}>0\}} \\ &= \delta Y_{t,T}^{down} + \delta Y_{t,T}^{up}. \end{aligned} \tag{3.13}$$

Equation (3.13) yields a simple downside and upside structure, whereby changes in the floating legs' values can be directly attributed to market downturns or market upturns within the rebalancing period. Additionally, the suggested decomposition ensures that downside and upside premiums add to the total premium for variance, skewness, and kurtosis. This results in easily interpretable, economically meaningful premiums, highlighting the compensation for risks associated with downside and upside higher moments.

3.3 Decomposed Higher-Moment Risk Premiums

3.3.1 Data and Methodology

The data for the empirical study consist of European S&P 500 index options obtained from the OptionMetrics data-base. It provides historical closing bid and ask quotes from the Chicago Board Options Exchange (CBOE) as well as interest rates, spot prices of the underlying index, and implied dividend yields from January 1996 to December 2019. I eliminate all options with bid and ask quotes equal to or smaller than zero. Moreover, I require that ask quotes exceed bid quotes; options must not have a special settlement, and they must meet standard no-arbitrage conditions. Since there is no continuum of options available, I use discretized versions of Equations (3.1)–(3.10) and their corresponding decompositions as it is standard in the literature (Jiang and Tian 2005; Kozhan et al. 2013; Khrashchevskyi 2020). However, I do not apply any cubic spline interpolation to create a smoother grid of option prices because I want to emphasize the economic risk premium interpretation by using only the information contained in option prices available in the data-base and not from interpolated prices.

Moment risk premiums are calculated on a monthly basis from options that have an approximate time to maturity of one month. To achieve this, data from the first trading day after the third Friday in a month—the first trading day after the standard expiration date of options traded at the CBOE—were used. I require these options to expire in the next month, which gives the demanded one-month time to maturity. Then, fixed legs are determined using only out-of-the-money forward options, midquotes between bid and ask, and discretized versions of Equations (3.5)–(3.7). Next, the floating legs' values (i.e., realized moments) within this month are calculated according to the floating leg definitions in Equations (3.8)–(3.10). With the assumption of daily rebalancing, this yields the following implementation of the floating leg calculation:

$$\begin{aligned}
rv_{t,T} &= \sum_{i=1}^{n-1} 2\left(e^{r_{i,i+1}} - 1 - r_{i,i+1}\right), \\
rs_{t,T} &= \sum_{i=1}^{n-1} 3\delta v_{i,T}^{E}(e^{r_{i,i+1}} - 1) + 6(2 - 2e^{r_{i,i+1}} + r_{i,i+1} + r_{i,i+1}e^{r_{i,i+1}}), \\
rk_{t,T} &= \sum_{i=1}^{n-1} 12\left(\delta v_{i,T}^{E} - \delta v_{i,T}^{V}\right)\left(e^{r_{i,i+1}} - 1\right) + 6v_{i+1,T}^{V}\left(e^{r_{i,i+1}} - 1\right)^2 \\
&\quad + \sum_{i=1}^{n-1} 12\left(r_{i,i+1}^2 + 2r_{i,i+1}e^{r_{i,i+1}} + 4r_{i,i+1} - 6e^{r_{i,i+1}} + 6\right),
\end{aligned} \tag{3.14}$$

where for downside and upside premiums, the decomposition in Equation (3.13) can be applied. Given the fixed legs and floating legs of higher-moment swaps, this paper defines model-free higher-moment risk premiums as differences between these legs, such that the risk premiums' expected values are negative. Rubinstein (1973) shows that utility functions of rational investors favor high even and low odd moments. Thus, hedging instruments, such as moment swaps, should have positive payoffs when realized variance and realized kurtosis are high and realized skewness is low. Furthermore, to ensure consistency with studies that examine the predictive power of (higher-)moment risk premiums (e.g., Bollerslev et al. 2009; Kilic and Shaliastovich 2019; Fan et al. 2020), higher-moment risk premiums are defined as ex-ante expected premiums:

$$\begin{aligned} VRP_t^{tot} &= rv_{t-1,t} - v_{t,T}^{L}, \\ SRP_t^{tot} &= is_{t,T} - rs_{t-1,t}, \\ KRP_t^{tot} &= rk_{t-1,t} - ik_{t,T}. \end{aligned} \tag{3.15}$$

This definition assumes that realized moments follow a random walk.[25] Essentially, the advantage of this definition is that both implied and realized moments are model-free and observable in t. It can be applied to the downside and upside components as well as to the total fixed and floating legs. An overview of all formulas for the decomposed higher-moment risk premiums can be found in Table A.1 in Appendix A.3.

3.3.2 Descriptive Analysis

Descriptive statistics of the decomposed higher-moment risk premiums used in this paper are provided in Table 3.1. Panel A shows that the total variance risk premium is negative on average and that the downside variance risk premium (VRP^{down}) even exceeds the total premium; the upside premium (VRP^{up}) takes positive values on average. The 25%, 50%, and 75% quantiles of both VRP^{tot} and VRP^{down} are consistently negative, whereas the 50% quantile of VRP^{up} is close to zero and the 75% quantile is positive. Nevertheless, all mean values are statistically significantly different from zero. Furthermore, the premiums are only slightly first-order autocorrelated. The highest autocorrelation coefficient is 0.1981 (VRP^{down}). So far, this aligns well with the existing literature and with the economic intuition that investors dislike downside variance risk and favor upside variance risk.

Figure 3.1 visually confirms these results. It illustrates a plot of the three variance premiums and depicts that VRP^{tot} and VRP^{down} in particular are highly correlated. Additionally, the upside variance risk premium tends to fluctuate around zero, which leads to a negative overall variance risk premium, as VRP^{down} is mostly below zero. Moreover, all premiums show sharp peaks during the 2008 financial crisis. However, the upside variance risk premium's peaks are the lowest of the three.

Panel B in Table 3.1 reports the results for the total skewness risk premium and its decomposed parts. Just as with the variance risk premiums, the total premium and the downside premium

[25] See Bollerslev et al. (2009) for a discussion on this assumption.

Table 3.1: Decomposed Higher-Moment Risk Premiums Summary Statistics

Panel A: Variance Risk Premium

	VRP_t^{tot}	VRP_t^{down}	VRP_t^{up}
Mean	−8.8903	−10.5611	1.6708
t-value	−6.3324	−8.6376	2.1241
SD	27.0369	16.9214	13.7339
AR1	0.0972	0.1981	0.0393
25%	−16.5524	−14.5710	−2.6155
50%	−7.7793	−8.2808	−0.1380
75%	−3.0555	−4.8835	2.5524
Skewness	0.7986	−0.8804	3.3673
Kurtosis	33.4419	40.3951	23.1293

Panel B: Skewness Risk Premium

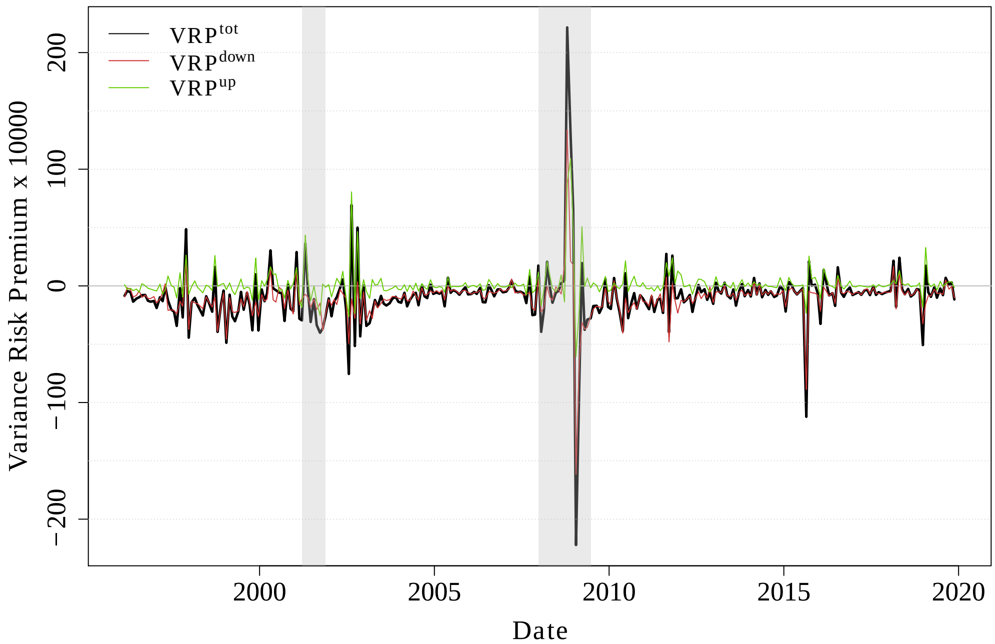
	SRP_t^{tot}	SRP_t^{down}	SRP_t^{up}
Mean	−1.8113	−4.1097	2.2984
t-value	−6.3738	−5.1560	2.8026
SD	6.0929	7.2528	5.4919
AR1	0.0587	0.4277	0.7258
25%	−2.3308	−4.1241	0.4492
50%	−1.2403	−2.3093	0.9334
75%	−0.5682	−1.2963	2.1581
Skewness	−4.8206	−6.9573	7.7019
Kurtosis	69.6162	67.8371	74.4625

Panel C: Kurtosis Risk Premium

	KRP_t^{tot}	KRP_t^{down}	KRP_t^{up}
Mean	−1.1637	−1.7458	0.5821
t-value	−4.4873	−3.2991	1.9765
SD	5.4049	5.3315	3.1474
AR1	−0.0447	0.3164	0.4079
25%	−0.8929	−1.1936	0.0189
50%	−0.4291	−0.6065	0.0686
75%	−0.1842	−0.2732	0.2244
Skewness	−5.6508	−7.9693	12.6954
Kurtosis	75.6297	75.8004	184.2868

Note: This table provides basic summary statistics of the variance risk premium (Panel A), the skewness risk premium (Panel B), and the kurtosis risk premium (Panel C). It shows the average premium (mean), the corresponding t-statistic adjusted for serial correlation according to Newey and West (1987), the standard deviation (SD) as well as the empirical 25%, 50% and 75% quantiles. Further, the first-order autocorrelation coefficient (AR1) as well as sample skewness and kurtosis are provided. For the sake of readability, premiums are scaled by a factor of 10000.

Figure 3.1: Decomposed Variance Risk Premium Plot

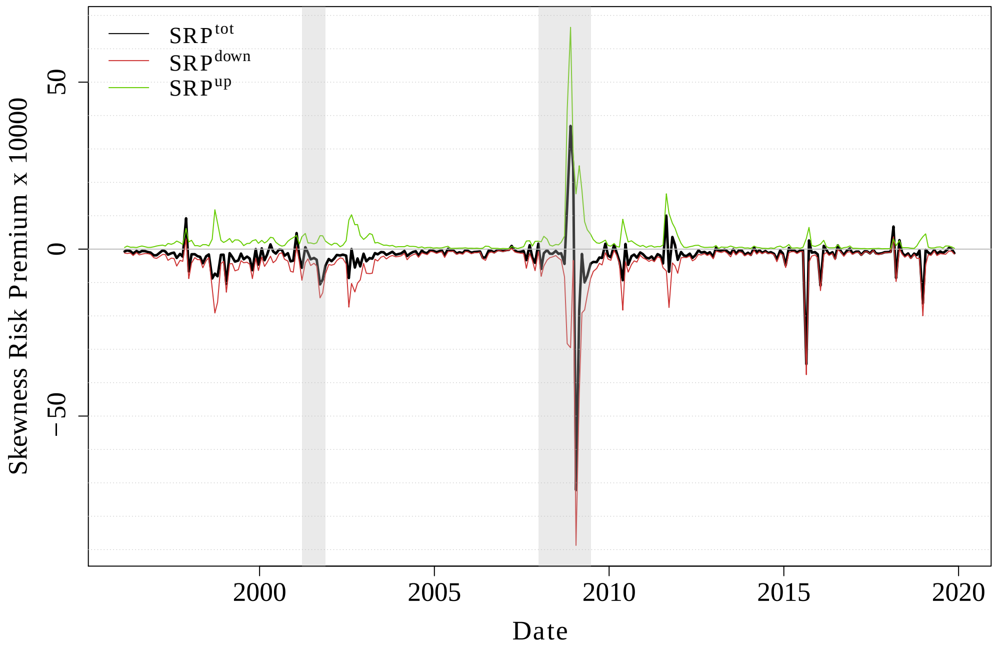


Note: This figure plots the total variance risk premium (black line) as well as the downside (red line) and the upside variance risk premium (green line) for the S&P 500 index from January 1996 to December 2019, respectively. The areas shaded in grey mark NBER recession periods.

are negative at statistically significant levels, whereas the upside premium is significantly positive. Owing to the total skewness risk premium's construction, a negative average premium may be anticipated, since investors are averse to negative skewness and are willing to pay a premium to hedge decreases in physical return skewness. The same argument applies to downside skewness, which is also perceived negatively by investors. A decrease in downside skewness, which can also be seen as a decrease in skewness of a return distribution that is truncated at 0, translates into an increase in undesired downside risks, which investors are willing to hedge. The opposite is true for upside skewness: here, investors seem to require compensation for exposure to assets that have low physical upside return skewness.

The 25%, 50%, and 75% quantiles of the empirical distributions are consistently negative for both SRP^{tot} and SRP^{down}, while they are all positive for SRP^{up}. The latter observation is in contrast to VRP^{tot}, for which only the 75% quantile is positive. SRP^{down} and SRP^{up} show exceptionally high autocorrelation, with the upside premium's coefficient of 0.7258. Figure 3.2, which plots the (decomposed) skewness risk premiums, further illustrates these aspects. A remarkable difference from the variance risk premium plot is that the downside and upside premiums are strongly negatively correlated. Moreover, all skewness premiums have lower standard deviations in absolute terms but higher in relation to their average values than the variance premiums. Furthermore, they exhibit more pronounced peaks than the variance premiums, with the sharpest peak during the 2008 financial crisis. Beyond that, peaks in downside premiums are larger in absolute terms than peaks in upside premiums.

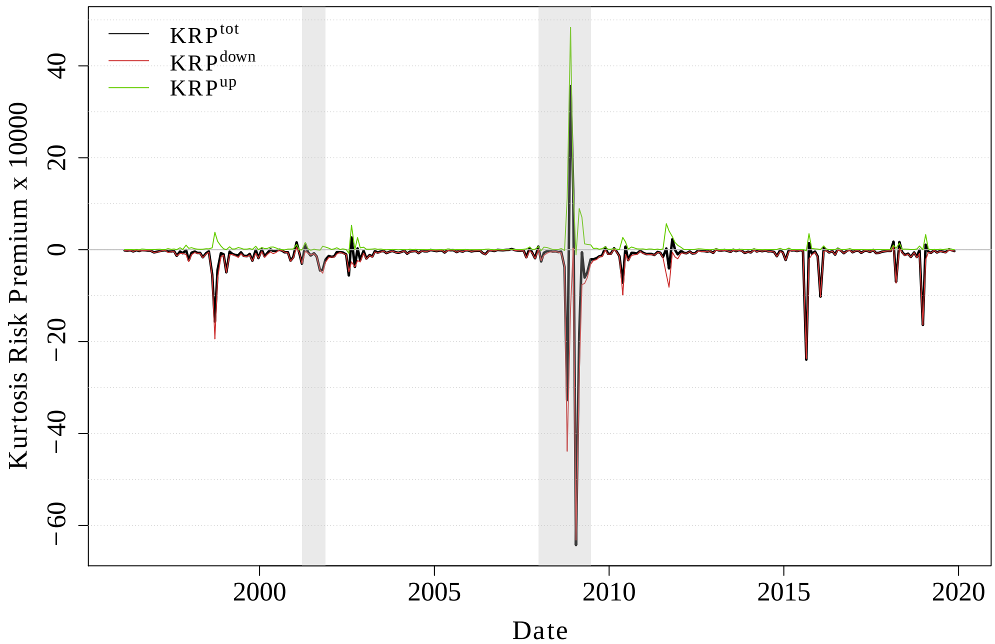
Figure 3.2: Decomposed Skewness Risk Premium Plot



Note: This figure plots the total skewness risk premium (black line) as well as the downside (red line) and the upside skewness risk premium (green line) for the S&P 500 index from January 1996 to December 2019, respectively. The areas shaded in grey mark NBER recession periods.

The results of the decomposed kurtosis risk premiums are presented in Panel C of Table 3.1. Economically, the same signs and mechanisms as for decomposed variance risk premiums may be anticipated. Investors dislike variation in kurtosis in general and variation in downside kurtosis in particular as it increases the probability of adverse return outcomes and results in a deterioration of the investment opportunity set. Thus, they are willing to pay a premium to hedge kurtosis risk. By contrast, investors might be willing to pay a premium for exposure to variation in upside kurtosis, which can be considered “good” kurtosis because it increases the likelihood of large positive tail outcomes. This economic intuition can be confirmed with a look at the mean premiums, which are statistically significantly negative for KRP^{tot} and KRP^{down} as well as significantly positive for KRP^{up}. Additionally, the first to third quartiles are consistently negative for KRP^{tot}, KRP^{down}, and positive for KRP^{up}, respectively, which is similar to the results presented in Panel B. Figure 3.3 illustrates a plot of the decomposed kurtosis risk premiums.

Figure 3.3: Decomposed Kurtosis Risk Premium Plot



Note: This figure plots the total kurtosis risk premium (black line) as well as the downside (red line) and the upside kurtosis risk premium (green line) for the S&P 500 index from January 1996 to December 2019, respectively. The areas shaded in grey mark NBER recession periods.

To conclude this descriptive analysis, Table 3.2 depicts the correlation matrix of the premiums. Some aspects must be highlighted here. First, almost all premiums are highly and statistically significantly correlated, with most coefficients being positive. A particularly high positive correlation of at least 0.8 is observed for some premium combinations, such as SRP^{tot} and KRP^{tot}, SRP^{down} and KRP^{down}, SRP^{up} and KRP^{up}. This emphasizes a strong similarity between skewness and kurtosis premiums and corroborates the visual impression from the first three figures that skewness risk premiums and kurtosis risk premiums appear to be very similar. Second, few premiums are significantly negative correlated. This applies to both upside and downside skewness and kurtosis premiums. Third, only three combinations of premiums do not exhibit a significant correlation in either direction. These are VRP^{up} and SRP^{down}, VRP^{up} and KRP^{down}, and SRP^{up} and KRP^{tot}.

Table 3.2: Decomposed Higher-Moment Risk Premium Correlation Matrix

	VRP_t^{tot}	VRP_t^{down}	VRP_t^{up}	SRP_t^{tot}	SRP_t^{down}	SRP_t^{up}	KRP_t^{tot}	KRP_t^{down}
VRP_t^{down}	0.9057***							
VRP_t^{up}	0.8527***	0.5509***						
SRP_t^{tot}	0.8276***	0.8064***	0.6358***					
SRP_t^{down}	0.4119***	0.5958***	0.0767	0.6740***				
SRP_t^{up}	0.3743***	0.1078*	0.6041***	0.2194***	−0.5729***			
KRP_t^{tot}	0.4843***	0.4598***	0.3869***	0.8405***	0.7295***	−0.0309		
KRP_t^{down}	0.2176***	0.3501***	−0.0030	0.5941***	0.9264***	−0.5643***	0.8282***	
KRP_t^{up}	0.4631***	0.1966***	0.6694***	0.4370***	−0.3165***	0.9028***	0.3143***	−0.2717***

Note: This table depicts Pearson's correlation coefficients between decomposed higher-moment risk premiums. Asterisks indicate statistical significance at the 1%(***), 5%(**), and 10%(*) level, respectively.

Overall, the descriptive analysis provides the following key results. Total premiums and downside premiums are negative on average, with the downside premiums being higher in absolute terms than total premiums. This is because upside premiums have positive average values that, to some degree, offset the downside premiums' negative values. Additionally, the higher the (order of the) moment, the larger the peaks that one can observe in the premiums' time series. This means that for the moments analyzed here, the empirical distributions become increasingly left-skewed and leptokurtic. Finally, Figures 3.1–3.3 reveal that premiums for skewness and kurtosis risk appear to be much more similar to each other than either is to the variance risk premiums, which is further supported by the results of the correlation matrix. This finding supports the suggestion of Neuberger and Payne (2021) that higher moments are more strongly affected by outliers in the underlying data.

3.4 Predicting S&P 500 Returns with Decomposed Higher-Moment Risk Premiums

The previous section's results suggest that the decomposition of higher-moment risk premiums uncovers information that is hidden in total higher-moment risk premiums because the effects of downside premiums and upside premiums might offset one another when only total premiums are considered. In this section, I utilize these findings and investigate the predictive power of decomposed higher-moment risk premiums both in an in-sample and and an out-of-sample study.

3.4.1 In-Sample Analysis

To investigate the predictability of future S&P 500 returns, I use predictive regressions that are set up as follows:

$$\frac{12}{h}\left(r^m_{t,t+h} - r^f_{t,t+h}\right) = \alpha_h + \beta^{'}_h X_t + \epsilon_{t,t+h}, \tag{3.16}$$

where h is the forecasting horizon of the regression, $r^m_{t,t+h} - r^f_{t,t+h}$ is the log market excess return over the forecasting period, and X_t is the matrix that contains monthly predictor variables. Thus, for $h > 1$, these regressions contain overlapping return observations, which are known to cause artificial autocorrelation in the regressions' error terms and therefore induce biases in the coefficients' standard errors (Hodrick 1992). I use the approach proposed by Britten-Jones et al. (2011) to transform regression models with overlap into models without overlap. For these models, statistical significance is then assessed using Newey and West (1987) robust standard errors to correct for heteroskedasticity and autocorrelation in the error terms with optimal lag length selection, according to Newey and West (1994).

First, the known prediction models from the literature are evaluated in terms of their predictive power when premiums are defined as in Section 3.2 and to provide three benchmark models to assess the potential gains in predictive power that result from decomposing higher-moment risk premiums. Second, I estimate prediction models that use different combinations of decomposed higher-moment risk premiums: (i) a model that includes only downside premiums, (ii) a model that includes only upside premiums, and (iii) a model that includes downside and upside premiums jointly. Third, the in-sample results are checked for plausibility by relating higher-moment risk premiums to changes in macroeconomic conditions.

Benchmark Models from the Literature The benchmark prediction models are presented in Table 3.3. Here, the annualized market excess return is regressed on the total variance risk premium (Panel A), on downside and upside variance risk premiums (Panel B), and on all total higher-moment risk premiums (Panel C). Panel A confirms the results already known from Bollerslev et al. (2009): the total variance risk premium's predictive power peaks at short horizons with an adjusted R^2 of around 5% and deteriorates at longer horizons, where the adjusted goodness of fit even partially takes negative values. The models as of $h = 12$ even have insignificant F-statistics. Likewise, the evidence for the decomposed premium model in Panel B corroborates findings reported by Kilic and Shaliastovich (2019). Decomposition of the premium yields slightly improved short-term predictability and improved longer-term predictability for up to 24 months. Moreover, the signs and significance of the regression coefficients also align with this study as the downside variance risk premium has a significant coefficient at almost every forecast horizon. Finally, consistent with Fan et al. (2020), Panel C shows the inclusion of skewness and kurtosis risk premiums increases the predictability of future market excess returns compared to the total variance risk premium prediction model, particularly at longer horizons of 12 months or longer. While the variance and kurtosis risk premiums have mostly negative regression coefficients, the skewness risk premium has positive coefficients. Moreover, the model in Panel C also outperforms the decomposed variance risk premium prediction model in terms of the adjusted R^2.

The benchmark prediction models' goodness of fit measures, however, are not as high as in earlier studies. This may be attributed to various factors. Unreported results indicate that the predictive power is stronger when I use a sample period from January 1996 to August 2014,

Table 3.3: Variance and Higher-Moment Risk Premium OLS Predictive Regressions (In-Sample)

Panel A: Total Variance Risk Premium

	Forecast horizon h						
	$h=1$	$h=3$	$h=6$	$h=9$	$h=12$	$h=18$	$h=24$
α	0.0069	0.0144	0.0264	0.0321	0.0349	0.0344	0.0319
	(0.0402)	(0.0321)	(0.0335)	(0.0357)	(0.0373)	(0.0380)	(0.0383)
VRP_t^{tot}	-39.7640^{***}	-29.7923^{***}	-15.6053^{***}	-8.5296^{***}	-3.8180	-2.0336	-2.7682
	(10.9820)	(5.6469)	(3.4568)	(3.1260)	(3.5420)	(2.9515)	(2.5331)
n	286	286	286	286	286	286	286
Adj. R^2	0.0330	0.0502	0.0329	0.0149	0.0023	-0.0011	0.0019
F-Stat	5.8766	8.5571	5.8571	3.1631	1.3350	0.8402	1.2721

Panel B: Decomposed Variance Risk Premium

	Forecast horizon h						
	$h=1$	$h=3$	$h=6$	$h=9$	$h=12$	$h=18$	$h=24$
α_h	-0.0197	-0.0055	-0.0039	0.0075	0.0141	0.0174	0.0138
	(0.0471)	(0.0387)	(0.0418)	(0.0406)	(0.0416)	(0.0415)	(0.0419)
VRP_t^{down}	-60.6695^{***}	-45.3312^{***}	-39.0623^{***}	-27.5255^{**}	-20.0513^{*}	-15.1013	-16.5797^{*}
	(17.5027)	(13.2159)	(13.7493)	(13.1244)	(10.9119)	(10.6730)	(9.5806)
VRP_t^{up}	-12.4072	-9.4721	14.9907	16.2350	17.5627	15.1841	15.2553^{*}
	(38.0080)	(15.8092)	(16.4802)	(14.9379)	(11.7761)	(10.1937)	(9.0871)
n	286	286	286	286	286	286	286
Adj. R^2	0.0332	0.0473	0.0345	0.0173	0.0056	0.0016	0.0051
F-Stat	4.2702	5.7331	4.4065	2.6806	1.5339	1.1551	1.4920

Panel C: Total Higher-Moment Risk Premiums

	Forecast horizon h						
	$h=1$	$h=3$	$h=6$	$h=9$	$h=12$	$h=18$	$h=24$
α_h	0.0118	0.0140	0.0253	0.0307	0.0331	0.0331	0.0308
	(0.0349)	(0.0315)	(0.0330)	(0.0347)	(0.0360)	(0.0366)	(0.0373)
VRP_t^{tot}	-70.4442	-48.5580^{***}	-35.3237^{***}	-26.8606^{***}	-16.4349^{*}	-11.2389	-9.2766^{*}
	(46.4866)	(18.0970)	(10.8565)	(9.3122)	(8.5601)	(7.2934)	(5.5751)
SRP_t^{tot}	111.2548	146.5252	165.3004^{***}	164.7562^{**}	127.4386^{**}	90.4954^{*}	68.9359^{*}
	(307.8590)	(109.7015)	(62.3654)	(63.5435)	(54.6782)	(47.5145)	(37.6968)
KRP_t^{tot}	102.5647	-88.4886	-114.9520^{***}	-128.2527^{***}	-115.8369^{***}	-80.3639^{**}	-66.3924^{**}
	(196.1159)	(67.1137)	(44.0143)	(43.4739)	(38.8758)	(35.3341)	(30.3342)
n	286	286	286	286	286	286	286
Adj. R^2	0.0476	0.0539	0.0532	0.0434	0.0285	0.0115	0.0150
F-Stat	4.5704	5.0746	5.0148	4.2470	3.0965	1.8344	2.0901

Note: This table presents evidence for predictive regressions of S&P 500 Index excess returns on the total variance risk premium (Panel A), downside and upside variance risk premiums (Panel B), and total higher-moment risk premiums (Panel C). The sample period spans from January 1996 to December 2019 and considered forecast horizons are $h = 1, 3, 6, 9, 12, 18, 24$ months. To correct for biases introduced in the standard errors due to overlapping return periods in the dependent variable, the regression model is transformed according to the approach proposed in Britten-Jones et al. (2011). Standard errors are reported in parentheses and adjusted for serial correlation and heteroskedasticity according to Newey and West (1987) with optimal lag length selection according to Newey and West (1994). Asterisks indicate statistical significance of two-sided t-tests at the 1%(***), 5%(**), and 10%(*) level, respectively.

which corresponds to that used in Kilic and Shaliastovich (2019). Kilic and Shaliastovich (2019), as a robustness check, also provide evidence for a longer sample period, which is weaker but still provides meaningful results that are similar to the results in this paper. Thus, a longer sample period tends to weaken the reported relation to some extent. Moreover, different premium definitions can cause results to diverge. As noted, this paper's definitions deviate from those offered in most of the existing literature to shed light on the predictive power of *economically* meaningful premiums as a prediction signal. Additionally, results might differ as a result of different regression approaches, the use of orthogonalized moment spreads, as in Fan et al. (2020), and different data sets. However, the economic intuition of the results remains unchanged.

Decomposed Higher-Moment Risk Premiums I next estimate three prediction models that employ decomposed higher-moment risk premiums. First, I use all downside moment risk premiums up to kurtosis as predictor variables. Second, I employ all upside moment risk premiums up to kurtosis. Third, I study a model that uses all downside and upside higher-moment risk premiums in combination.

The downside risk premium model is presented in Panel A of Table 3.4. Solely considering downside higher-moment risk premiums as predictor variables yields improved predictions compared to the (decomposed) variance risk premium benchmark models at almost every horizon but particularly at horizons of 6 months or longer. Furthermore, downside higher-moment risk premiums improve the prediction of S&P 500 excess returns at a forecast horizon of 12 months or longer, compared to the total higher-moment risk premium prediction model. Thus, downside higher-moment risk premiums seem to incorporate valuable information about future market excess returns that go beyond the information that can be extracted from the variance risk premium and all total higher-moment risk premiums.

By contrast, upside higher-moment risk premiums yield inferior prediction results than total higher-moment risk premiums at every considered forecast horizon, as depicted in Panel B of Table 3.4. At a one- and three-month prediction horizon, this model also produces inferior results to those of (decomposed) variance risk premium predictors. Moreover, upside higher-moment risk premiums are inferior to downside higher-moment risk premiums at every considered forecast horizon. With few exceptions, in both models, the variance and kurtosis risk premiums have negative coefficient estimates while the skewness risk premiums have positive coefficient estimates, which aligns with the results from the benchmark prediction models.

Table 3.4: Downside and Upside Higher-Moment Risk Premium OLS Predictive Regression (In-Sample)

Panel A: Downside Higher-Moment Risk Premiums

	Forecast horizon h						
	$h=1$	$h=3$	$h=6$	$h=9$	$h=12$	$h=18$	$h=24$
α_h	−0.0335	0.0022	0.0077	0.0179	0.0246	0.0285	0.0218
	(0.0440)	(0.0330)	(0.0332)	(0.0318)	(0.0312)	(0.0315)	(0.0338)
VRP_t^{down}	−57.4011*	−73.4275***	−35.1652**	−19.4640	−10.3073	−9.8745	−5.8314
	(29.2941)	(18.3404)	(15.1892)	(15.8049)	(16.2670)	(14.4920)	(12.1836)
SRP_t^{down}	−179.8708	161.4081*	45.9143	33.7403	32.4724	39.8879	0.6062
	(160.3983)	(97.4094)	(77.8988)	(97.8660)	(108.3468)	(104.0037)	(86.8040)
KRP_t^{down}	336.5421*	−157.3177	−80.6139	−85.0557	−91.6712	−79.0980	−37.3619
	(190.5717)	(96.2601)	(77.3827)	(98.5281)	(111.2422)	(99.6609)	(83.9657)
n	286	286	286	286	286	286	286
Adj. R^2	0.0461	0.0474	0.0482	0.0407	0.0342	0.0208	0.0252
F-Stat	4.4547	4.5587	4.6244	4.0366	3.5317	2.5152	2.8467

Panel B: Upside Higher-Moment Risk Premiums

	Forecast horizon h						
	$h=1$	$h=3$	$h=6$	$h=9$	$h=12$	$h=18$	$h=24$
α_h	0.0604	0.0368	0.0266	0.0242	0.0226	0.0258	0.0221
	(0.0433)	(0.0325)	(0.0330)	(0.0332)	(0.0328)	(0.0341)	(0.0380)
VRP_t^{up}	−90.9522**	−46.0119***	−34.0937***	−25.0546***	−14.8368**	−5.2300	−8.0137*
	(39.7464)	(12.7741)	(8.5144)	(6.5625)	(6.8322)	(6.3776)	(4.2442)
SRP_t^{up}	−123.4044	70.0427	77.7022	78.7751	78.0117	52.3066	62.6287
	(288.1967)	(115.2255)	(98.0458)	(77.4765)	(69.8003)	(59.6683)	(52.9772)
KRP_t^{up}	437.9106	−72.6076	24.8393	23.2171	−4.0042	−18.9800	−23.3116
	(385.4481)	(182.0495)	(135.6354)	(111.8602)	(102.2153)	(75.4955)	(59.8836)
n	286	286	286	286	286	286	286
Adj. R^2	0.0199	0.0350	0.0356	0.0308	0.0143	−0.0028	0.0096
F-Stat	2.4533	3.5933	3.6406	3.2743	2.0409	0.8019	1.6900

Note: This table presents evidence for predictive regressions of S&P 500 Index excess returns on downside higher-moment risk premiums (Panel A) and upside higher-moment risk premiums (Panel B). The sample period spans from January 1996 to December 2019 and considered forecast horizons are $h = 1, 3, 6, 9, 12, 18, 24$ months. To correct for biases introduced in the standard errors due to overlapping return periods in the dependent variable, the regression model is transformed according to the approach proposed in Britten-Jones et al. (2011). Standard errors are reported in parentheses and adjusted for serial correlation and heteroskedasticity according to Newey and West (1987) with optimal lag length selection according to Newey and West (1994). Asterisks indicate statistical significance of two-sided t-tests at the 1%(***), 5%(**), and 10%(*) level, respectively.

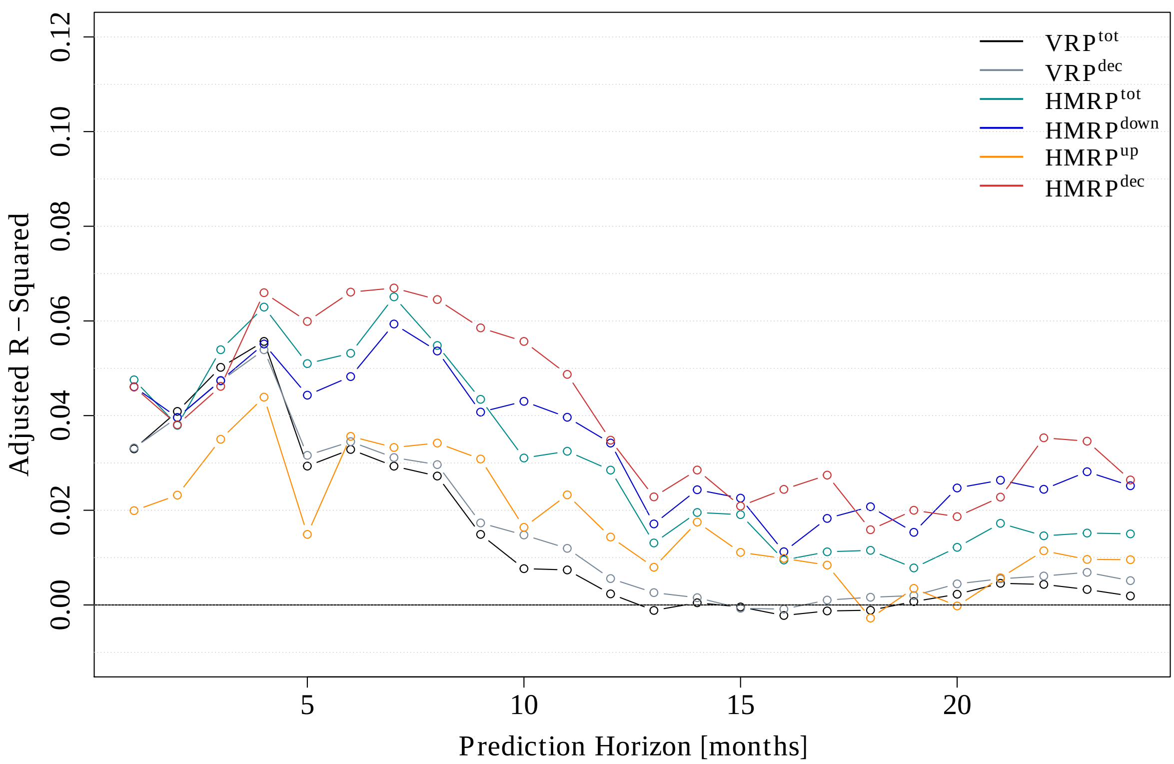
Table 3.5: Decomposed Higher-Moment Risk Premium OLS Predictive Regression (In-Sample)

	Forecast horizon h						
	$h=1$	$h=3$	$h=6$	$h=9$	$h=12$	$h=18$	$h=24$
α_h	−0.0381	−0.0186	−0.0103	0.0047	0.0146	0.0168	0.0116
	(0.0500)	(0.0394)	(0.0362)	(0.0337)	(0.0333)	(0.0334)	(0.0344)
VRP_t^{down}	−108.8602	−120.5138**	−107.8078***	−84.5961***	−63.5125**	−54.2327**	−42.7749**
	(89.0862)	(54.7729)	(34.7392)	(32.1635)	(32.1213)	(25.9080)	(21.4614)
VRP_t^{up}	−51.0701	−11.5757	−9.7145	−10.5767	−3.7746	3.6718	1.9242
	(60.5994)	(18.4403)	(10.9511)	(10.4465)	(9.4586)	(7.8244)	(5.6725)
SRP_t^{down}	174.4653	417.1339	437.5002**	389.9901**	316.1470*	262.8394*	187.8878
	(456.4591)	(271.5681)	(175.8419)	(192.0159)	(184.4387)	(144.2092)	(121.3786)
SRP_t^{up}	293.2773	295.2608*	259.6428**	207.9134*	141.7878	122.0367	110.1262
	(306.2757)	(152.8023)	(123.7200)	(110.2134)	(112.3065)	(87.7656)	(77.9376)
KRP_t^{down}	110.5486	−292.8208	−352.5977**	−341.9266*	−304.2773	−242.5278	−172.0549
	(377.8420)	(241.5390)	(173.4303)	(202.7854)	(198.3125)	(154.7014)	(129.7769)
KRP_t^{up}	−41.7943	−231.0406	−61.9561	−5.8186	21.7160	−16.8233	−23.6447
	(268.4822)	(208.6373)	(170.6451)	(184.5891)	(190.1824)	(140.2888)	(115.0972)
n	286	286	286	286	286	286	286
Adj. R^2	0.0460	0.0462	0.0661	0.0585	0.0348	0.0159	0.0264
F-Stat	2.9713	2.9781	3.8913	3.5403	2.4736	1.6597	2.1080

Note: This table presents evidence for predictive regressions of S&P 500 Index excess returns on decomposed higher-moment risk premiums. The model jointly employs downside and upside premiums up to the fourth moment as predictor variables. The sample period spans from January 1996 to December 2019 and considered forecast horizons are $h = 1, 3, 6, 9, 12, 18, 24$ months. To correct for biases introduced in the standard errors due to overlapping return periods in the dependent variable, the regression model is transformed according to the approach proposed in Britten-Jones et al. (2011). Standard errors are reported in parentheses and adjusted for serial correlation and heteroskedasticity according to Newey and West (1987) with optimal lag length selection according to Newey and West (1994). Asterisks indicate statistical significance of two-sided t-tests at the 1%(***), 5%(**), and 10%(*) level, respectively.

When all decomposed higher-moment risk premiums for the return prediction are considered (shown in Table 3.5), the other forecasting models—benchmark as well as downside and upside—can be outperformed. The prediction goodness of fit peaks at the six-month forecast horizon and subsequently deteriorates slowly. Nevertheless, it is the best prediction model at most horizons. Only at horizons of 12 months or longer, this model produces results that are similar to those of the downside higher-moment risk premium prediction model. The coefficient estimates still show mostly negative (positive) values for variance and kurtosis (skewness) risk premiums, but particularly the downside premiums' coefficients show to differ from zero at statistically significant levels. Furthermore, the coefficient estimates of downside premiums tend to have larger absolute values than those of upside premiums.

Figure 3.4: Adjusted R^2 of In-Sample OLS Predictive Regressions



Note: This figure shows the adjusted R^2 of various in-sample OLS predictive regressions of the annualized market excess return on specific higher-moment risk premium combinations. Results are shown for regressions on the total variance risk premium (VRP^{tot}), a joint model with both decomposed variance risk premiums (VRP^{dec}), and a model with all total higher-moment risk premiums up to the fourth moment (HMRP^{tot}). Moreover, goodness-of-fit measures for a model with, downside higher-moment risk premiums (HMRP^{down}), upside higher-moment risk premiums (HMRP^{up}), and a model that jointly employs all downside and upside higher-moment risk premiums up to kurtosis (HMRP^{dec}) as predictors are plotted. Subject of the study are prediction horizons from 1 to 24 months.

A visual comparison of all the mentioned in-sample prediction models can be found in Figure 3.4. It shows the adjusted R^2 for the benchmark prediction models and the decomposed higher-moment risk premium prediction models estimated for forecast horizons of 1 to 24 months.[26] Figure 3.4 illustrates that in-sample market excess return predictions can be enhanced by decomposing total higher-moment risk premiums into their downside and upside components and using these premiums in tandem as predictor variables. Moreover, it emphasizes the finding that the predictive power of downside risk premiums is stronger than that of upside premiums.

Thus, the evidence so far suggests that (i) the decomposition of total higher-moment risk premiums into their downside and upside components increases the in-sample predictive power for market excess returns compared to the known benchmark prediction models from the literature; (ii) upside premiums are inferior to downside premiums in predicting subsequent market excess returns; (iii) the joint model of all decomposed higher-moment risk premiums yields the best predictions, particularly at forecast horizons of up to 12 months, whereas subsequent predictions are of similar quality to those from the downside premium model.

Predicting Changes in Macroeconomic Conditions As summarized by Rapach and Zhou (2013), stock return predictability is closely linked to fluctuations in the business cycle and is not necessarily an indicator of market inefficiency. The authors argue that asset prices are functions of state variables of the real economy and that the real economy itself is subject to fluctuations in the business cycle. Now, if the quantity and price of aggregate risk are linked to business cycle fluctuations, time-varying risk and time-varying risk premiums may be anticipated. Variables that are capable of predicting macroeconomic conditions should thus also be able to predict returns, which makes return predictability reasonable without necessarily being inconsistent with rational asset pricing and market efficiency (Cochrane 2007). Hence, a predictor variable that is related to macroeconomic risk and able to predict economic fluctuations is a more credible predictor because its predictability can be traced back to time-varying risk premiums caused by changes in aggregate risk (Rapach and Zhou 2013; Lin et al. 2018).

Regarding higher-moment risk premiums, Khrashchevskyi (2020) provides evidence of the connection between higher-moment risk premiums and macroeconomic risk. To further assess the eligibility of decomposed higher-moment risk premiums as return predictors, I perform a predictive regression—as suggested by Cochrane (2007) and similarly performed by Lin et al. (2018)—to determine whether decomposed higher-moment risk premiums can forecast changes in

[26] Full regression results for unreported specifications are available on request.

macroeconomic conditions:

$$\Delta Y_{t,t+h} = \alpha + \beta_h^{'} X_t + \epsilon_{t,t+h}, \tag{3.17}$$

where $\Delta Y_{t,t+h}$ is the change in a macroeconomic variable between t and $t+h$, h is the prediction horizon measured in months, and X_t is a vector of predictor variables. The choice of macroeconomic variables Y_t follows Lin et al. (2018): I use the smoothed recession probability (SRP), industrial production growth (IPG), Treasury Bill rates (TBL), default yield spreads (DFY), Chicago Fed National Activity Index (CFNAI), and the Aruoba et al. (2009) business conditions index (ADSI).

Table 3.6 presents the in-sample R^2s of the predictive regressions for three different forecasting horizons and six different sets of predictors. The forecast horizons analyzed are $h = 3$, $h = 12$, and $h = 24$ to assess the predictive power for changes in macroeconomic conditions across a broad range of horizons. Predictors are chosen in accordance with the main in-sample analysis described above and comprise the three benchmark models and the three decomposed higher-moment risk premium models.

Panel A illustrates that downside higher-moment risk premiums can predict changes in macroeconomic variables at a three-month horizon. The predictive power of HMRP_t^{up} is weaker than that of HMRP_t^{down} for the three-month prediction horizon. Moreover, the joint model HMRP_t^{dec} in most cases substantially increases the predictive power for changes in macroeconomic conditions compared to the benchmark predictors. The results presented in Panels B and C are similar but slightly weaker than the evidence for the three-month prediction horizon. Some variables, such as IPG and CFNAI, can no longer be predicted at longer horizons. The predictive power of upside higher-moment risk premiums, however, improves at longer forecasting horizons. All in all, decomposition increases the predictability of changes in macroeconomic conditions, providing evidence that moment risk premiums are related to changes in macroeconomic conditions and that decomposition of these premiums is highly informative with respect to predictions of macroeconomic conditions. To sum up, the analysis supports the idea that decomposed higher-moment risk premiums are suitable return predictors because their predictive power can be credibly traced back to time-varying risk premiums caused by changes in aggregate risk or risk aversion.

Additional Robustness Checks To assess the robustness of the previous results, I conduct several further robustness checks. First, I additionally estimate the regression model in Equation (3.16)

Table 3.6: Prediction of Macroeconomic Conditions

Panel A: Forecast horizon $h = 3$

	SRP	IPG	TBL	DFY	CFNAI	ADSI
VRP_t^{tot}	0.0079	−0.0038	−0.0001	−0.0060	0.0003	0.0262
VRP_t^{dec}	0.0334	−0.0028	−0.0016	0.0701	0.0067	0.0399
$HMRP_t^{tot}$	0.0217	0.0038	−0.0041	−0.0090	0.0059	0.0199
$HMRP_t^{down}$	0.0354	0.0000	0.0242	0.0181	0.0077	0.0386
$HMRP_t^{up}$	0.0010	−0.0096	0.0341	0.0092	0.0008	0.0102
$HMRP_t^{dec}$	0.0372	0.0154	0.0880	0.0935	0.0442	0.0481

Panel B: Forecast horizon $h = 12$

	SRP	IPG	TBL	DFY	CFNAI	ADSI
VRP_t^{tot}	−0.0026	−0.0054	−0.0032	−0.0027	−0.0063	−0.0068
VRP_t^{dec}	0.0828	−0.0087	−0.0064	0.0685	−0.0059	0.0093
$HMRP_t^{tot}$	0.1120	−0.0118	−0.0089	0.0515	−0.0091	0.0115
$HMRP_t^{down}$	0.1314	−0.0118	0.0313	0.0616	−0.0070	0.0180
$HMRP_t^{up}$	0.1352	−0.0069	0.0349	0.0635	−0.0029	0.0249
$HMRP_t^{dec}$	0.1339	−0.0135	0.0803	0.0704	−0.0077	0.0198

Panel C: Forecast horizon $h = 24$

	SRP	IPG	TBL	DFY	CFNAI	ADSI
VRP_t^{tot}	−0.0028	−0.0070	−0.0014	−0.0070	−0.0070	−0.0065
VRP_t^{dec}	0.0530	−0.0101	−0.0041	0.0412	−0.0071	0.0034
$HMRP_t^{tot}$	0.0524	−0.0140	−0.0012	0.0217	−0.0114	−0.0027
$HMRP_t^{down}$	0.0620	−0.0133	0.0368	0.0273	−0.0107	0.0013
$HMRP_t^{up}$	0.0583	−0.0137	0.0049	0.0318	−0.0107	0.0019
$HMRP_t^{dec}$	0.0519	−0.0196	0.0414	0.0363	−0.0204	−0.0058

Note: This table reports in-sample R^2s of predictive regressions where the dependent variable is the change in macroeconomic conditions and the independent variables comprise various combinations of (decomposed) higher-moment risk premiums. Considered variables that gauge macroeconomic conditions are smoothed recession probability (SRP), industrial production growth (IPG), treasury bill rates (TBL), default yield spreads (DFY), the Chicago Fed National Activity Index (CFNAI), and the Aruoba et al. (2009) business conditions index (ADSI). Results are shown for regressions on the total variance risk premium (VRP^{tot}), a joint model with both decomposed variance risk premiums (VRP^{dec}), and a model with all total higher-moment risk premiums up to the fourth moment ($HMRP^{tot}$). Moreover, goodness-of-fit measures for a model with all downside and upside higher-moment risk premiums up to kurtosis ($HMRP^{dec}$) as well as all downside higher-moment risk premiums ($HMRP^{down}$) are depicted. The regression models are transformed according to the approach proposed in Britten-Jones et al. (2011). Analyzed forecast horizons are $h = 3$, $h = 12$, and $h = 24$ and the sample period is from January 1996 to December 2019.

with a more robust least absolute deviation (LAD) approach that puts less weight on more extreme observations than an ordinary least squares (OLS) estimator. This ensures that the results are not driven by the extreme peaks that can be observed in the time series of higher-moment risk premiums. The results from these estimations are similar to those reported in Tables 3.3–3.5 and confirm earlier conclusions. Second, I assess the robustness of my results by applying other option filters and data preparation methods. The results show to be robust to additional option filters, such as trading volume > 0 and option midquote > 0.125, as well as to spline-interpolated option data. Beyond that, I can confirm my conclusions when moment spreads are calculated with the methodology of Bakshi et al. (2003) and when these spreads are used as predictors in the predictive regression model.[27]

3.4.2 Out-of-Sample Analysis

Useful and economically significant predictors should not only be able to predict subsequent excess market returns in-sample but should also survive an out-of-sample assessment. A commonly applied method to evaluate out-of-sample predictive power is the comparison of a nested prediction model with a naïve benchmark predictor by means of the out-of-sample R^2 (R^2_{OOS}) (Campbell and Thompson 2008; Welch and Goyal 2008), which is calculated as follows:

$$R^2_{OOS} = 1 - \frac{\sum_{t=1}^{T-h} \left(r_{t,t+h} - \hat{r}_{t,t+h}\right)}{\sum_{t=1}^{T-h} \left(r_{t,t+h} - \bar{r}_{t,t+h}\right)}. \tag{3.18}$$

Here, $r_{t,t+h}$ is the realized market return between t and $t+h$, while $\hat{r}_{t,t+h}$ and $\bar{r}_{t,t+h}$ are the returns predicted by the nested model and the naïve predictor, respectively. R^2_{OOS} is defined in the interval $(-\infty, 1]$, where positive values indicate superior out-of-sample prediction power of the nested model compared to the benchmark predictor. The statistical significance of this measure can be formally assessed by using the test statistic of Clark and West (2007). I use the first 84 months as an initial sample to calibrate the first prediction model and then use an expanding window approach wherein I include every new observation to estimate the next prediction model and the next predicted return. All sets of predictors that were analyzed in the foregoing section are subject to evaluation.

To address the concern of uncertain and unstable out-of-sample prediction models due to overfitted in-sample models (Rapach et al. 2010), I use forecast combinations to exploit the predictive power of decomposed higher-moment risk premiums in the out-of-sample study. The goal of the

[27] Results of the additional robustness checks are available upon request.

following analysis is to construct a combination forecast that assigns equal weights to prediction signals deriving from (i) (decomposed) variance and from (ii) (decomposed) skewness and kurtosis because skewness and kurtosis in particular are highly correlated with one another. This is intended to ensure that extreme observations are not assigned too much weight. Therefore, the (decomposed) variance risk premium prediction enters with 50% into the combination forecast, while (decomposed) skewness and kurtosis risk premiums enter with 25%, respectively, such that the following combination forecasts are constructed:

$$\begin{aligned} Comb_{t,t+h}^{tot/down/up} &= 0.5\hat{r}_{t,t+h}^{VRP_t^{tot/down/up}} + 0.25\hat{r}_{t,t+h}^{SRP_t^{tot/down/up}} + 0.25\hat{r}_{t,t+h}^{KRP_t^{tot/down/up}} \\ Comb_{t,t+h}^{dec} &= 0.25\hat{r}_{t,t+h}^{VRP_t^{down}} + 0.25\hat{r}_{t,t+h}^{VRP_t^{up}} + 0.125\hat{r}_{t,t+h}^{SRP_t^{down}} + 0.125\hat{r}_{t,t+h}^{SRP_t^{up}} \\ &+ 0.125\hat{r}_{t,t+h}^{KRP_t^{down}} + 0.125\hat{r}_{t,t+h}^{KRP_t^{up}}, \end{aligned} \tag{3.19}$$

where $\text{Comb}_{t,t+h}$ is the forecast combination for market excess return predictions between t and $t+h$ and $\hat{r}_{t,t+h}$ is the return prediction of the respective premium that is stated in the superscript. With this construction, variance risk premium information contributes 50% to the combination forecast, while skewness and kurtosis risk premium information together contribute the remaining 50%.[28]

As an additional robustness check, I not only compare OLS predictions with the historical mean as a simple benchmark predictor, but I also use LAD regressions and compare them with the historical median of market returns. The latter approach is important because peaks in higher-moment risk premiums represent highly influential observations in OLS estimation, as gauged by Cook's distance. This effect is even stronger for a smaller sample size, as in the out-of-sample assessment. Hence, a more robust estimation methodology, such as LAD, provides information that is less distorted by highly influential observations because it downweighs the informational content in these observations compared to OLS regressions. The evidence for this out-of-sample study is shown in Table 3.7.

Panel A depicts the results of the OLS model predictions. At the one-month forecasting horizon, no model manages to outperform the historical mean prediction as every R^2_{OOS} is negative. However, at $h = 3$, VRP_t^{tot} and VRP_t^{dec} realize R^2_{OOS} values higher than 0.07 that are statistically significantly different from zero. This evidence is supportive of the in-sample results, as these two prediction models make their best in-sample predictions at a three-month forecasting horizon.

[28]Unreported results indicate that the out-of-sample results are robust to other forecast combination constructions, such as mean and median forecasts (Rapach et al. 2010), as well as Stock and Watson (2004) weighing schemes, which take prediction errors into account.

Table 3.7: R^2_{OOS} of Predictive Regressions

Panel A: Ordinary Least Squares Predictive Regression – Benchmark: Mean

	Forecast horizon h						
	$h=1$	$h=3$	$h=6$	$h=9$	$h=12$	$h=18$	$h=24$
VRP_t^{tot}	−0.0225	0.0754*	0.0330*	0.0242*	−0.0036	−0.0106	0.0037
VRP_t^{dec}	−0.0683	0.0729*	0.0399*	0.0010	−0.0069	−0.0021	0.0125
Comb_t^{tot}	−0.0315	0.0278*	−0.0290	0.0237	0.0044	−0.0061	0.0124
Comb_t^{down}	−0.0050	0.0006	0.0362	0.0379	0.0165	0.0144	0.0265
Comb_t^{up}	−0.1514	−0.0478	0.0040	−0.0099	−0.0005	0.0128	−0.0096
Comb_t^{dec}	−0.0458	−0.0181	0.0380	0.0183	0.0114	0.0239	0.0132

Panel B: Least Absolute Deviation Predictive Regression – Benchmark: Median

	Forecast horizon h						
	$h=1$	$h=3$	$h=6$	$h=9$	$h=12$	$h=18$	$h=24$
VRP_t^{tot}	−0.0362	0.0868*	0.0283*	0.0246*	−0.0464	−0.0692	−0.0314
VRP_t^{dec}	−0.0748	0.0653*	0.0149**	0.0155**	−0.0061	−0.0400	0.0499**
Comb_t^{tot}	−0.0798	0.0113*	−0.0222	0.0151	−0.0106	−0.0255	0.0012
Comb_t^{down}	−0.0671	0.0069**	0.0372*	0.0466	0.0427*	0.0052*	0.0469*
Comb_t^{up}	−0.2392	0.0099	−0.1390	−0.0212	0.0040	−0.0078	0.0129
Comb_t^{dec}	−0.0993	0.0225*	−0.0048	0.0210	0.0329	0.0338	0.0384*

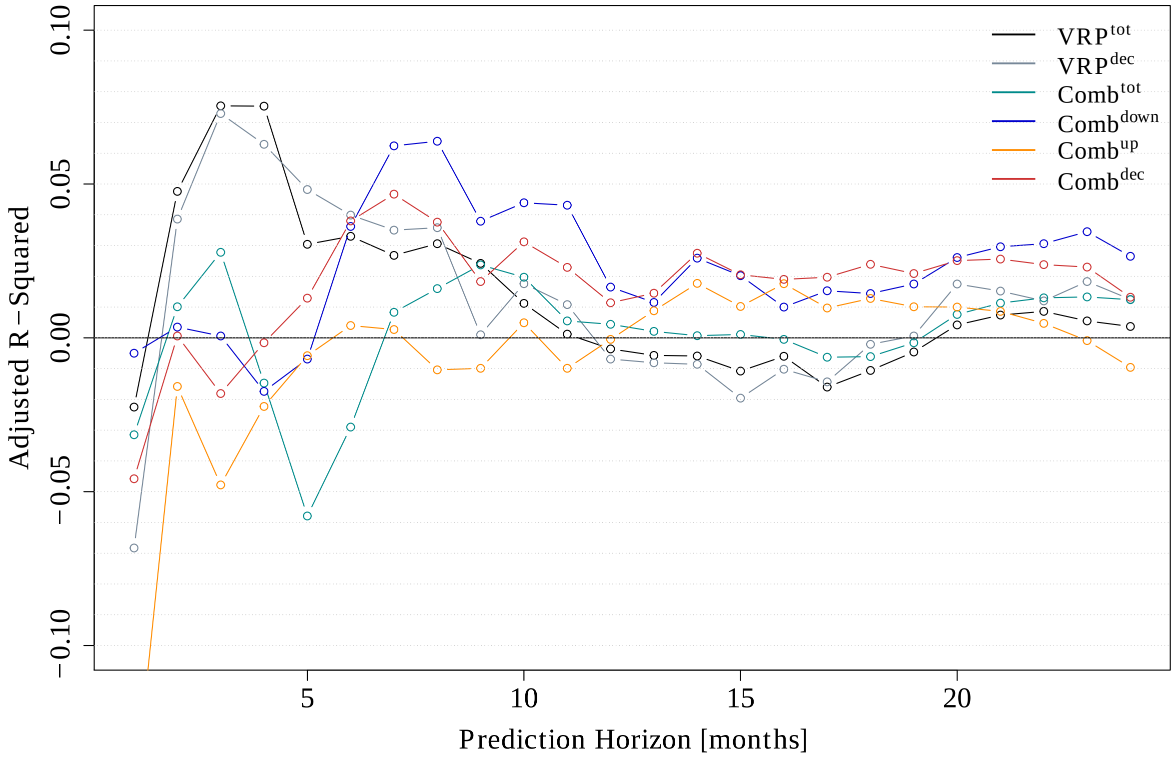
Note: This table shows the out-of-sample goodness-of-fit measure R^2_{OOS} for various prediction models at various forecasting horizons. The prediction models are the total variance risk premium (VRP_t^{tot}), a joint model that uses downside and upside variance risk premiums (VRP_t^{dec}), and four combination forecasts Comb_t that are constructed according to Equation (3.19). These combination forecasts comprise individual predictions of various decomposed higher-moment risk premiums. First, all total higher-moment risk premiums up to kurtosis are combined into Comb_t^{tot}. Second (Third), all downside (upside) higher-moment risk premiums up to kurtosis are combined into Comb_t^{down} (Comb_t^{up}). Fourth, a combination forecast is constructed that comprises all downside and upside premiums (Comb_t^{dec}). Panel A states the results of the out-of-sample study for OLS predictions that are compared with the historical mean market excess return as a benchmark predictor. Panel B shows results for LAD predictions that are evaluated in comparison with the historical median of market excess returns as a benchmark predictor. Statistical significance is assessed by the formal test of Clark and West (2007) and asterisks indicate significance of one-sided (upper-tail) t-tests at the 1%(***), 5%(**), and 10%(*) level, respectively.

At longer horizons, their out-of-sample predictive power deteriorates. VRP_t^{tot} outperforms the historical mean at six and nine months but afterwards only poorly predicts subsequent market excess returns. VRP_t^{dec} is also reasonably good at $h = 6$ but loses its out-of-sample predictive power at longer horizons. By contrast, the combination forecast of the total higher-moment risk premiums generates valuable predictions only at three-month and nine-month forecasting horizons.

Decomposed higher-moment risk premium combination forecasts provide mixed evidence of out-of-sample return predictability. The upside premium combination forecast Comb_t^{up} is a weak out-of-sample predictor and is often inferior to the naïve benchmark predictor, as most R^2_{OOS} are negative. This affirms the in-sample evidence where upside higher-moment risk premiums do not predict the market excess returns too well, either. By contrast, Comb_t^{down} is inferior to VRP_t^{tot} and VRP_t^{dec} at short horizons up to six months but subsequently has superior predictive power with R^2_{OOS} that are steadily above zero. This indicates a robust improvement compared to the historical mean predictor. The combination forecast Comb_t^{dec}, which incorporates all downside and upside higher-moment risk premiums, produces better results than out-of-sample variance risk premium prediction models as of a six-month forecasting horizon; however at most horizons it is inferior to Comb_t^{down}. This is noteworthy because it contradicts the in-sample evidence, where the most comprehensive prediction model HMRP_t^{dec} achieved the best predictions. Thus, the out-of-sample evidence indicates that this model's superior in-sample predictive power is due to an overfitted regression model. Figure 3.5, which is the out-of-sample equivalent to Figure 3.4, graphically compares the R^2_{OOS} of the discussed out-of-sample prediction models.

The more robust LAD regression methodology, which is shown in Panel B of Table 3.7, corroborates the OLS predictions' findings. However, four key findings should be stressed. First, variance risk premium prediction models, both total and decomposed, produce the best predictions at shorter forecasting horizons, particularly for subsequent three-month market excess returns. Second, upside higher-moment risk premiums forfeit their meager in-sample predictive power in the out-of-sample study. Third, Comb_t^{down} appears to be a better out-of-sample predictor than Comb_t^{tot} and produces valuable return predictions as of $h = 6$. Fourth, it seems reasonable to assume the existence of informational content in extreme observations of decomposed higher-moment risk premiums but that this information is limited to some degree. LAD regressions improve the out-of-sample predictability compared to OLS regressions as VRP_t^{dec} and Comb_t^{down} in particular appear to have much more prediction models with significant R^2_{OOS}. That is to say, OLS predictions are, to a certain extent, biased by influential observations that deteriorate a

Figure 3.5: R^2_{OOS} of OLS Predictive Regressions



Note: This figure depicts the R^2_{OOS} of various out-of-sample market excess return prediction models versus the historic mean as a benchmark predictor. Results are shown for predictions with the total variance risk premium (VRP^{tot}) and a joint prediction model with both decomposed variance risk premiums (VRP^{dec}). Moreover, the figure shows results for four combination forecasts based on individual higher-moment risk premium predictive models as shown in Equation 3.19 (Rapach et al. 2010). The first combination forecasts aggregates predictions of all total higher-moment risk premiums up to the fourth moment (Comb^{tot}), whereas the second combines predictions of all downside (Comb^{down}), and the third of all upside higher-moment risk premiums (Comb^{up}) up to the fourth moment. Lastly, out-of-sample goodness-of-fit measures for a model with all downside and upside higher-moment risk premiums up to kurtosis (Comb^{down}) are shown. Subject of the out-of-sample analysis are prediction horizons running from 1 to 24 months with an extending windows approach and an initial calibration sample of 84 months.

model's predictive power. Hence, putting less weight on extreme observations can increase the quality of predictions.

Overall, this section refines the in-sample evidence and emphasizes the relevance of decomposing higher-moment risk premiums when it comes to assessing their predictive power for future stock returns, particularly for longer-term return predictions. Especially (decomposed) variance risk premiums and downside combination forecasts turn out to be suitable out-of-sample stock return predictors. The evidence suggests that the best predictions for subsequent market excess returns at short horizons can be achieved with total and decomposed variance risk premiums as predictors. The employment of decomposed moment risk premiums is particularly reasonable for prediction horizons of six or nine months or longer. In this case, the results indicate that downside premiums are best suited as predictors for these forecasting horizons.

3.5 Conclusion

This paper examines decompositions of higher-moment risk premiums up to the fourth moment in a consistent framework, wherein premiums have a clear economic interpretation as profits from a trading strategy. To facilitate this analysis, I introduce a new measure of kurtosis risk premium and propose a simple method of decomposing these risk premiums. In doing so, this paper is the first to analyze decomposed skewness and kurtosis risk premiums in the equity index option market.

In a comprehensive empirical forecasting study, I present strong in-sample and out-of-sample evidence that information in decomposed higher-moment premiums can be exploited for predictions of subsequent market excess returns. Moreover, the results survive a series of robustness checks, and the predictors are shown to be capable of predicting future changes in macroeconomic conditions, which justifies their use as market return prediction signals. The empirical findings suggests that (decomposed) variance risk premium predictors tend to yield the best prediction results at short forecasting horizons of up to six or nine months, whereas downside higher-moment risk premiums in particular yield good predictions at longer prediction horizons. By contrast, upside higher-moment risk premiums do not provide exploitable information for subsequent market excess returns.

A Appendix

A.1 Kurtosis Swap Fixed Leg Replicating Portfolio

From Equation (3.3), define

$$H\left[F_{T,T}\right] = 12\left(\left(\ln\frac{F_{T,T}}{F_{t,T}}\right)^2 + 2\ln\frac{F_{T,T}}{F_{t,T}}\frac{F_{T,T}}{F_{t,T}} + 4\ln\frac{F_{T,T}}{F_{t,T}} - 6\frac{F_{T,T}}{F_{t,T}} + 6\right) \tag{A.1}$$

and calculate first and second derivatives with respect to $F_{T,T}$:

$$\begin{aligned} H'\left[F_{T,T}\right] &= 12\left(2\ln\frac{F_{T,T}}{F_{t,T}}\frac{1}{F_{T,T}} - 4\frac{1}{F_{t,T}} + 2\ln\frac{F_{T,T}}{F_{t,T}}\frac{1}{F_{t,T}} + 4\frac{1}{F_{T,T}}\right), \\ H''\left[F_{T,T}\right] &= 24\left(\frac{1}{F_{t,T}F_{T,T}} - \frac{1}{F_{T,T}^2} - \ln\frac{F_{T,T}}{F_{t,T}}\frac{1}{F_{T,T}^2}\right). \end{aligned} \tag{A.2}$$

Hence,

$$\begin{aligned} H\left[F_{t,T}\right] &= 0, \\ H'\left[F_{t,T}\right] &= 0, \\ H''\left[K\right] &= 24\left(\frac{1}{F_{t,T}K} - \frac{1}{K^2} - \frac{1}{K^2}\ln\frac{K}{F_{t,T}}\right) \\ &= 24\left(\frac{K - F_{t,T}\left(1 + \ln\frac{K}{F_{t,T}}\right)}{F_{t,T}K^2}\right). \end{aligned} \tag{A.3}$$

Applying the Bakshi and Madan (2000) and Carr and Madan (2001) spanning theorem yields

$$\begin{aligned} H[F_{T,T}] = 24\Bigg(&\int_0^{F_{t,T}} \frac{K - F_{t,T}\left(1 + \ln\frac{K}{F_{t,T}}\right)}{F_{t,T}K^2}(K-S)^+ dK \\ &+ \int_{F_{t,T}}^{\infty} \frac{K - F_{t,T}\left(1 + \ln\frac{K}{F_{t,T}}\right)}{F_{t,T}K^2}(S-K)^+ dK\Bigg). \end{aligned} \tag{A.4}$$

Taking expectations under the risk-neutral measure gives

$$\begin{aligned} G_{t,T}^{K} &= \mathbb{E}_t^{\mathbb{Q}}\left[12\left(\left(\ln\frac{F_{T,T}}{F_{t,T}}\right)^2 + 2\ln\frac{F_{T,T}}{F_{t,T}}\frac{F_{T,T}}{F_{t,T}} + 4\ln\frac{F_{T,T}}{F_{t,T}} - 6\frac{F_{T,T}}{F_{t,T}} + 6\right)\right] \\ &= \frac{24}{B_{t,T}}\Bigg\{\int_0^{F_{t,T}} \frac{K - F_{t,T}\left(1 + \ln\frac{K}{F_{t,T}}\right)}{F_{t,T}K^2} P_{t,T}(K) dK \\ &\quad + \int_{F_{t,T}}^{\infty} \frac{K - F_{t,T}\left(1 + \ln\frac{K}{F_{t,T}}\right)}{F_{t,T}K^2} C_{t,T}(K) dK\Bigg\}. \end{aligned} \tag{A.5}$$

A.2 Kurtosis Swap Floating Leg

The kurtosis swap g-swap and its derivatives with respect to $r_{t,T}$ are

$$\begin{aligned} g^{K}(r_{t,T}) &= 12\left(r_{t,T}^2 + 2r_{t,T}e^{r_{t,T}} + 4r_{t,T} - 6e^{r_{t,T}} + 6\right), \\ g^{K'}(r_{t,T}) &= 12\left(2r_{t,T} - 4e^{r_{t,T}} + 2r_{t,T}e^{r_{t,T}} + 4\right), \\ g^{K''}(r_{t,T}) &= 12\left(2 - 2e^{r_{t,T}} + 2r_{t,T}e^{r_{t,T}}\right). \end{aligned} \tag{A.6}$$

Forward prices of the derivatives are

$$\begin{aligned} G_{t,T}^{K'} &= \mathbb{E}_t^{\mathbb{Q}}\left[g^{K'}(r_{t,T})\right] = \mathbb{E}_t^{\mathbb{Q}}\left[12\left(2r_{t,T} - 4e^{r_{t,T}} + 2r_{t,T}e^{r_{t,T}} + 4\right)\right] \\ &= 12\mathbb{E}_t^{\mathbb{Q}}\left[2\left(r_{t,T}e^{r_{t,T}} - e^{r_{t,T}} + 1\right) - 2\left(e^{r_{t,T}} - 1 - r_{t,T}\right)\right] = 12G_{t,T}^{E} - 12G_{t,T}^{V}, \\ G_{t,T}^{K''} &= \mathbb{E}_t^{\mathbb{Q}}\left[g^{K''}(r_{t,T})\right] = \mathbb{E}_t^{\mathbb{Q}}\left[12\left(2 - 2e^{r_{t,T}} + 2r_{t,T}e^{r_{t,T}}\right)\right] = 12G_{t,T}^{E}, \end{aligned} \tag{A.7}$$

where $G_{t,T}^{E}$ is the forward price of the entropy contract introduced by Neuberger (2012). This gives the realized kurtosis for continuous rebalancing:

$$\begin{aligned} dY_{t,T}^{K} &= dG_{t,T}^{K'}df_{t,T} + \frac{1}{2}\left(G_{t,T}^{K''} - G_{t,T}^{K'}\right)\left(df_{t,T}\right)^2 \\ &= 12d\left(G_{t,T}^{E} - G_{t,T}^{V}\right)df_{t,T} + 6G_{t,T}^{V}\left(df_{t,T}\right)^2, \end{aligned} \tag{A.8}$$

and the realized kurtosis for discrete rebalancing:

$$\begin{aligned} \delta Y_{t,T}^{K} &= \mathbb{E}_{t+\delta t}^{\mathbb{Q}}\left[g^{K}\left(r_{t+\delta t,T} + \delta f_{t,T}\right) - g^{K}\left(r_{t+\delta t,T}\right) - \left(e^{\delta f_{t,T}} - 1\right)\mathbb{E}_t^{\mathbb{Q}}\left[g^{K'}\left(r_{t,T}\right)\right]\right] \\ &= \mathbb{E}_{t+\delta t}^{\mathbb{Q}}\big[12\big(\left(r_{t+\delta t,T} + \delta f_{t,T}\right)^2 + 2\left(r_{t+\delta t,T} + \delta f_{t,T}\right)e^{r_{t+\delta t,T}+\delta f_{t,T}} \\ &\qquad + 4\left(r_{t+\delta t,T} + \delta f_{t,T}\right) - 6e^{r_{t+\delta t,T}+\delta f_{t,T}} + 6\big)\big] \\ &- \mathbb{E}_{t+\delta t}^{\mathbb{Q}}\left[12\left(r_{t+\delta t,T}^2 + 2r_{t+\delta t,T}e^{r_{t+\delta t,T}} + 4r_{t+\delta t,T} - 6e^{r_{t+\delta t,T}} + 6\right)\right] \\ &- \left(e^{\delta f_{t,T}} - 1\right)G_{t,T}^{K'} \\ &= 12\left(\delta G_{t,T}^{E} - \delta G_{t,T}^{V}\right)\left(e^{\delta f_{t,T}} - 1\right) + 6G_{t+\delta t,T}^{V}\left(e^{\delta f_{t,T}} - 1\right)^2 \\ &+ 12\left(\delta f_{t,T}^2 + 2\delta f_{t,T}e^{\delta f_{t,T}} + 4\delta f_{t,T} - 6e^{\delta f_{t,T}} + 6\right). \end{aligned} \tag{A.9}$$

A.3 Decomposed Higher-Moment Risk Premium Formulas

Table A.1: Decomposed Higher-Moment Risk Premium Formulas

Panel A: Prerequisites

Entropy Contract	$G^E_{t,T} = \frac{2}{B_{t,T}} \left\{ \int_{F_{t,T}}^{\infty} \frac{C_{t,T}(K)}{KF_{t,T}} dK + \int_0^{F_{t,T}} \frac{P_{t,T}(K)}{KF_{t,T}} dK \right\}$
Entropy Implem.	$v^E_{t,T} = 2 \sum_{K_i \leq F_{t,T}} \frac{P_{t,T}(K_i)}{B_{t,T}F_{t,T}K_i} \Delta I(K_i)$
	$+ 2 \sum_{K_i \geq F_{t,T}} \frac{C_{t,T}(K_i)}{B_{t,T}F_{t,T}K_i} \Delta I(K_i)$

Panel B: Variance

g-function	$g^V(r_{t,T}) = 2(e^{r_{t,T}} - 1 - r_{t,T})$
Fixed Leg	$G^V_{t,T} = \frac{2}{B_{t,T}} \left\{ \int_0^{F_{t,T}} \frac{P_{t,T}(K)}{K^2} dK + \int_{F_{t,T}}^{\infty} \frac{C_{t,T}(K)}{K^2} dK \right\}$
Floating Leg	$\delta Y^V_{t,T} = 2\left(e^{\delta f_{t,T}} - 1 - \delta f_{t,T}\right)$
Risk Premium	$\text{VRP}^{tot}_t = \delta Y^V_{t-1,t} - G^V_{t,T}$
Fixed Leg Implem.	$v^L_{t,T} = 2 \sum_{K_i \leq F_{t,T}} \frac{P_{t,T}(K_i)}{B_{t,T}K_i^2} \Delta I(K_i) + 2 \sum_{K_i \geq F_{t,T}} \frac{C_{t,T}(K_i)}{B_{t,T}K_i^2} \Delta I(K_i)$
Floating Leg Implem.	$rv_{t,T} = \sum_{i=1}^{n-1} 2(e^{r_{i,i+1}} - 1 - r_{i,i+1})$
Risk Premium	$\text{VRP}^{tot}_t = rv_{t-1,t} - v^L_{t,T}$

Panel C: Skewness

g-function	$g^S(r_{t,T}) = 6(2 + r_{t,T} - 2e^{r_{t,T}} + r_{t,T}e^{r_{t,T}})$
Fixed Leg	$G^S_{t,T} = \frac{6}{B_{t,T}} \left\{ \int_{F_{t,T}}^{\infty} \frac{(K - F_{t,T})C_{t,T}(K)}{K^2 F_{t,T}} dK \right.$
	$\left. - \int_0^{F_{t,T}} \frac{(F_{t,T} - K)P_{t,T}(K)}{K^2 F_{t,T}} dK \right\}$
Floating Leg	$\delta Y^S_{t,T} = 3\delta G^E_{t,T}\left(e^{\delta f_{t,T}} - 1\right) + 6\left(2 - 2e^{\delta f_{t,T}} + \delta f_{t,T} + \delta f_{t,T} e^{\delta f_{t,T}}\right)$
Risk Premium	$\text{SRP}^{tot}_t = G^S_{t,T} - \delta Y^S_{t-1,t}$
Fixed Leg Implem.	$is_{t,T} = 6 \sum_{K_i > F_{t,T}} \frac{(K_i - F_{t,T})C_{t,T}(K_i)}{B_{t,T}K_i^2 F_{t,T}} \Delta I(K_i)$
	$- 6 \sum_{K_i \leq F_{t,T}} \frac{(F_{t,T} - K_i)P_{t,T}(K_i)}{B_{t,T}K_i^2 F_{t,T}} \Delta I(K_i)$
Floating Leg Implem.	$rs_{t,T} = \sum_{i=1}^{n-1} 3\delta v^E_{i,T}(e^{r_{i,i+1}} - 1)$
	$+ 6(2 - 2e^{r_{i,i+1}} + r_{i,i+1} + r_{i,i+1}e^{r_{i,i+1}})$
Risk Premium	$\text{SRP}^{tot}_t = is_{t,T} - rs_{t-1,t}$

Table A.1: continued

Panel D: Kurtosis

g-function	$g^K(r_{t,T}) = 6\left(e^{2r_{t,T}} - 2r_{t,T} - 5 + 4e^{r_{t,T}} - 4r_{t,T}e^{r_{t,T}}\right)$
Fixed Leg	$G^K_{t,T} = \dfrac{24}{B_{t,T}}\left\{\int_0^{F_{t,T}} \dfrac{K - F_{t,T}\left(1 + \ln \frac{K}{F_{t,T}}\right)}{F_{t,T}K^2} P_{t,T}(K)dK\right.$
	$\left. + \int_{F_{t,T}}^{\infty} \dfrac{K - F_{t,T}\left(1 + \ln \frac{K}{F_{t,T}}\right)}{F_{t,T}K^2} C_{t,T}(K)dK\right\}$
Floating Leg	$\delta Y^K_{t,T} = 12\left(\delta G^E_{t,T} - \delta G^V_{t,T}\right)\left(e^{\delta f_{t,T}} - 1\right) + 6G^V_{t+\delta t,T}\left(e^{\delta f_{t,T}} - 1\right)^2$
	$+ 12\left(\delta f^2_{t,T} + 2\delta f_{t,T}e^{\delta f_{t,T}} + 4\delta f_{t,T} - 6e^{\delta f_{t,T}} + 6\right)$
Risk Premium	$\text{KRP}^{tot}_t = \delta Y^K_{t-1,t} - G^K_{t,T}$
Fixed Leg Implem.	$ik_{t,T} = 24\sum_{K_i \leq F_{t,T}} \dfrac{K_i - F_{t,T}\left(1 + \ln \frac{K_i}{F_{t,T}}\right) P_{t,T}(K_i)}{B_{t,T}K_i^2 F_{t,T}} \Delta I(K_i)$
	$+ 24\sum_{K_i > F_{t,T}} \dfrac{K_i - F_{t,T}\left(1 + \ln \frac{K_i}{F_{t,T}}\right) C_{t,T}(K_i)}{B_{t,T}K_i^2 F_{t,T}} \Delta I(K_i)$
Floating Leg Implem.	$rk_{t,T} = \sum_{i=1}^{n-1} 12\left(\delta v^E_{i,T} - \delta v^V_{i,T}\right)\left(e^{r_{i,i+1}} - 1\right) + 6v^V_{i+1,T}\left(e^{r_{i,i+1}} - 1\right)^2$
	$+ \sum_{i=1}^{n-1} 12\left(r^2_{i,i+1} + 2r_{i,i+1}e^{r_{i,i+1}} + 4r_{i,i+1} - 6e^{r_{i,i+1}} + 6\right)$
Risk Premium	$\text{KRP}^{tot}_t = rk_{t-1,t} - ik_{t,T}$

4 How to Harvest Variance Risk Premiums for the Long-term Investor?

Joint work with Olaf Korn and Gabriel J. Power.

Abstract

Derivative positions that aim to earn variance risk premiums are exposed to sharp price declines during market crises, which calls into question their suitability for the long-term investor. Our paper systematically discusses the problems associated with the design of long-term variance-based investments. Various design elements are proposed to address some of these problems. An empirical study of investment strategies based on these design elements yields significant differences across strategies in terms of risk and return for the S&P 500 index options market. Overall, our results show that variance strategies can be attractive to the long-term investor if properly designed.

Acknowledgements: We thank Marco Erling, Florian Reibis, Jörg Zimmermann, as well as participants of the 2021 CFR research seminar for helpful comments and suggestions. Vitus Benson provided excellent research assistance.

4.1 Introduction

The variance risk premium is a well-known phenomenon in options markets. A large literature provides empirical evidence for the existence of a negative variance risk premium in equity index options, such as the S&P 500 options (e.g., Carr and Wu 2009; Kozhan et al. 2013). Moreover, the realized premium has a well-documented structure: it exhibits low negative values most of the time and takes on very high positive values during rare extreme events. However, little attention has been paid in the literature to the harvesting of the variance risk premium for the long-term investor. The structure of variance risk with its rare extreme events makes this a difficult—if not impossible—task. The goal of our work is to provide new insights into whether and how investors can use the variance risk premium for long-term capital accumulation. More specifically, this is the first study to systematically compare different variance-based investment strategies in terms of their suitability for long-term investors.

The paper makes conceptual and empirical contributions. First, strategies for earning the variance risk premium are distinguished from other variance strategies. Then, three problems that arise when variance risk premiums are to be exploited for the long-term investor are highlighted: (i) The payoff problem: Which payoff profiles are appropriate? Which instruments should be used to create them?; (ii) The leverage problem: Which risk level should be chosen? How can the ex-ante variance risk of different strategies be measured and compared?; (iii) The finite maturity problem: Which maturities of derivatives should be chosen? When and how often should positions be rolled over? Starting from the three problems, we propose different strategies for earning the variance risk premium that address potential solutions. The analysis of these strategies contributes to a better understanding of the specifics of the variance risk premium in the context of long-term investments.

In an empirical study for the S&P 500, we examine the various strategies in terms of their suitability for capital accumulation for the long-term investor. We use data from January 1996 to December 2020, so that the two most significant stock market crashes in recent decades—the 2008 financial crisis and the Covid-19 pandemic in early 2020—are included in the sample. This is particularly important for variance strategies. For the strategies, we use realistic assumptions about transaction costs and CBOE margin requirements for option positions. We use two variants of strategies to set the risk exposure: First, we use the maximum possible exposure by levering positions until capital is fully tied up. Second, we use an equal ex-ante factor exposure for all strategies. To compare exposure, we propose a model-free measure of nonlinearity (gamma) that can be determined from current option prices alone.

The empirical analysis provides the following main results: (i) Maximum exposure strategies differ greatly in terms of return and risk profile. In some cases, the strategies exhibit excessive risk and there is even a total loss of the invested capital; (ii) Equal exposure strategies still differ from each other, although significantly less than maximum exposure strategies. The remaining differences are due to the different instruments on which the strategies are based, resulting in different payoff profiles and transaction costs; (iii) All variance strategies have a significant positive correlation with the market, but this correlation differs across strategies due to different payoff profiles; (iv) Throughout the whole sample period, variance strategies have continuously earned premiums. This distinguishes them from other factor strategies based on Fama's and French's (2015) five factors or the momentum factor, which have not shown a significant upward trend since the 2008 financial crisis; (v) The variance factor, although correlated with the market, translates into an attractive factor strategy for long-term investors, both as a stand-alone factor and as a complement to a market investment. Overall, our study shows that although variance strategies exhibit extreme distributions with high negative skewness and excess kurtosis, they recover quite quickly from large drawdowns and continue to consistently earn premiums. The results suggest that these strategies, when properly designed, are attractive to long-term investors.

4.2 Literature Overview

The objective of this paper is to seek effective ways to harvest the variance risk premium for the long-term investor. Formally defined, the variance risk premium is the deviation between the expected (physical) variance realized over the life of the options and the option-implied (risk-neutral) variance. More loosely speaking, the variance risk premium is the expected return on the "variance factor". Although this variance factor is far from uniformly defined in the literature, approaches to obtaining its factor premium have in common that they provide exposure to variance changes. Nevertheless, these approaches can differ greatly from one another. In the following, we briefly review the literature on strategies aimed at earning variance premiums and distinguish our paper from this literature.

A first approach looks at trading strategies that hold options in combination with their underlying, usually a broad market index. Since options have non-linear payoffs, they naturally provide exposure to changes in variance. Whaley (2002) describes and analyzes the Chicago Board Options Exchange (CBOE) BuyWrite Index, which consists of a short-position in an out-of-the money (OTM) S&P 500 index call and a long-position in the S&P 500. Ungar and Moran (2009) outline a put-write strategy, which is the basis for the CBOE S&P 500 PutWrite Index. These strategies have in common that they led to an outperformance of the S&P 500 on a risk-adjusted basis. Other strategies, such as PutWrite and BuyWrite strategies with varying moneyness of the options or with additional caps and floors are assessed by Clark and Dickson (2019). However, all these strategies represent a portfolio approach rather than a pure variance strategy, as they are not delta-hedged but in most cases fully collateralized with a position in the index. This leads to a mixture of market exposure and variance exposure, as outlined in Israelov and Nielsen (2015). The approach of our paper is to study the properties of different pure variance strategies first and then to relate the performance of these strategies to the underlying and to other equity-based factors.

Another strand of literature utilizes variance exposure as an additional element in a broader portfolio context and examines its impact on overall portfolio performance. Brière et al. (2010) explore long-term equity and bond portfolio strategies and add short-positions in variance swaps to include variance exposure. They find that the additional variance exposure boosted portfolio returns but added little diversification benefits. Fallon et al. (2015) employ a similar approach and investigate the role of additional volatility risk in institutional investment portfolios. They find that adding small amounts of exposure to variance risk substantially enhanced long-term returns at the cost of increased short-term tail risk in their sample. Other studies, such as Ge (2016a) and Ge (2017) likewise examine the effect of variance exposure on portfolio returns and portfolio performance in crisis periods. Overall, this literature provides interesting results on the effects of variance exposure in an asset allocation context. However, it does not analyze the trade-offs between alternative design elements of the “variance factor”. Moreover, it does not explore the properties of variance strategies themselves but looks only on their impact on specific portfolios.

A different idea to earn premiums associated with variance exposure is to engage into VIX trading strategies. These strategies are typically built on VIX futures (Simon and Campasano 2014) or VIX options (Simon 2017) and aim to roll down the term-structure of option-implied volatility. As the term-structure of implied volatility is typically in contango, investors can enter into a short-position in longer-term futures to roll down the term-structure of VIX futures and, on average, earn the corresponding premium (Ge 2016b). However, in our understanding, this premium is not as much a variance risk premium as it is a VIX term-structure premium. In particular, these VIX strategies earn money via changes in implied (risk-neutral) volatility between two points in time and not via the deviation between implied and expected realized (physical) volatility, i.e., the variance risk premium. Of course, changes in implied volatility might well be correlated with realized volatility. Nevertheless, this type of strategy earns a premium which is conceptually distinct. In our paper, we concentrate on strategies that aim to harvest the variance risk premium.

Lastly, there are few papers that examine strategies aimed to earn pure variance risk premiums. Fallon and Park (2016) analyze a strategy that sells synthetically derived capped one-month S&P 500 variance swaps. Specifically, they employ a stochastic volatility model with jumps (SVJ) to estimate the swap rates. The authors find the corresponding strategy to have very high Sharpe ratios and severe but infrequent crash risk. However, Fallon and Park (2016) do not investigate the long-term performance of such a strategy and the impact of rare crash events on the capital accumulation of the long-term investor. Ge (2016b) comes closest to our approach by examining the long-term performance of, among others, a variance swap investment strategy. He finds that this strategy provides decent returns but a very large maximum drawdown. However, the comparison of different strategies in Ge's study is based on an ex-post standardization of their risks. This is not feasible for real investments. Our paper, in contrast, uses two different ways to determine a meaningful leverage of different strategies on an ex-ante basis. Moreover, our study is the first to include the period of the Covid-19 pandemic. This period and the corresponding stock market shock is potentially very important for the understanding of long-term variance strategies.

4.3 Strategies, Data, and Study Design

4.3.1 Three Problems

The design of strategies that seek to harvest variance risk premiums for the long-term investor entails three major problems. The first one is what we call the *payoff problem*. In principle, variance exposure can be generated by selling any convex payoff structure. However, different payoff profiles may be more or less suitable to achieve certain desirable goals. First, a payoff profile should provide sufficient factor exposure that is priced in the market. Second, a payoff profile should produce low correlations with other risk factors to create potential for diversification benefits in a portfolio context. Third, a payoff profile should limit extreme negative returns in order to avoid large drawdowns. Fourth, a payoff profile should be implementable with low transaction costs. This is crucial for long-term investors, as they profit from compound interest throughout the investment period. There are certainly trade-offs between the achievement of these four goals. These trade-offs are at the heart of the payoff problem, i.e. the problem to choose suitable payoff profiles.

The second problem is what we call the *leverage problem.* Variance strategies use derivatives to sell convex payoff profiles. Since some derivative strategies require no initial capital (e.g., swap contracts) and others actually generate capital (e.g., delta-hedged puts), the question of the appropriate amount of leverage arises. The leverage problem can lead to some similar trade-offs as the payoff problem. On the one hand, strategies should provide sufficient exposure to variance risk; on the other hand, they should limit extreme losses that call into question the long-term success of a strategy—all in a cost-effective manner. However, the leverage problem also raises issues that are distinct from the payoff problem. There may be different strategies with similar or even identical payoffs that use different instruments. Therefore, potential leverage may be more constrained for one strategy than another due to differences in initial capital or margin requirements. In addition, differences in leverage between strategies can lead to sharply different factor exposures, making meaningful comparisons difficult.

Third, there is the *finite maturity problem.* Because of the finite maturities of derivatives, long-term investors must periodically roll their positions. This entails further considerations. In particular, investors must decide which maturities to choose, when to roll, and how many instruments to use simultaneously. One of the potential trade-offs is that longer-term contracts may have less factor exposure but need to be traded less frequently, reducing transaction costs. Shorter-term contracts, however, may be more liquid, resulting in lower transaction costs per trade.

4.3.2 Variance Strategies

We now present strategies that are in principle suited for harvesting variance risk premiums and address some of the trade-offs associated with the three problems. We begin with a short position in an at-the-money (ATM) *straddle* as a first intuitive approach to selling protection against rising variances. The short straddle limits market exposure because it combines a short call (with negative delta) with a short put (with positive delta).[29] Moreover, it consists of only two ATM instruments, which means that transaction costs are limited. The stylized payoff profile of a short straddle is shown in part (a) of Figure 4.1. It highlights that this payoff is generally consistent with an instrument showing variance exposure. However, large negative or positive

[29]One could even set the strike price of the options as such that the beta of the straddle is exactly zero.

price movements of the underlying would result in large negative returns. Even medium price movements already lead to losses.

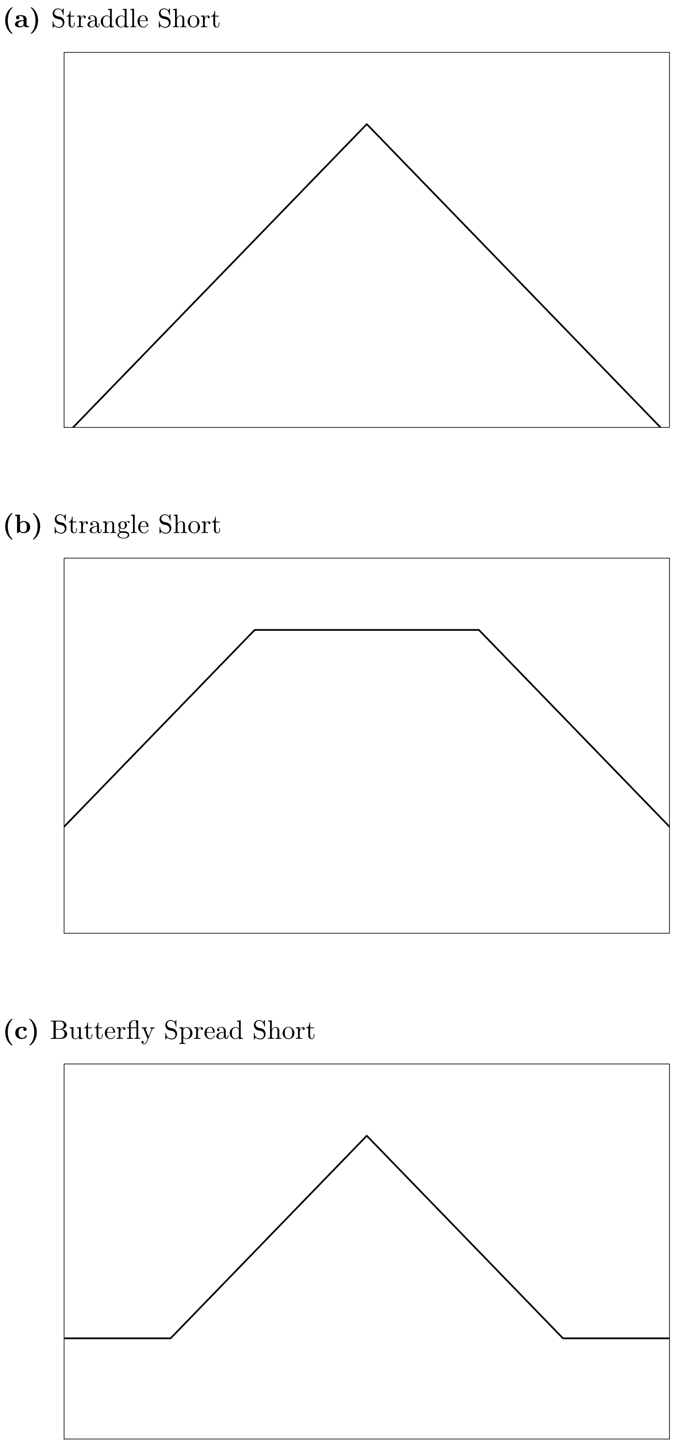
Alternatively, one could shift the strike prices of the options from ATM to out-of-the-money (OTM) to avoid losses in case of medium price changes of the underlying. In addition, this shift in strike prices reduces the magnitude of losses in case of large price moves of the underlying, as compared to the straddle. A corresponding payoff profile is depicted in part (b) of Figure 4.1, which shows a short *strangle*. However, the strangle also has some disadvantages. First, OTM options have less variance exposure than ATM options. Second, the maximum payoff is lower than for a corresponding straddle. Finally, large negative returns can still occur since the payoff function has no lower bound.

To counter the problem of extreme losses, one could add a floor to the payoff profile. By adding a long OTM call option and a long OTM put option, the straddle turns into a *butterfly spread*, as shown in part (c) of the Figure 4.1. Of course, adding a floor is also possible for the strangle. Such a portfolio is called a *condor strangle*. However, limiting the downside risk also comes with drawbacks. Since two long positions in options enter the portfolio together with short positions, the overall variance exposure of the portfolio is reduced. Moreover, the additional long positions in calls and puts incur costs—both transaction costs and costs of capital for the option premiums.

Another approach to harvesting the variance risk premium is to sell *delta-hedged call or put* options, i.e., puts or calls hedged with positions in the market index. Delta-hedged options have a similar payoff structure as a straddle. Therefore, they have similar advantages, i.e., limited correlation with the market factor and only two instruments making up the portfolio, but also similar disadvantages, i.e., potentially large downside risk. However, since delta-hedged option portfolios contain different instruments than a straddle, the long-term performance of delta-hedged options may differ significantly from the long-term performance of a straddle strategy. This may be due to different potential for leverage or different transaction costs and margin requirements.

Finally, an investor can gain variance exposure through *variance swaps*. The advantage of a variance swap is that it is the most direct way to earn the variance risk premium while limiting correlation with the market factor, since variance swaps are market neutral by construction. However, a variance swap can be viewed as a fairly complex option portfolio with potentially high transaction costs and leverage constraints. Moreover, since variance is calculated as the

Figure 4.1: Stylized Payoffs: Straddle, Strangle, and Butterfly Spread



Note: This figure shows stylized payoffs of a short position in a straddle (part (a)), a short position in a strangle (part (b)) and a short position in a butterfly spread (part (c)).

squared difference from the mean, a variance swap could be prone to extreme losses if the realized variance reaches peaks.

Overall, we have selected seven different but related portfolios: Straddle, strangle, butterfly spread, condor strangle, delta-hedged call, delta-hedged put, and variance swaps. Since all of them offer convex payoff structures, selling these portfolios has the potential to earn variance risk premiums. In addition, differences in their designs are particularly targeted to provide insights on potential trade-offs with respect to variance exposure, downside risk, transaction costs, capital requirements, and margins.

Having introduced different portfolios and their payoff profiles, the next question is what quantities should be held in each portfolio to obtain a good risk-return profile for the long-term investor. We follow two ways to approach this issue in our study. The first is the *maximum exposure* approach. Here, the investor generates the maximum possible exposure for each strategy by levering portfolio positions until her capital is fully invested in long positions and margins for short positions. In essence, the maximum exposure approach is a full invest of the available capital in the variance strategy. It provides information on the impact of different instruments on potential leverage. The second way is the *equal exposure* approach. Here, all strategies are levered until each strategy has the same ex-ante factor exposure. If one strategy requires less capital to achieve this exposure than another, the additional funds are invested in a risk-free account. Equal exposure strategies help us compare different strategies on an ex-ante risk basis.

Equal exposure strategies require an ex-ante measure of factor exposure. The literature suggests different measures associated with the non-linearity of options, in particular vega and gamma (Cremers et al. 2015). However, vega in general depends on the valuation model used, and specifically, the Black-Scholes vega measures variance risk in a setting where no variance risk premium exists. Consequently, using vega creates potential problems in the context of our study. Therefore, we use the gamma of the option as a general measure of non-linearity or convexity. Since we want to avoid model dependence and, in particular, a measure that builds on a model where variance risk premiums do not even exist, we apply a model-free approach. We determine gamma and, for the sake of consistency, also delta in this way. The derivation of such model-free greeks is based on the option-implied risk-neutral return (RND) distribution and is discussed in more detail in the next subsection.

4.3.3 Data, RNDs and Greeks

In our empirical study, we use European S&P 500 index options data which stems from the OptionMetrics IvyDB U.S. data base. The data base provides historical closing quotes from the CBOE, as well as implied volatilities, interest rates, spot prices of the underlying, and implied dividend yields. Our sample period starts in January 1996 and ends in December 2020, that is, it spans 299 months. We filter our option data set with standard filters from the literature (Goyal and Saretto 2009; Cao and Han 2013). We require best bid and best ask quotes as well as the bid-ask-spread to be non-negative and discard options with special settlement and options that do not have a.m. settlement. Moreover, the implied volatility must be greater than zero and we require options to survive standard no-arbitrage conditions. For every month in our data period, we only use data of the first trading day after the third Friday of that respective month. The latter is the standard expiration date of options traded at the CBOE. Additionally, we require the options to mature in the next month. This retains options with approximately one month to maturity. Our final sample consists of 23,231 call options and 23,697 put options for 299 months with an average time to maturity of approximately 28 days.

Model-free deltas and gammas require an estimate of the risk-neutral distribution of the S&P 500 index. We determine the RND based on the approach outlined in Figlewski (2010). At every point in time we select out-of-the money forward put and call options that survive our filter criteria and convert their midquotes to the corresponding Black and Scholes (1973) implied volatilities. As put and call options might trade at slightly different implied volatilities, we smooth out potential jumps in implied volatilities at the transition point from put options to call options. In an interval of 2.5% around the at-the-money forward price we apply the blending approach suggested in Figlewski (2010) to achieve a smoothed weighted-average implied volatility.[30] Next, we perform a quartic spline to interpolate the implied volatility curve across a broad range of observed strike prices. We then compute a fine grid of 12,000 equally spaced strike prices on the interval $[0.001, 3 \cdot X_t]$, where X_t is the current index level, and determine their corresponding implied volatilites. These implied volatilities are then converted back into put option prices, which we use to estimate the risk-neutral density by approximating derivatives of the put price function with respect to the strike price according to Breeden and Litzenberger (1978). This procedure is only valid between the lowest and the highest available strike prices. For strike

[30]For this blending approach we have to employ in-the-money (ITM) options that are only used to smooth the implied volatility curve and discarded afterwards.

prices beyond the observed range we estimate a generalized extreme value (GEV) distribution to extend the risk-neutral density to the left and right tail.[31]

Given the RND, we compute model-free deltas and gammas according to a simple comparative static analysis. First, we assume that the entire RND is shifted upward by an infinitesimally small amount ε. Next, we reprice options under this shifted RND and observe the price change. Likewise, we consider a downward shift of the entire RND and perform the same repricing procedure. Lastly, we determine the average of these two price changes, divide it by ε and use this measure as the model-free delta. The same idea is applied to gamma, which is the sensitivity of delta with respect to changes in the price of the underlying. Formally, this analysis can be found in greater detail in Appendix B.1. Given the assumptions we make for the derivation, model-free deltas of long call options are simply the probability mass to the right of the strike price, whereas long put options' model-free deltas are the negative value of the probability mass to the left of the strike price. Model-free gammas are simply the probability mass "at" the strike price for both long call and long put options.

4.3.4 Implementation of Strategies

General Specifications We now turn to the concrete design of the trading strategies, which are based on the portfolios we presented in Subsection 4.3.2. As we use S&P 500 options with a maturity of approximately one month and hold them until expiration, we have to open new positions every month. Positions are established every Monday after the third Friday of the month. We index each strategy to a level of 100 as of January 1996 and then determine the cumulative wealth through the end of our data period in December 2020.

In terms of transaction costs, we use the ask price for option purchases (long positions) and the bid price for option sales (short positions). However, it is likely that institutional investors can trade at terms better than the quoted spread (Mayhew 2002; De Fontnouvelle et al. 2003), which is why we choose an effective spread of 25% of the quoted spread as our baseline scenario.[32] This is an even more conservative approach than that used by other papers examining option trading strategies—Goyal and Saretto (2009) and Cao and Han (2013) even assume an effective spread of 0% as the base case, i.e., trading is possible at the midquote. Transaction costs are incurred each time an option or the underlying is bought or sold. However, cash-settled options

[31] This approach is described in more detail in Figlewski (2010).

[32] In practice, institutional investors also use limit order strategies to reduce transaction costs below the quoted spread.

do not incur additional transaction costs at expiration, which is why we choose to let options expire. We assume that transactions in the underlying, which are necessary for delta hedging, are possible at a quoted bid-ask spread of 3 basis points.

In addition, we require margin accounts to collateralize the positions. For option trades, we apply the CBOE margin rules, which provide detailed guidelines for individual options and option portfolios. Cao et al. (2021), for example, also use the CBOE margin rules for their option trading strategies. An overview of the CBOE margin requirements we use in our study is provided in Appendix B.2. We use the margin rules for initial margins and assume that there are no maintenance margins or other adjustments required within the trading month. For short positions in the underlying index, we assume that 150% of the short sale proceeds must be deposited, which is equivalent to the Federal Reserve Board's "Regulation T". Finally, the margin requirements for variance swaps, which require no capital at inception because they are priced so that their initial value is zero, are set in accordance with the "margin requirements for non-centrally cleared derivatives" of the "Basel Committee on Banking Supervision" (BCBS). The BCBS requires that 15% of the variance notional be deposited as initial margin. The variance notional is the notional amount by which the difference between the floating leg and the fixed leg is multiplied. We assume that margin accounts bear interest at the risk-free rate.

For maximum exposure strategies, we calculate the maximum possible exposure by levering the positions until the capital is fully used to buy long positions and provide margins on short positions. For equal exposure strategies, we lever the positions until each strategy has the same factor exposure. In each period, we select the strategy that has the lowest gamma when it is fully levered according to the maximum exposure approach. All other strategies are then scaled down to this exposure and the remaining capital is invested at the risk-free rate.

Straddle and Butterfly Spread For the straddle strategy, we choose both a call option and a put option with strike prices closest to the ATM forward point. A straddle has its own margin rules according to the CBOE. For the butterfly spread, we add two options to the straddle. We choose an additional call option with a model-free delta of 0.05 and an additional put option with a model-free delta of -0.05. Both options enter the straddle as long positions. The choice of deltas is inspired by the CBOE S&P 500 Iron Butterfly Index[33], which is a hypothetical option trading strategy calculated by the CBOE that sells butterfly spreads. The margins for

[33] `https://www.cboe.com/us/indices/dashboard/BFLY/`.

the butterfly spread are determined as the margins of a short call spread plus the margins of a short put spread, as there are no separate rules for butterfly spreads. Both the straddle and the butterfly spread may have a residual index exposure that we delta-hedge with an index position at initiation. At maturity, the option positions are cash-settled and the hedge positions in the index are closed, which is true for all strategies we consider in our study.

Strangle and Condor Strangle The way we specify the strangle and condor strangle strategies is very similar to the straddle and butterfly spread strategies. We choose a call option and a put option with expiration next month that have model-free deltas of 0.20 and -0.20, respectively. Selling these two options results in a strangle. For the condor strangle, we again add two options. Consistent with the butterfly spread, we choose an additional call option with a model-free delta of 0.05 and an additional put option with a model-free delta of -0.05. This choice is similar to the CBOE S&P 500 Iron Condor Index[34], which is a hypothetical options trading strategy from the CBOE selling condor strangles. The CBOE provides its own margin rules for the strangle, while for the condor strangle we combine the margins of a short call spread and the margins of a short put spread. Again, any remaining delta of the strategies is hedged via index positions.

Delta-Hedged Call and Delta-Hedged Put For the delta-hedged call and put strategies, we select the call and put options with expiration in the next month that are closest to the ATM forward point. The strategies sell these options at the best bid (taking into account the effective spread) and hedge the resulting delta exposure with a position in the underlying index. Thus, the short call requires a long position in the underlying, while the short put requires a short position in the underlying, for which additional margins must be deposited. To keep transaction costs low, the delta hedge is set up at initiation and is not readjusted until maturity.

Variance Swap The variance swap strategy requires the calculation of the variance swap rate. This swap rate is determined according to Kozhan et al. (2013), which leads to consistent pricing based on the same options for all strategies in our study. We select all OTM forward call and put options to calculate swap rates. Transaction costs are accounted for by using bid quotes instead of midquotes to calculate variance swap rates. The strategy sells variance swaps at the variance swap rate and holds this position until maturity. By design, variance swaps are delta-neutral

[34] `https://www.cboe.com/us/indices/dashboard/CNDR/`.

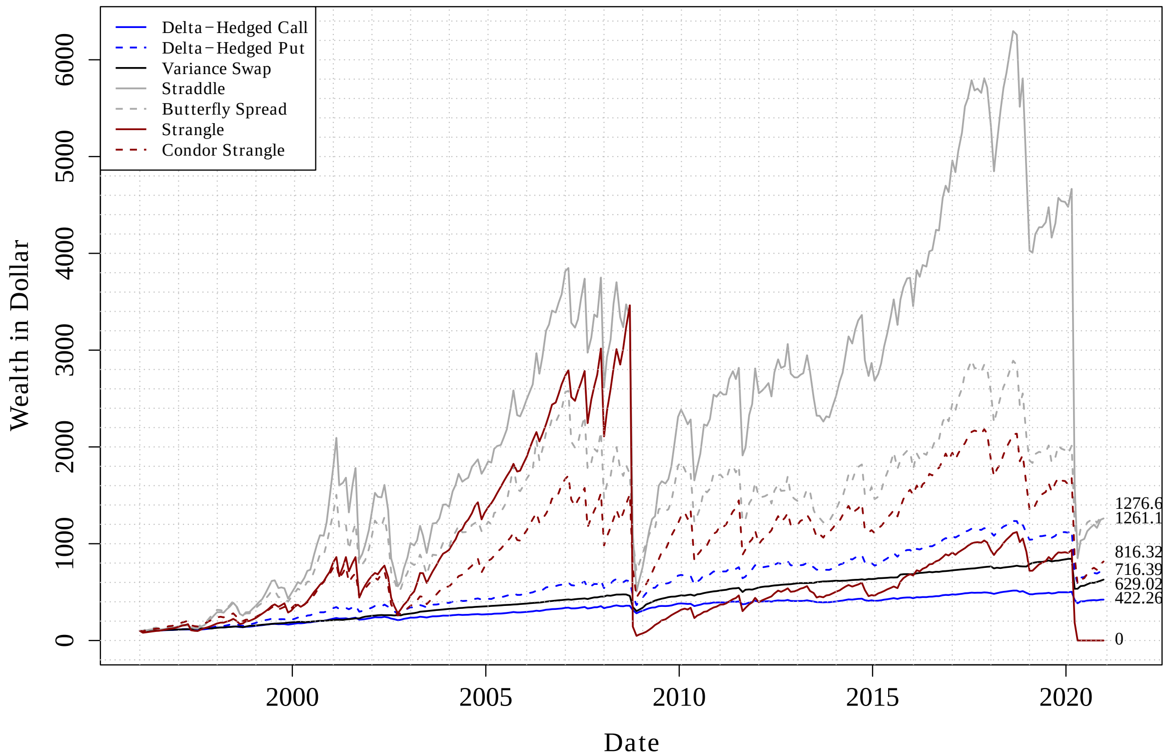
at inception. At maturity, the realized variance over the last month is taken to determine the variance swap's payoff.

4.4 Empirical Results

4.4.1 Maximum Exposure Strategies

The first set of results relates to maximum exposure strategies and aims at a better understanding of the leverage problem. Maximum exposure strategies show whether the capital requirements of an initial investment (if any) plus margins already set reasonable leverage constraints for strategies and lead to desirable risk-return patterns over longer horizons. If not, the question is whether the constraints are already too tight to earn significant premiums or, on the contrary, too loose to protect long-term investors from excessive risk.

Figure 4.2: Maximum Exposure Strategies: Cumulative Wealth



Note: This figure shows the cumulative wealth development of seven maximum exposure variance strategies for an initial investment of $100. The data period covers January 1996 to December 2020.

Figure 4.2 provides the accumulated wealth over time as created by different strategies. For each strategy, the initial budget in January 1996 is $100. The figure shows large differences among strategies. One group of extremely risky strategies consists of the straddle and the strangle.

These strategies reach very high wealth levels by 2007, but collapse during the financial crisis in October 2008. While the straddle recovers quickly from this shock, the strangle barely survives and does not have enough capital to recover quickly. It is eventually hit by the market turmoil due to the Covid-19 pandemic and is forced into bankruptcy in March 2020. The straddle also takes a massive hit from the pandemic. The butterfly spread and the condor strangle seem to dampen the extremes of the straddle and the strangle somewhat. However, they are still very risky and are also hit hard by the variance shocks of the financial crisis and the Covid-19 pandemic. Another interesting observation is the different behavior of the delta-hedged put and the delta-hedged call. The delta-hedged put fluctuates more and shows a steeper trend. In contrast, the delta-hedged call has a very smooth path, but does not seem to generate enough exposure for a significant upward movement, especially after the financial crisis. Finally, the variance swap combines a very smooth path, except for the most extreme months of the financial crisis and the Covid-19 pandemic, with a clear upward trend.

Table 4.1 characterizes the performance of the different strategies via summary statistics. Panel A provides information on the sample moments of monthly returns and Panel B shows information on downside risk. The first three measures of downside risk (VaR, CVaR, Max Loss) still take a monthly perspective and refer to (potential) losses in the following month. This is sufficient for a monthly investment horizon. However, for the long-term investor, the characteristics of the entire path are also important. In particular, the ability of a strategy to recover from an intermediate downturn is crucial. Therefore, Panel B offers four different drawdown statistics. The maximum drawdown (Max DD) is the maximum percentage loss of a strategy from its current maximum value to a trough; and the average drawdown (Average DD) shows how far, on average over all months of the 25-year period, a strategy is from its previous maximum. Drawdown length indicates how many months it takes to reach a new wealth maximum at a given point in time. For this measure, we report the maximum number of months (Max DD Length) and the average (over all months of the 25-year period) number of months (Average DD Length). For completeness, Panel C of Table 4.1 repeats the final wealth levels at the end of the data period from Figure 4.2 and converts them to annual geometric average returns.

Several differences between the strategies are evident in the figures in Table 4.1. First, when comparing straddle and strangle, straddle returns have a higher mean, a lower standard deviation, are less skewed to the left, and have a lower kurtosis. In terms of downside risk, the straddle is less risky than the strangle by all measures of risk. Therefore, the idea of replacing a straddle with a strangle to achieve risk reduction, especially with respect to large downturns, does not

Table 4.1: Return- and Risk-Statistics of Maximum Exposure Strategies

Panel A: Basic Monthly Summary Statistics

	Delta-Hedged Call	Delta-Hedged Put	Variance Swap	Straddle	Butterfly Spread	Strangle	Condor Strangle
Mean	0.0051	0.0080	0.0067	0.0182	0.0140	0.0157	0.0129
Standard Dev.	0.0243	0.0494	0.0309	0.1215	0.0986	0.1395	0.0975
Skewness	−2.6440	−2.9729	−8.1252	−2.2880	−1.4947	−4.0907	−2.7520
Exc. Kurtosis	12.9533	17.1776	88.7758	9.4223	3.6283	23.1548	8.6940
Sharpe Ratio	0.1266	0.1195	0.1516	0.1330	0.1208	0.0976	0.1112

Panel B: Downside Risk Statistics

	Delta-Hedged Call	Delta-Hedged Put	Variance Swap	Straddle	Butterfly Spread	Strangle	Condor Strangle
VaR (95%)	−0.0302	−0.0629	−0.0076	−0.1716	−0.1767	−0.1949	−0.1900
CVaR (95%)	−0.0697	−0.1448	−0.0700	−0.3632	−0.2832	−0.4448	−0.3228
Max Loss	−0.1670	−0.3860	−0.3676	−0.6729	−0.4106	−1.0000	−0.5085
Average DD	−0.0372	−0.0755	−0.0433	−0.1934	−0.2295	−0.2733	−0.2006
Max DD	−0.2550	−0.5266	−0.3676	−0.8684	−0.7238	−1.0000	−0.7384
Average DD Length	5.2683	5.7000	4.3750	12.8000	19.0714	18.8462	14.5882
Max DD Length	28	29	23	110	122	148	112

Panel C: Annualized Geometric Mean Return and Terminal Wealth

	Delta-Hedged Call	Delta-Hedged Put	Variance Swap	Straddle	Butterfly Spread	Strangle	Condor Strangle
Ann. Geom. Return [%]	5.95	8.22	7.66	10.71	10.76	−100.00	8.79
Term. Wealth [$]	422.26	716.39	629.02	1261.13	1276.63	0.00	816.32

Note: This table provides return- and risk-statistics of seven maximum exposure variance strategies. Panel A depicts basic summary statistics of monthly returns, while Panel B is dedicated to downside risk statistics. In particular, it considers asymmetric risk metrics for monthly returns such as the 95% value-at-risk (VaR), the 95% conditional value-at-risk (CVar), and the maximum loss (max loss). Additionally, it shows path-dependent drawdown measures: The average drawdown (Average DD), the maximum drawdown (Max DD), the average drawdown length (Average DD length), and the maximum drawdown length (Max DD Length). Panel C shows the annualized geometric return and the terminal wealth of the strategies.

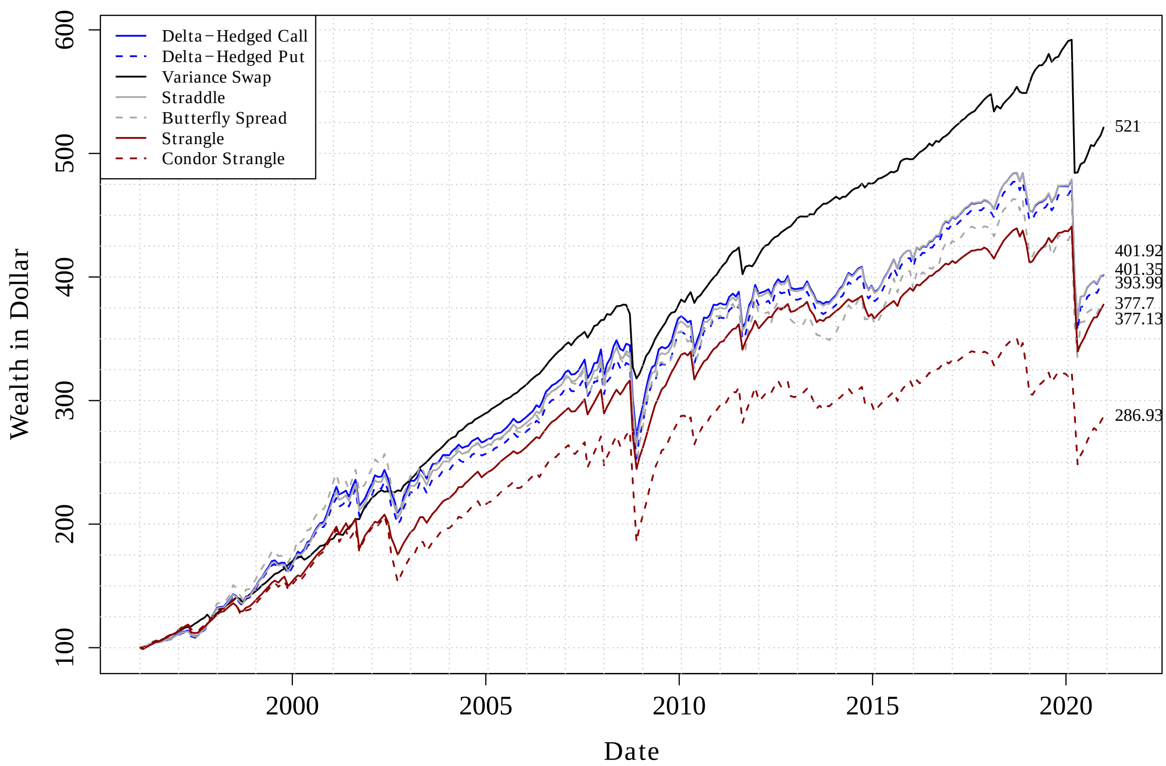
work, at least when taking maximum exposure. Both straddle and strangle strategies are highly speculative, with high mean returns but massive downside risk. Second, the comparison between straddle and butterly spread and between strangle and condor strangle provides insight into the impact of floors imposed via long positions in OTM calls and puts. These floors result in less negative skewness and less kurtosis, but they are also costly in terms of mean returns. Because of the latter effect, it is questionable whether such floors are an effective way to reduce risk. In terms of downside risk, the butterfly spread and the condor strangle are much riskier than the delta-hedged put, the delta-hedged call, and the variance swap. Third, the monthly returns of the variance swap show by far the most negative skewness and the highest kurtosis of all variance strategies. However, these characteristics do not necessarily imply high downside risk. The reason is that the standard deviation and the dynamics of the path are also important. Although the variance swap experiences some heavy losses (the maximum monthly loss is almost 37%), the relatively short drawdown length shows that the strategy can recover quite well from such losses.

In summary, there are clear differences between the various maximum exposure strategies. The delta-hedged call seems to generate relatively low factor exposure, even compared to the delta-hedged put. Other strategies create large exposure and massive risks (straddle, strangle); still others seem to have variance exposure that already offers promising risk-return profiles (delta-hedged put, variance swap) and relatively quick recovery from setbacks. In factor investing, however, investors seek to manage and control exposure. They look for a strategy that most efficiently exploits a given exposure, which can be studied using equal exposure strategies.

4.4.2 Equal Exposure Strategies

Figure 4.3 and Table 4.2 show the wealth accumulation and summary statistics of different equal exposure strategies. As expected, the standardization of the strategies in terms of ex-ante gamma leads to a more homogeneous picture. In terms of differences and similarities among the strategies and the benefits of particular design elements, the following observations are most striking. First, there is no clear evidence that a strangle helps reduce the risk of a straddle. Strangle returns are even more left-skewed and leptokurtic. In terms of downside risk, the different measures point in different directions. The straddle has a lower (less negative) VaR and CVaR, but a higher maximum loss. The maximum drawdown and maximum drawdown length are lower for the strangle, but the average drawdown and average drawdown length are lower for the straddle. Second, the floors introduced by the butterfly spread and the condor strangle improve monthly skewness and kurtosis, but are not effective in reducing downside risk, as a comparison with straddle and strangle shows. Third, delta-hedged put, delta-hedged call, and straddle have very similar paths and distributional properties. This finding implies that the different behavior of delta-hedged calls, delta-hedged puts, and straddles under maximum exposure strategies is due to their different potential for leverage. The delta-hedged call strategy requires a long position in the index to hedge the short call, which consumes a lot of capital. In contrast, the delta-hedged put strategy involves shorting the index. This short position induces margin requirements, but the potential leverage of a delta-hedged put is much greater compared to the call. A (delta-hedged) straddle requires a very small investment (long or short) in the index. In addition, the CBOE's margin requirements for a short straddle are much lower than the sum of the margin requirements for a short call and a short put. Therefore, straddles can achieve much higher gammas than the other strategies. Fourth and finally, the variance swap has the highest return of all the variance strategies while showing the lowest downside risk for the majority of the risk measures.

Figure 4.3: Equal Exposure Strategies: Cumulative Wealth



Note: This figure shows the cumulative wealth development of seven equal exposure variance strategies for an initial investment of $100. The data period covers January 1996 to December 2020.

Table 4.2: Return- and Risk-Statistics of Equal Exposure Strategies

Panel A: Basic Monthly Summary Statistics

	Delta-Hedged Call	Delta-Hedged Put	Variance Swap	Straddle	Butterfly Spread	Strangle	Condor Strangle
Mean	0.0050	0.0049	0.0057	0.0050	0.0049	0.0047	0.0040
Standard Dev.	0.0238	0.0248	0.0156	0.0242	0.0290	0.0226	0.0296
Skewness	−2.7361	−2.7109	−7.8579	−2.7220	−1.9814	−3.5094	−3.1060
Exc. Kurtosis	13.9543	13.7783	83.5056	13.7999	9.4847	16.0801	13.8736
Sharpe Ratio	0.1213	0.1151	0.2317	0.1202	0.0974	0.1175	0.0653

Panel B: Downside Risk Statistics

	Delta-Hedged Call	Delta-Hedged Put	Variance Swap	Straddle	Butterfly Spread	Strangle	Condor Strangle
VaR (95%)	−0.0294	−0.0316	−0.0067	−0.0303	−0.0397	−0.0323	−0.0536
CVaR (95%)	−0.0691	−0.0724	−0.0374	−0.0702	−0.0812	−0.0738	−0.0966
Max Loss	−0.1670	−0.1705	−0.1821	−0.1673	−0.1697	−0.1417	−0.2064
Average DD	−0.0352	−0.0355	−0.0229	−0.0346	−0.0515	−0.0447	−0.0579
Max DD	−0.2457	−0.2535	−0.1821	−0.2467	−0.2892	−0.2293	−0.3270
Average DD Length	5.1951	5.1220	4.1538	5.0476	6.6765	5.5758	6.8485
Max DD Length	28	28	18	28	33	17	41

Panel C: Annualized Geometric Mean Return and Terminal Wealth

	Delta-Hedged Call	Delta-Hedged Put	Variance Swap	Straddle	Butterfly Spread	Strangle	Condor Strangle
Ann. Geom. Return [%]	5.74	5.66	6.85	5.74	5.47	5.48	4.32
Term. Wealth [$]	401.35	393.99	521.00	401.92	377.13	377.70	286.93

Note: This table provides return- and risk-statistics of seven equal exposure variance strategies. Panel A depicts basic summary statistics of monthly returns, while Panel B is dedicated to downside risk statistics. In particular, it considers asymmetric risk metrics for monthly returns such as the 95% value-at-risk (VaR), the 95% conditional value-at-risk (CVar), and the maximum loss (max loss). Additionally, it shows path-dependent drawdown measures: The average drawdown (Average DD), the maximum drawdown (Max DD), the average drawdown length (Average DD length), and the maximum drawdown length (Max DD Length). Panel C shows the annualized geometric return and the terminal wealth of the strategies.

4.4.3 Why Strategies Differ: Payoff or Costs?

Even with the same gamma exposure, some differences remain between the strategies. Where do these differences come from? A first possible reason is the different payoff structures of the strategies. A second one is potential differences in transaction costs. Since new derivative positions must be set up regularly due to the finite maturity of options, the latter may be substantial.

The similarity of the payoffs should be reflected in high correlations of returns. Table 4.3 shows such correlations between the monthly returns of the different strategies. Delta-hedged call, delta-hedged put, and straddle have very high correlations of over 0.99, consistent with theoretical considerations.[35] Strangle, condor strangle, and butterfly spread are slightly less correlated, consistent with their modified payoff profiles. However, they still show correlations with the first group above 0.9. The most striking observation from Table 4.3 is the relatively low correlation (below 0.7) between the variance swap and all other strategies. This indicates the uniqueness of the variance swap.

Table 4.3: Equal Exposure Strategies' Return Correlation Matrix

	Delta-Hedged Call	Delta-Hedged Put	Variance Swap	Straddle	Butterfly Spread	Strangle
Delta-Hedged Put	0.9960					
Variance Swap	0.6631	0.6584				
Straddle	0.9990	0.9990	0.6608			
Butterfly Spread	0.9723	0.9751	0.6043	0.9746		
Strangle	0.9306	0.9311	0.6236	0.9319	0.8601	
Condor Strangle	0.9093	0.9156	0.5352	0.9133	0.9000	0.9536

Note: This table depicts the correlations between the monthly returns of seven equal exposure variance strategies.

Can the uniqueness of the variance swap really be attributed to its payoff profile? At first glance, this question is difficult to answer. The payoff of a variance swap depends on the realized variance over the price path, while the payoff of the other strategies depends only on the price of the underlying (index level) at the maturity date of the options. However, the initial replicating portfolio of the variance swap can provide some intuition. The payoff function of this options portfolio allows a direct comparison with the payoff profile of the other strategies. As an example,

[35] In a world without transaction costs, where put-call parity holds exactly, the payoff profiles of the three strategies are identical for the same gamma exposure.

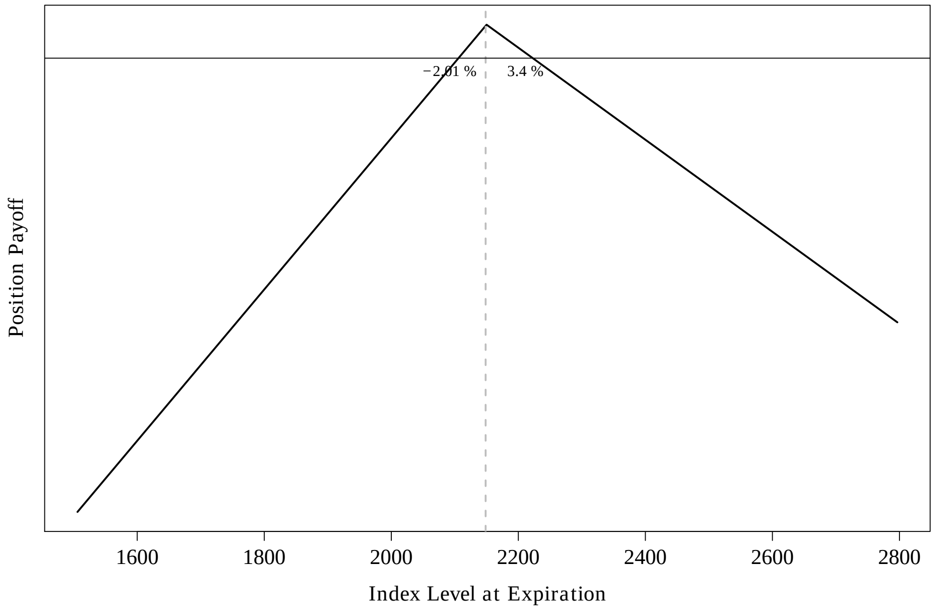
Figure 4.4 shows the payoff functions of the delta-hedged put and the replicating portfolio of the variance swap for the setup date 24/10/2016. This picture is also quite typical for other setup dates.[36] The payoff function of the delta-hedged put is a piecewise linear function, with a kink at the forward price. This structure implies that losses grow linearly with the distance between the index and the forward price. The variance swap, on the other hand, shows a strongly non-linear payoff profile. There is a wide range around the forward price where the payoff function is almost flat. This feature also leads to a relatively large gap between the two break-even points (-3.62% and 3.36%), compared to the delta-hedged put (-2.01% and 3.4%). The variance swap makes similar profits over a wide range around the ATM point. This is consistent with the strategy's relatively smooth path of cumulated wealth. There is also a strong asymmetry in the payoff of the variance swap. When the index level rises, the losses remain relatively moderate, which also contributes to the smoothness of the path. Only when the index loses massively does the variance swap realize losses that seem to grow exponentially with the index decline. The latter observation is consistent with the high negative skewness and high maximum loss of the variance swap strategy.

Another possible reason for differences between strategies is transaction costs. An intuitive guess is that strategies using more than two option contracts, or OTM and ITM options, are more affected by transaction costs than strategies using only one or two ATM options. The first group consists of the strangle, butterfly spread, condor strangle, and the replicating portfolio of the variance swap, while the second group includes the delta-hedged call, delta-hedged put, and straddle. Figure 4.5 confirms the conjecture. Part (a) shows the cumulative wealth of the different strategies when the base case effective spread is reduced from 25% to zero; and part (b) shows the cumulative wealth for an effective spread of 50%. Moving from 50% effective spread to zero transaction costs improves the final wealth (after 25 years) of the condor strangle by 30%, the butterfly spread by 24%, and the variance swap by 18%. In contrast, the delta-hedged call, put and straddle only show an increase in wealth of about 13%. However, the variance swap retains the highest final wealth of all strategies even when the effective spread is 50%. Condor strangle, strangle and butterfly spread still deliver the lowest terminal wealth even without transaction costs. So the differences in the performance of strategies are certainly not only caused by different transaction costs.

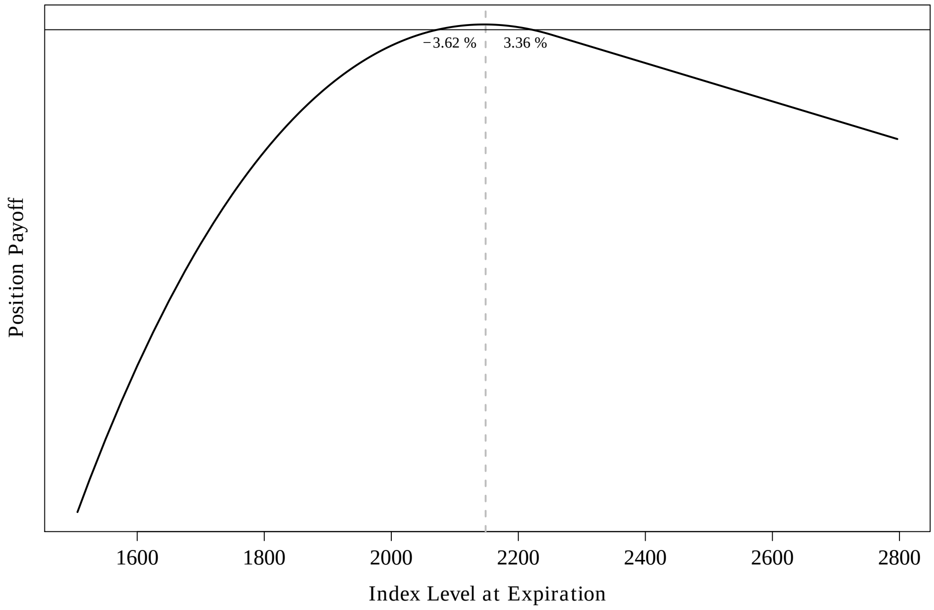
[36]The payoff function depends on the forward price, the available strike prices and the deltas of the options on the setup date. Therefore, the payoff function varies over time.

Figure 4.4: Payoffs of Delta-Hedged Put and Variance Swap

(a) Delta-Hedged Put



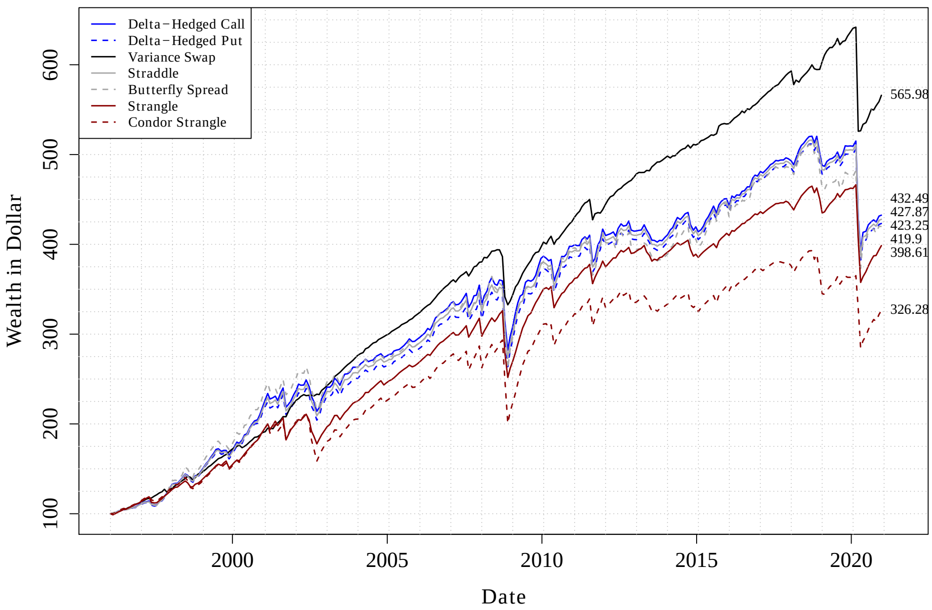
(b) Variance Swap



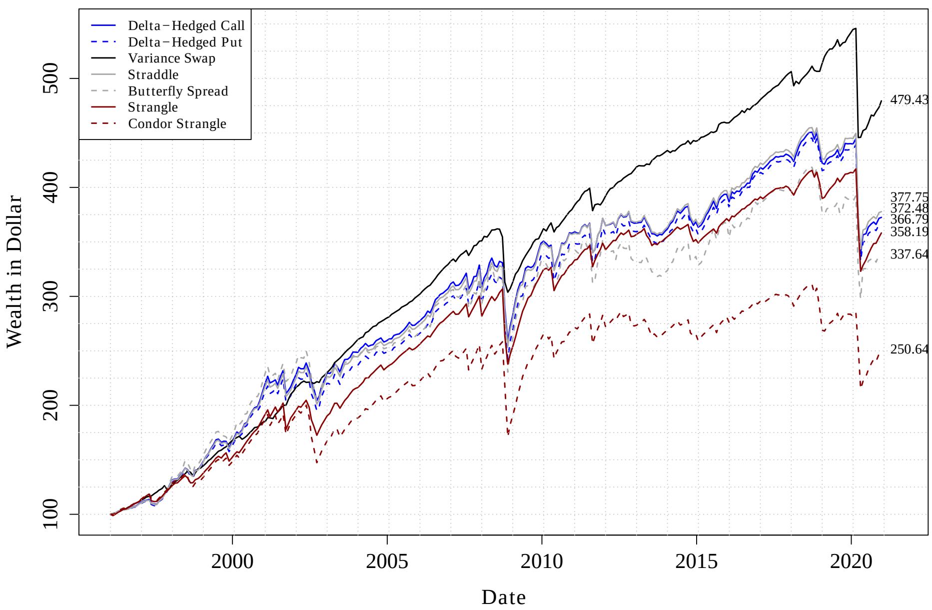
Note: This figure shows payoff diagrams of a delta-hedged put option (part (a)) and a variance swap (part (b)) from October 24, 2016, that is, at initiation of the position. The x-axis shows the index level at expiration and the y-axis the corresponding payoffs of the delta-hedged put option and the variance swap, respectively. The horizontal solid line depicts the break-even payoff and the vertical dashed line the index level at initiation.

Figure 4.5: Equal Exposure Strategies: The Impact of Transaction Costs

(a) Effective Spread: 0%



(b) Effective Spread: 50%



Note: This figure shows the cumulative wealth development of seven equal exposure variance strategies for an initial investment of $100. The data period covers January 1996 to December 2020. Part (a) shows strategies with an effective spread of 0% and part (b) the same strategies with an effective spread of 50%.

4.4.4 Variance Strategies and Market Movements

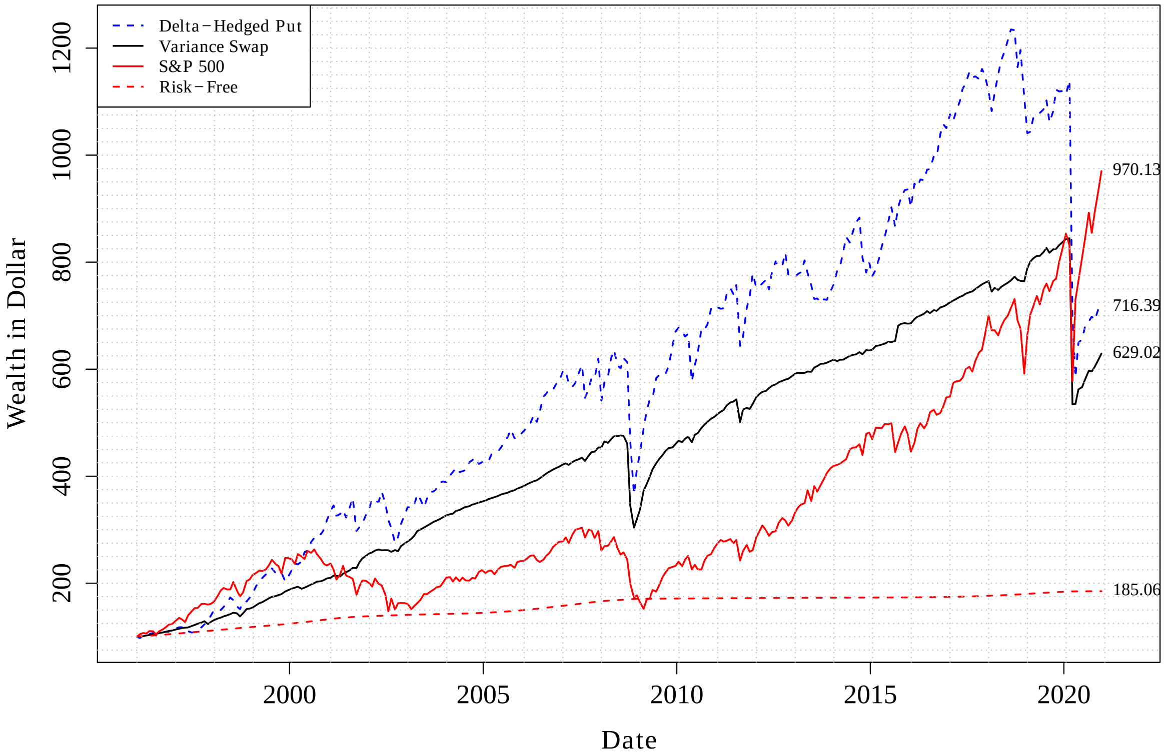
For a deeper understanding of the differentiating properties of variance strategies, we look more closely at their relationship to market (index) movements. The performance of variance strategies is closely linked to the market for at least two reasons. First, in our study, the index is the underlying of the derivatives positions that implement variance strategies. Therefore, by construction, payoffs are linked to market movements, which can be clearly seen in Figures 4.1 and 4.4. The choice between alternative profiles constitutes the payoff problem that we study. Second, previous research has shown an inverse relationship between market movements and variance changes, i.e., variances tend to increase when the market goes down (Black 1976; Christie 1982; Bollerslev et al. 2006). However, different variance strategies might be affected differently by this effect.

Since the additional design elements of strangle, condor strangle, and butterfly spread have not led to improvements in the variance strategies, we focus only on two strategies in the following subsections. The first, the delta-hedged put, represents the group consisting of delta-hedged call and put and the straddle. The second strategy is the variance swap. To give these strategies a chance to exploit their full "variance exposure", we consider their maximum exposure variants.

Figure 4.6 shows the cumulative wealth of a buy-and-hold index investment in the S&P 500, along with the two variance strategies. An investment in the risk-free instrument is also shown as a reference point. Since 2003, the trend behavior of the market and the two variance strategies has been similar. Setbacks of the index are clearly visible in the delta-hedged put strategy. This is obviously true for the financial crisis in 2008 and the Covid-19 shock in 2020, but it is also true for the market downturn in 2002 (aftermath of September 11 attacks) and the market drop at the end of 2018. In contrast, the variance swap strategy seems to be mainly affected only by the two very large market shocks. This observation suggests a higher correlation between market returns and delta-hedged put returns compared to variance swap returns. However, exactly the opposite is the case, with monthly return correlations of 0.38 (market and delta-hedged put) and 0.52 (market and variance swap).

To better understand this phenomenon, Figure 4.7 plots monthly delta-hedged put returns (part (a)) and variance swap returns (part (b)) as a function of market returns. For the delta-hedged put, there is a clear positive relationship between market returns and put returns when the market falls. However, when the market rises, the relationship is negative. Overall, the non-linear relationship between market returns and put returns results in a moderately positive correlation.

Figure 4.6: Maximum Exposure Strategies vs. Market



Note: This figure depicts the cumulative wealth development of four different trading strategies for an initial investment of $100. The strategies are the maximum exposure delta-hedged put strategy (dashed blue line), the variance swap strategy (solid black line), a buy and hold investment strategy in the S&P 500 total return index (solid red line), and an investment in risk-free assets (dashed red line). The data period covers January 1996 to December 2020.

The overall positive relationship is due to the fact that large negative market movements occur more frequently than large positive market movements. For the variance swap, the relationship is nearly flat over much of the market's return distribution. Only for very large negative market returns is the relationship clearly positive. Since there is no offsetting effect for rising markets, the correlation between the market and variance swap is positive and higher than for the delta-hedged put. These results are well in line with the examples of the payoff profiles in Figure 4.4. They show that the two variance strategies have significantly different characteristics in terms of their relationship with the market, which is important information for potential investors.

4.4.5 "Variance" as an Investment Style

So far, we have studied and compared different ways to earn variance risk premiums. In a broader perspective, the strategies under study are just examples of factor investments, because "stock index variance" can be seen as one of (potentially) many equity-based factors. In this section, we relate variance strategies to other factor investments. Specifically, we consider the S&P 500 index (market), long-short portfolios of the remaining four factors of the Fama and French (2015) five-factor model (Size (SMB), Value (HML), Profitability (RMW), Investment (CMA)), plus Momentum (MOM).[37] Factor returns are constructed such that their return periods coincide with the roll-over periods from our option strategies since the original monthly Fama and French (2015) factor returns are based on calendar months. We achieve this by computing geometric returns over the specified investment horizon for every individual portfolio and eventually determine the factor returns according to the definitions from Fama and French (2015). Data on the additional factors is from Kenneth French's website.[38]

Figure 4.8 shows the cumulative wealth development of the various factor investments. One observation is very striking. After the financial crisis, since mid-2009, none of the SMB, HML, RMW, CMA, and MOM strategies has an upward trend.[39] Only the market and the two variance strategies show significant upward movements during this period. This distinguishes the variance strategies from the other five strategies that try to earn additional premiums in the stock market besides the market risk premium. Also in the period before the financial crisis, the variance strategies together with the momentum strategy show the strongest upward trend.

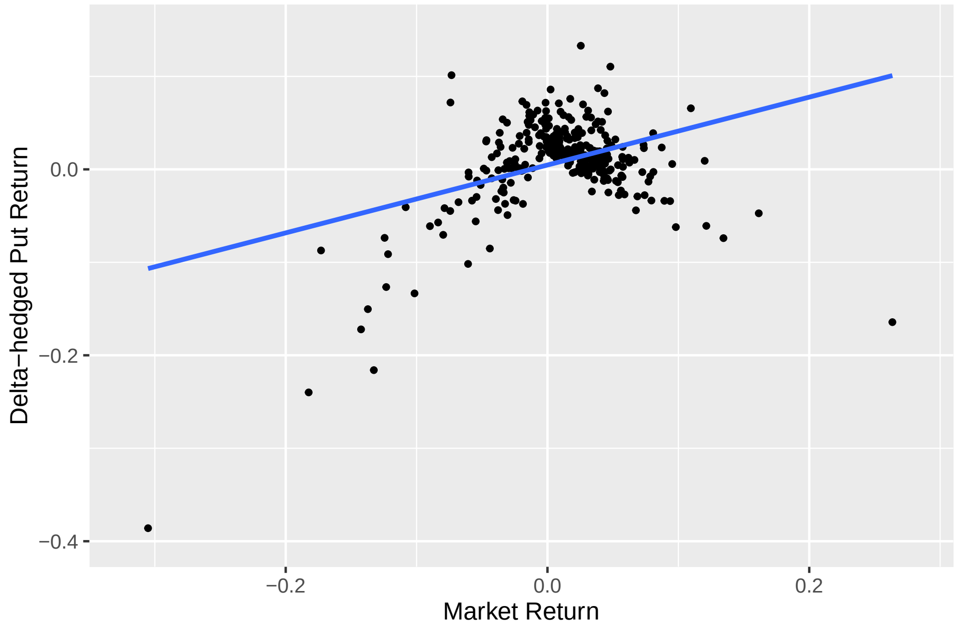
[37]Since long-short portfolios do not require an initial investment, the initial capital of $100 is invested in an account earning the risk-free rate.

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

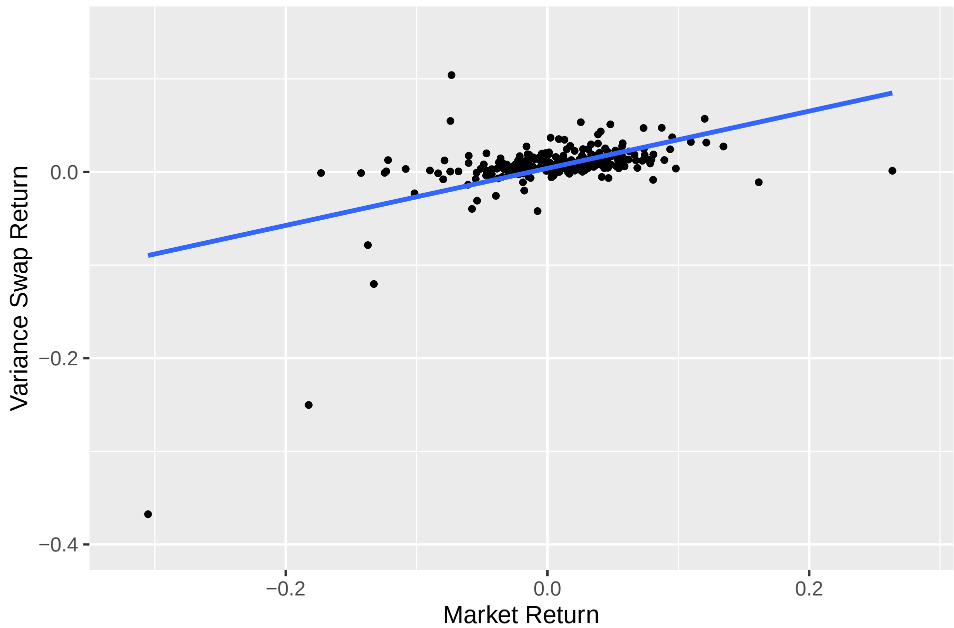
[39]When interpreting the performance of these five factor strategies, one has to keep in mind that they provide an overly optimistic view of the corresponding styles. The reason is that no transaction costs of portfolio revisions are taken into account in the performance calculations, as is the case with the variance strategies.

Figure 4.7: Return Relations Between Variance Strategies and Market

(a) Delta-Hedged Put

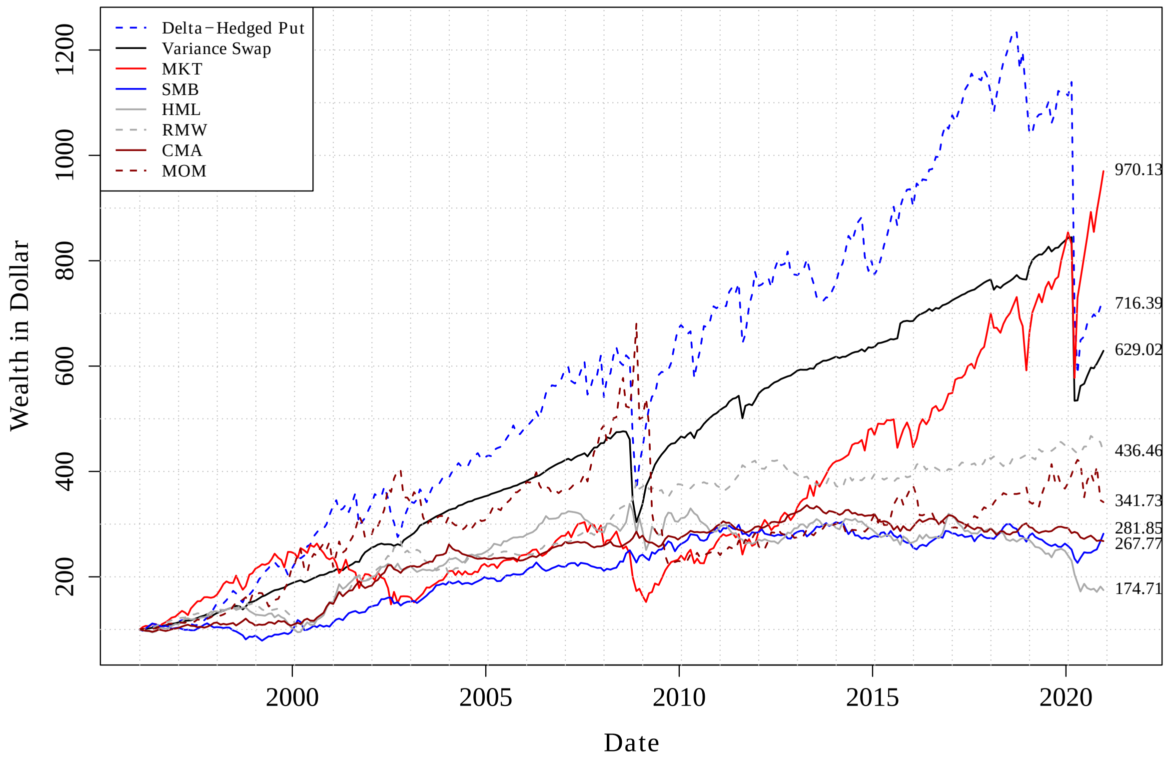


(b) Variance Swap



Note: This figure provides return scatter plots that show the contemporaneous relationship between monthly market returns and monthly delta-hedged put returns (part (a)) and monthly variance swap returns (part (b)), respectively. The blue lines indicate a linear trend, estimated via least squares.

Figure 4.8: Alternative Factor Investment Strategies: Cumulative Wealth



Note: This figure provides the cumulative wealth development of two maximum exposure variance strategies (delta-hedged put and variance swap) and six factor investment strategies that refer to the Fama and French (2015) five-factor model and the momentum factor. The data period covers January 1996 to December 2020. MKT is the market factor, SMB is the size factor, HML is the value factor, RMW is the profitability factor, CMA is the investment factor, and MOM is the momentum factor.

Table 4.4 presents return and risk statistics of the different factor investments. In our view, the most relevant reference point for the variance strategies is the market investment, since it is the only one that generates significant premiums after the financial crisis. Looking at the monthly return statistics, market and variance strategies do not differ strongly in terms of mean returns and standard deviations. However, the variance strategies show more negative skewness and higher excess kurtosis. The variance swap is the most extreme in this respect. These properties seem to support the idea that variance strategies are "picking up nickels in front of a steamroller".

But what do these monthly return moments mean for the long-term investor? Most of the time, the cumulative wealth of variance strategies is above the wealth level of the market investment, and the downside risk statistics provide an additional perspective on this question. As shown in Panel B, the variance swap has lower downside risk than the market according to all measures, except for the maximum loss. Variance strategies do occasionally experience extreme losses, but these losses are of the same magnitude as the extreme losses of a market investment. Moreover, variance strategies are able to recover from previous losses relatively quickly, i.e., they have the ability to generate strong upside movements in a relatively short period of time. This characteristic is evident after the financial crisis and also after the low point of the Covid-19 pandemic.[40] In this sense, variance strategies are able to pick up more than "nickles" in such periods.

Finally, we examine the monthly co-movement of variance strategies with other factor portfolios. In doing so, we seek a better understanding of the economic conditions under which variance premiums are small or large. Table 4.5 shows the results of regressions of the monthly returns of variance strategies on either the market returns or the returns of all six factor portfolios. There is a positive and statistically significant relationship between the variance strategies and the market. As seen in the previous section, the linear approximation does not fully reveal the true non-linear structure, at least for the delta-hedged put. However, an overall positive co-movement is economically intuitive. It is consistent with large market downturns (within a month) being larger than large market upturns, and the notion of increasing variance in falling equity markets.

The only other significant loading is a positive one on the size factor. Strategies selling insurance against high market volatility tend to perform poorly when small caps also perform relatively poorly. This is economically plausible. When times become more volatile, it is likely that small

[40]Unfortunately, our data set ends in December 2020, but some recovery after the Covid-19 shock is already evident since March 2020.

Table 4.4: Return- and Risk-Statistics of Alternative Factor Investment Strategies

Panel A: Basic Monthly Summary Statistics

	Delta-Hedged Put	Variance Swap	MKT	SMB	HML	RMW	CMA	MOM
Mean	0.0080	0.0067	0.0090	0.0039	0.0025	0.0053	0.0035	0.0057
Standard Dev.	0.0494	0.0309	0.0525	0.0298	0.0347	0.0266	0.0217	0.0539
Skewness	−2.9729	−8.1252	−0.9071	−0.0434	0.2812	0.3554	0.5697	−1.6832
Exc. Kurtosis	17.1776	88.7758	6.7505	2.3063	2.6527	6.0900	1.4345	8.4620
Sharpe Ratio	0.1195	0.1516	0.1332	0.0622	0.0116	0.1215	0.0676	0.0670

Panel B: Downside Risk Statistics

	Delta-Hedged Put	Variance Swap	MKT	SMB	HML	RMW	CMA	MOM
VaR (95%)	−0.0629	−0.0076	−0.0747	−0.0399	−0.0432	−0.0302	−0.0298	−0.0900
CVaR (95%)	−0.1448	−0.0700	−0.1323	−0.0609	−0.0728	−0.0566	−0.0396	−0.1491
Max Loss	−0.3860	−0.3676	−0.3052	−0.1341	−0.1278	−0.1078	−0.0548	−0.3401
Average DD	−0.0755	−0.0433	−0.0791	−0.0643	−0.0710	−0.0571	−0.0616	−0.1025
Max DD	−0.5266	−0.3676	−0.4993	−0.2948	−0.4942	−0.3278	−0.2032	−0.6714
Average DD Length	5.7000	4.3750	7.1875	11.2500	13.7368	12.4000	15.4706	13.3500
Max DD Length	29	23	74	82	148	65	93	146

Panel C: Annualized Geometric Mean Return and Terminal Wealth

	Delta-Hedged Put	Variance Swap	MKT	SMB	HML	RMW	CMA	MOM
Ann. Geom. Return [%]	8.22	7.66	9.55	4.25	2.26	6.09	4.03	5.06
Term. Wealth [$]	716.39	629.02	970.17	281.85	174.71	436.46	267.77	341.73

Note: This table provides return- and risk-statistics of two maximum exposure variance strategies (delta-hedged put and variance swap) and 6 other equity-based factor investment strategies. Returns of factor investment strategies are calculated according to the Fama and French (2015) five-factor model, such that the return periods coincide with the return periods from the variance strategies. MKT is the market factor, SMB is the size factor, HML is the value factor, RMW is the profitability factor, CMA is the investment factor, and MOM is the momentum factor. Panel A depicts basic summary statistics of monthly returns, while Panel B is dedicated to downside risk statistics. In particular, it considers asymmetric risk metrics for monthly returns such as the 95% value-at-risk (VaR), the 95% conditional value-at-risk (CVar), and the maximum loss (max loss). Additionally, it shows path-dependent drawdown measures: The average drawdown (Average DD), the maximum drawdown (Max DD), the average drawdown length (Average DD length), and the maximum drawdown length (Max DD Length). Panel C shows the annualized geometric return and the terminal wealth of the strategies.

Table 4.5: Maximum Exposure Strategies vs. Alternative Factor Investment Strategies: OLS Regression

	Delta-Hedged Put	Variance Swap	Delta-Hedged Put	Variance Swap
α	0.0034	0.0025	0.0029	0.0022
	(1.2467)	(1.1412)	(1.0350)	(0.9005)
β_{MKT}	0.3648***	0.3080***	0.3276***	0.2845***
	(3.8692)	(2.9263)	(4.4556)	(2.9022)
β_{SMB}			0.3236**	0.2231**
			(1.9799)	(2.3091)
β_{HML}			0.2907	0.1759
			(1.5783)	(1.3132)
β_{RMW}			0.0973	0.0385
			(0.6987)	(0.3905)
β_{CMA}			−0.1770	−0.1323
			(−0.7123)	(−1.5327)
β_{MOM}			−0.0031	0.0259
			(−0.0388)	(0.4390)
n	299	299	299	299
Adj. R^2	0.1480	0.2718	0.2072	0.3304
F-statistic	52.7617	112.2456	13.9801	25.5060

Note: This table shows the ordinary least squares (OLS) regression results from a regression of the monthly returns of two maximum exposure variance strategies (delta-hedged put, variance swap) on the returns other factor investment strategies. Specifically, we consider the S&P 500 index (MKT), long-short portfolios of the remaining four factors of the Fama and French (2015) five-factor model (Size (SMB), Value (HML), Profitability (RMW), Investment (CMA)), plus Momentum (MOM). Newey and West (1987) robust standard errors with 12 lags are reported in parentheses and asterisks indicate significance at the 1% (***), 5% (**), and 10% (*) level, respectively.

caps will struggle more than large caps because they are more vulnerable on average.[41] Taken together, the market and size factors capture 20% (delta-hedged put) or even 33% (variance swap) of the variance strategies' return fluctuations.

Are variance strategies valuable additions of the long-term investor's opportunity set of stock-related factors? Given the results of Table 4.5, the answer is ambiguous. Looking at the alphas of the delta-hedged put and the variance swap in the full regression model, they are 0.29% and 0.22% per month, respectively, which is economically significant compared with an average market return of 0.9%. However, these alphas are not statistically significant. One reason is the high residual variance caused by only two observations, October 2008 and March 2020. Moreover, due to the non-robustness of the OLS estimator, these two months are also highly influential for the resulting estimates of the factor loadings.[42] We therefore re-estimate the regression models with a more robust method, the least absolute deviation (LAD) estimator, which minimizes the mean absolute residual error.[43] Table 4.6 presents the results. A first important finding is that LAD regressions lead to much lower factor loadings for the market and the size factor, which become even insignificant for the delta-hedged put. This finding suggests potentially high diversification benefits if a delta-hedged put strategy complements a market investment, given that such benefits are measured in terms of mean absolute deviations instead of standard deviations.[44] A second important finding is that the LAD regressions attribute a higher proportion of the variance strategies' mean returns to alpha, as compared to the OLS regressions. The estimated alphas of 0.95% and 0.49% per month of the delta-hedged put and variance swap strategies, respectively, are clearly economically and statistically significant.

Overall, the results suggest that variance strategies are attractive factor strategies for long-term investors. These strategies can provide alternatives to a market investment that perform similarly overall but have clearly different downside risk characteristics. They can also be useful complements to a market investment to provide diversification benefits and significant alpha.

[41] This conjecture is supported by, for example, Duffee (1995) and Ang and Chen (2002).

[42] Cook's distances (Cook (1977)) for these two observations reach from 10 to 226 times the average Cook's distance in the respective models, indicating highly influential observations.

[43] See, for example, Hill and Holland (1977). To account for potential heteroskedasticity and autocorrelation in the LAD framework, we determine standard errors with a block bootstrap method. We use a block length of 12, to be consistent with the number of lags applied for the Newey and West (1987) standard errors of Table 4.5.

[44] Moving from a pure market investment to a portfolio that invests 50% in the market and 50% in the delta-hedged put strategy reduces the monthly standard deviation by about 20 percent but the mean absolute deviation by more than 40 percent. As the results by Goldstein and Taleb (2007) suggest, even finance professionals consider the mean absolute deviation a more intuitive dispersion measure than the standard deviation.

Table 4.6: Maximum Exposure Strategies vs. Alternative Factor Investment Strategies: LAD Regression

	Delta-Hedged Put	Variance Swap	Delta-Hedged Put	Variance Swap
α	0.0105***	0.0051***	0.0095***	0.0049***
	(0.0027)	(0.0008)	(0.0030)	(0.0009)
β_{MKT}	0.1246	0.0915***	0.0852	0.0826***
	(0.1341)	(0.0239)	(0.1362)	(0.0255)
β_{SMB}			0.1447	0.0736***
			(0.1079)	(0.0275)
β_{HML}			−0.0120	0.0200
			(0.1192)	(0.0323)
β_{RMW}			0.1752	0.0021
			(0.1618)	(0.0329)
β_{CMA}			−0.0844	−0.0016
			(0.2227)	(0.0437)
β_{MOM}			−0.0640	−0.0160
			(0.0656)	(0.0267)
n	299	299	299	299
R^2_{Pseudo}	0.0054	0.0777	0.0216	0.0999

Note: This table shows the least absolute deviation (LAD) regression results from a regression of the monthly returns of two maximum exposure variance strategies (delta-hedged put, variance swap) on the returns other factor investment strategies. Specifically, we consider the S&P 500 index (MKT), long-short portfolios of the remaining four factors of the Fama and French (2015) five-factor model (Size (SMB), Value (HML), Profitability (RMW), Investment (CMA)), plus Momentum (MOM). Newey and West (1987) robust standard errors with 12 lags are reported in parentheses and asterisks indicate significance at the 1% (***), 5% (**), and 10% (*) level, respectively.

4.5 Conclusion

The variance risk premium is a well-documented empirical phenomenon. In this paper, we analyze whether and how investors can exploit this premium for long-term capital accumulation. Our paper identifies three general problems that arise in long-term variance-based investment strategies. Certain design elements that could mitigate some of these problems are suggested, and corresponding trading strategies are proposed. To determine the variance risk to which strategies are exposed ex-ante, we either consider the maximum exposure based on capital requirements or equalize exposure via a model-free measure of convexity (gamma).

In an empirical study for the S&P 500 index options market, we analyze the performance of different strategies. We compare them to each other and to equity-based factor investing strategies. The analysis shows that variance strategies differ substantially in some aspects of risk and return, are significantly positively correlated with the market, and consistently earn premiums over the entire study period. The latter distinguishes variance strategies from other factor strategies, which have not generated premiums since the 2008 financial crisis. However, variance strategies can be hit hard by extreme stock market crashes, but also have the potential to recover quickly from these shocks. All in all, the empirical results show that variance strategies can be attractive to the long-term investor—both as an alternative and as a complement to a market investment—if properly designed. Future research could explore other design elements of variance strategies to further improve long-term variance-based investing. For example, comparing the use of options with different maturities or varying the roll-over periods would shed more light on the finite maturity problem.

B Appendix

B.1 Model-Free Deltas and Gammas

As Black-Scholes greeks depend on the unrealistic assumptions of the Black-Scholes model, we compute deltas and gammas in a model-free fashion. For illustrative purposes, consider the discrete future stock price S_i with future states of the world $i = 1, \ldots, \bar{i}, \ldots, N$, and the corresponding probabilities q_i. Further, assume that the discount rate is $r = 0$. Now, shift the prices S_i by an infinitesimally small amount ε, such that the new prices and corresponding probabilities are

$$S_1 + \varepsilon, \ldots, S_N + \varepsilon,$$
$$q_1, \ldots, q_N.$$

The value of a plain-vanilla call option C^E with $X = S_{\bar{i}}$ prior to the shift is

$$C^E(X) = \sum_{i=\bar{i}+1}^{N} (S_i - X)\, q_i. \tag{B.1}$$

After the shift, the value is

$$\begin{aligned} C^E(X) &= \sum_{i=\bar{i}}^{N} (S_i + \varepsilon - X)\, q_i \\ &= \sum_{i=\bar{i}+1}^{N} (S_i - X)\, q_i + \sum_{i=\bar{i}+1}^{N} \varepsilon q_i + (S_{\bar{i}} + \varepsilon - X)\, q_{\bar{i}} \\ &= \sum_{i=\bar{i}+1}^{N} (S_i - X)\, q_i + \sum_{i=\bar{i}}^{N} \varepsilon q_i. \end{aligned} \tag{B.2}$$

The delta of any option O is defined as the change in value of the option divided by the change in value of the underlying, that is $\Delta = dO/dS$. Equation (B.2) shows that for $dS = \varepsilon$, the change in value of any option O is $dO = \sum_{i=\bar{i}}^{N} \varepsilon q_i$. Thus,

$$\Delta_C^{mf} = \frac{dO}{dS} = \frac{\sum_{i=\bar{i}}^{N} \varepsilon q_i}{\varepsilon} = \sum_{i=\bar{i}}^{N} q_i \tag{B.3}$$

In the case of a plain vanilla call option, this toy example shows that the model-free delta, which can be derived from the whole risk neutral distribution, is equal to the cumulative probability of all future outcomes that are above the strike price. Conversely, it can be shown that the model-free delta for put options is

$$\Delta_P^{mf} = \frac{dO}{dS} = \frac{-\sum_{i=1}^{\bar{i}} \varepsilon q_i}{\varepsilon} = -\sum_{i=1}^{\bar{i}} q_i. \tag{B.4}$$

Going one step further, it is easy to see that a model-free gamma γ^{mf}—the change in the delta of an option induced by a change in the underlying's price—is simply the level of the probability density function $q_{\bar{i}}$ (PDF) for both call options and put options:

$$\gamma_C^{mf} = \gamma_P^{mf} = q_{\bar{i}} \tag{B.5}$$

B.2 CBOE Margin Requirements

The following appendix provides a summary of relevant CBOE margin requirements that are used to construct the variance strategies. Further information and sample calculations can be found on the homepage of the CBOE.[45]

Short Call Initial Margin Requirement:

- 100% of option proceeds, plus 15% of aggregate underlying index value (number of contracts × index level × $100) less out-of-the-money amount, if any
- minimum requirement is option proceeds plus 10% of the aggregate underlying index value
- proceeds received from sale of call(s) may be applied to the initial margin requirement
- after position is established, ongoing maintenance margin requirement applies, and an increase (or decrease) in the margin required is possible

Short Put Initial Margin Requirement:

- 100% of option proceeds, plus 15% of aggregate underlying index value (number of contracts × index level × $100) less out-of-the-money amount, if any
- minimum requirement is option proceeds plus 10% of the put's aggregate strike price (number of contracts × strike price × $100)
- proceeds received from sale of puts(s) may be applied to the initial margin requirement
- after position is established, ongoing maintenance margin requirement applies, and an increase (or decrease) in the margin required is possible

Short Straddle Initial Margin Requirement:

- short call(s) or short put(s) requirement, whichever is greater, plus the option proceeds of the other side
- proceeds from sale of entire straddle may be applied to initial margin requirement
- after position is established, ongoing maintenance margin requirement applies, and an increase (or decrease) in the margin required is possible

[45] `https://www.cboe.com/us/options/strategy_based_margin/`.

Short Strangle Initial Margin Requirement:

- short call(s) or short put(s) requirement, whichever is greater, plus the option proceeds of the other side
- proceeds from sale of entire strangle may be applied to initial margin requirement
- after position is established, ongoing maintenance margin requirement applies, and an increase (or decrease) in the margin required is possible

Short Call Spread Initial Margin Requirement:

- the amount by which the short call aggregate strike price is below the long call aggregate strike price (aggregate strike price = number of contracts × strike price × $100)
- long call(s) must be paid for in full
- proceeds received from sale of short call(s) may be applied to the initial margin requirement
- the short call(s) may expire before the long call(s) and not affect margin requirement

Short Put Spread Initial Margin Requirement:

- the amount by which the long put aggregate strike price is below the short put aggregate strike price (aggregate strike price = number of contracts × strike price × $100)
- long put(s) must be paid for in full
- proceeds received from sale of short put(s) may be applied to the initial margin requirement
- the short put(s) may expire before the long put(s) and not affect margin requirement

5 Conclusion

This dissertation presents novel evidence on moment risk premiums in option markets. In the past two decades, moment risk premiums have been of particular interest in financial economics research. There are some stylized facts about these premiums that have been established throughout the past years: The literature has found that premiums for higher moments exist, they are time varying and highly correlated with each other, they can be decomposed into downside and upside related components, they are extremely affected be rare extreme events, and they can be used to predict future stock market returns.[46]

However, the premiums' special structure and their particular characteristics raise other pivotal issues, which have not yet been addressed in the literature. From these issues, three main research questions arise, to which the individual chapters of this dissertation make important contributions:

(i) How can moment risk premiums be measured and quantified in a robust and flexible manner and what are these robust premiums' characteristics?

(ii) What are the characteristics of decomposed higher-moment risk premiums and how can a better knowledge of their structure be utilized to predict market returns?

(iii) Is it possible for long-term investors to harvest the variance risk premium, and if so, what are feasible trading strategies to achieve this goal?

The first study in this dissertation (Chapter 2) provides answers and evidence for the first major research question. It investigates quantile-based alternatives to traditional moment swaps, such as variance swaps and skewness swaps. It introduces the q-quantile swap as a basic building block and shows that portfolios of these swaps can be used to quantify risk premiums associated with any partition of the risk-neutral return distribution in a robust fashion. In an empirical study

[46]See, for example, Bollerslev, Tauchen, and Zhou (2009); Carr and Wu (2009); Kozhan, Neuberger, and Schneider (2013); Kilic and Shaliastovich (2019); Khrashchevskyi (2020); Fan, Xiao, and Zhou (2020).

for the S&P 500 options market, the quantile-based approach can identify two option-specific premiums: One premium is related to tail extremity and one to the asymmetry in the return distribution. Moreover, the results show that it definitely matters how moment risk premiums are measured and they warn us to be cautious when interpreting traditional higher-moment risk premiums, such as skewness and kurtosis, as they are strongly driven by tail extremity.

The second study (Chapter 3) examines decompositions of higher-moment risk premiums up to the fourth moment. In doing so, this paper is the first to analyze decomposed skewness and kurtosis risk premiums in the equity index option market and provides evidence for the second main research question. Additionally, a new measure of kurtosis risk premium is proposed, which is also decomposed into its downside and upside components. In an empirical forecasting study, this chapter presents strong in-sample and out-of-sample evidence that information in decomposed higher-moment premiums can be exploited for predictions of subsequent market excess returns. The empirical findings suggests that (decomposed) variance risk premium predictors tend to yield the best prediction results at short forecast horizons of up to six or nine months, whereas downside higher-moment risk premiums in particular yield good predictions at longer horizons.

The third study (Chapter 4) analyzes whether and how investors can exploit the variance risk premium for long-term capital accumulation. In doing so, it contributes to answering the third major research question. The study identifies three general problems encountered with long-term variance-based investment strategies. Specific design elements that could alleviate some of these problems are suggested, and corresponding trading strategies are proposed. The strategies are then compared with each other and with equity-based factor investing strategies in an empirical study. The analysis shows that variance strategies differ considerably in some aspects of risk and return and consistently earn premiums over the entire sample period. The latter distinguishes variance strategies from other factor strategies that have not generated premiums since the financial crisis in 2008. All in all, the empirical results show that variance strategies can be attractive for the long-term investor—both as an alternative and as a complement to a market investment—if they are properly designed.

Altogether, this dissertation contributes to the literature on moment risk premiums in option markets by providing new theoretical approaches and empirical results on the measurement, the structure, and the investment implications of moment risk premiums. The core findings of this thesis can be condensed as follows: (i) The dissertation finds a novel risk premium associated with asymmetry when moment risk premiums are measured with a quantile-based approach,

(ii) traditional moment risk premiums are driven by tail extremity and need to be interpreted carefully, (iii) downside moment risk premiums exceed upside moment risk premiums and carry information about subsequently realized market risk premiums, (iv) the variance risk premium can be exploited to accumulate capital, (v) and trading strategies based on the variance risk premium require thoughtful design decisions since their performances critically depend on a proper design.

The findings in this thesis are relevant for both researchers and practitioners in the fields of (empirical) asset pricing, asset management, and risk management. Based on this dissertation, there are several ways to conduct further research. A first natural way to extend the research is to apply the methods established in this thesis to other asset classes, such as commodities or currencies, and to other countries. Also cryptocurrencies are highly interesting in this respect since the analysis of option-implied information for these assets is only at its beginning (Alexander and Imeraj 2020; Woebbeking 2021). Additionally, the cross-section of equity options is another particularly interesting field. A broad strand of literature has been investigating variance risk premiums in the cross-section of equity options (Carr and Wu 2009; Schürhoff and Ziegler 2011; Gourier 2016). However, the existence of higher-moment risk premiums in the cross-section is still questionable, and this thesis provides a multitude of tools to measure these premiums in the cross-section. Moreover, this dissertation focuses on monthly options. Yet, it is of particular interest for future research to extend the maturities under consideration to longer horizons in order to analyze structural effects and the term structure of premiums. Beyond that, options with different maturities or with varying roll-over periods could be utilized to shed more light on the finite maturity problem (c.f. Chapter 4).

For practitioners in risk management, the findings in this dissertation help to identify relevant risks and risk factors associated with moment risk, which is an important groundwork to developing risk management strategies for these risks. Asset managers, such as institutional investors, are on the one hand provided with improved market return predictions and on the other hand with guidance for constructing long-term trading strategies that constantly earn premiums and that have an attractive risk-return profile.

Bibliography

Alexander, C. and A. Imeraj (2020). The Bitcoin VIX and its variance risk premium. *Journal of Alternative Investments 23*(4), 84–109.

Ammann, M. and A. Feser (2019a). Option-implied value-at-risk and the cross-section of stock returns. *Review of Derivatives Research 22*(3), 449–474.

Ammann, M. and A. Feser (2019b). Robust estimation of risk-neutral moments. *Journal of Futures Markets 39*(9), 1137–1166.

Ammann, M. and M. Moerke (2021). Credit variance risk premiums. Working paper, Swiss Institute of Banking and Finance, St. Gallen.

Ang, A. and J. Chen (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics 63*(3), 443–494.

Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing.* Markham Pub. Co.

Aruoba, S. B., F. X. Diebold, and C. Scotti (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics 27*(4), 417–427.

Bakshi, G. and N. Kapadia (2003). Delta-hedged gains and the negative market volatility risk premium. *Review of Financial Studies 16*(2), 527–566.

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

Bakshi, G. and D. Madan (2000). Spanning and derivative security valuation. *Journal of Financial Economics 55*(2), 205–238.

Bakshi, G. and D. Madan (2006). A theory of volatility spreads. *Management Science 52*(12), 1945–1956.

Bakshi, G., D. Madan, and G. Panayotov (2010). Returns of claims on the upside and the viability of U-shaped pricing kernels. *Journal of Financial Economics 97*(1), 130–154.

Bali, T. G., S. J. Brown, S. Murray, and Y. Tang (2017). A lottery-demand-based explanation of the beta anomaly. *Journal of Financial and Quantitative Analysis 52*(6), 2369–2397.

Bali, T. G., N. Cakici, and R. F. Whitelaw (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics 99*(2), 427–446.

Bali, T. G. and S. Murray (2013). Does risk-neutral skewness predict the cross-section of equity option portfolio returns? *Journal of Financial and Quantitative Analysis 48*(4), 1145–1171.

Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177–181.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*(3), 637–654.

Bollerslev, T., J. Litvinova, and G. Tauchen (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics 4*(3), 353–384.

Bollerslev, T., J. Marrone, L. Xu, and H. Zhou (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence. *Journal of Financial and Quantitative Analysis 49*(3), 633–661.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies 22*(11), 4463–4492.

Bollerslev, T. and V. Todorov (2011). Tails, fears, and risk premia. *Journal of Finance 66*(6), 2165–2211.

Bollerslev, T., V. Todorov, and L. Xu (2015). Tail risk premia and return predictability. *Journal of Financial Economics 118*(1), 113–134.

Bowley, A. L. (1920). *Elements of Statistics*, Volume 2. P.S. King.

Breeden, D. T. and R. H. Litzenberger (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business 51*(4), 621–651.

Brière, M., A. Burgues, and O. Signori (2010). Volatility exposure for strategic asset allocation. *Journal of Portfolio Management 36*(3), 105–116.

Britten-Jones, M., A. Neuberger, and I. Nolte (2011). Improved inference in regression with overlapping observations. *Journal of Business Finance & Accounting 38*(5-6), 657–683.

Broll, M. (2016). The skewness risk premium in currency markets. *Economic Modelling 58*, 494–511.

Campbell, J. Y. and S. B. Thompson (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies 21*(4), 1509–1531.

Cao, J. and B. Han (2013). Cross section of option returns and idiosyncratic stock volatility. *Journal of Financial Economics 108*(1), 231–249.

Cao, J., B. Han, X. Zhan, and Q. Tong (2021). Option return predictability. *Review of Financial Studies*, forthcoming.

Carr, P. and K. Lewis (2004). Corridor variance swaps. *Risk 17*(2), 67–72.

Carr, P. and D. Madan (2001). Optimal positioning in derivative securities. *Quantitative Finance 1*(1), 19–37.

Carr, P. and L. Wu (2009). Variance risk premiums. *Review of Financial Studies 22*(3), 1311–1341.

Chang, B. Y., P. Christoffersen, and K. Jacobs (2013). Market skewness risk and the cross section of stock returns. *Journal of Financial Economics 107*(1), 46–68.

Chang, B. Y., P. Christoffersen, K. Jacobs, and G. Vainberg (2012). Option-implied measures of equity risk. *Review of Finance 16*(2), 385–428.

Christie, A. A. (1982). The stochastic behavior of common stock variances—Value, leverage and interest rate effects. *Journal of Financial Economics 10*(4), 407–432.

Clark, S. P. and M. Dickson (2019). Performance expectations of basic options strategies may be different than you think. *Journal of Asset Management 20*(2), 91–102.

Clark, T. E. and K. D. West (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics 138*(1), 291–311.

Cochrane, J. H. (2007). Financial markets and the real economy. In R. Mehra (Ed.), *Handbook of the Equity Risk Premium*. Elsevier BV.

Conrad, J., R. F. Dittmar, and E. Ghysels (2013). Ex ante skewness and expected stock returns. *Journal of Finance 68*(1), 85–124.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics 19*(1), 15–18.

Coval, J. D. and T. Shumway (2001). Expected option returns. *Journal of Finance 56*(3), 983–1009.

Cremers, M., M. Halling, and D. Weinbaum (2015). Aggregate jump and volatility risk in the cross-section of stock returns. *Journal of Finance 70*(2), 577–614.

Cuesdeanu, H. and J. C. Jackwerth (2018). The pricing kernel puzzle in forward looking data. *Review of Derivatives Research 21*(3), 253–276.

Da Fonseca, J. and E. Dawui (2021). Semivariance and semiskew risk premiums in currency markets. *Journal of Futures Markets 41*(3), 290–324.

Da Fonseca, J. and Y. Xu (2017). Higher moment risk premiums for the crude oil market: A downside and upside conditional decomposition. *Energy Economics 67*, 410–422.

Da Fonseca, J. and Y. Xu (2019). Variance and skew risk premiums for the volatility market: The VIX evidence. *Journal of Futures Markets 39*(3), 302–321.

De Fontnouvelle, P., R. P. H. Fishe, and J. H. Harris (2003). The behavior of bid-ask spreads and volume in options markets during the competition for listings in 1999. *Journal of Finance 58*(6), 2437–2463.

Doran, J. S., D. Jiang, and D. R. Peterson (2011). Gambling preference and the new year effect of assets with lottery features. *Review of Finance 16*(3), 685–731.

Duffee, G. R. (1995). Stock returns and volatility—A firm-level analysis. *Journal of Financial Economics 37*(3), 399–420.

Elyasiani, E., L. Gambarelli, and S. Muzzioli (2020). Moment risk premia and the cross-section of stock returns in the European stock market. *Journal of Banking & Finance 111*, 105732.

Fallon, W. and J. Park (2016). An asset class characterization of the U.S. equity index volatility risk premium. *Journal of Portfolio Management 43*(1), 72–84.

Fallon, W., J. Park, and D. Yu (2015). Asset allocation implications of the global volatility premium. *Financial Analysts Journal 71*(5), 38–56.

Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business 38*(1), 34–105.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Fan, Z., X. Xiao, and H. Zhou (2020). Moment risk premia and stock return predictability. *Journal of Financial and Quantitative Analysis*, 1–27.

Feunou, B., M. R. Jahan-Parvar, and C. Okou (2018). Downside variance risk premium. *Journal of Financial Econometrics 16*(3), 341–383.

Figlewski, S. (2010). Estimating the implied risk-neutral density for the U.S. market portfolio. In T. Bollerslev, J. R. Russell, and M. Watson (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, pp. 323–353. New York: Oxford University Press.

Figlewski, S. (2018). Risk-neutral densities: A review. *Annual Review of Financial Economics 10*, 329–359.

Ge, W. (2016a). Incorporating smart beta into portfolios: A case study with the volatility risk premium. *Journal of Index Investing 7*(1), 17–24.

Ge, W. (2016b). A survey of three derivative-based methods to harvest the volatility premium in equity markets. *Journal of Investing 25*(3), 48–58.

Ge, W. (2017). Stress-testing volatility risk premium harvesting strategies based on S&P 500 index options. *Journal of Index Investing 8*(1), 37–46.

Goldstein, D. G. and N. N. Taleb (2007). We don't quite know what we are talking about. *Journal of Portfolio Management 33*(4), 84–86.

Gourier, E. (2016). Pricing of idiosyncratic equity and variance risks. Working paper, Queen Mary University of London, London.

Goyal, A. and A. Saretto (2009). Cross-section of option returns and volatility. *Journal of Financial Economics 94*(2), 310–326.

Hain, M., M. Uhrig-Homburg, and N. Unger (2018). Risk factors and their associated risk premia: An empirical analysis of the crude oil market. *Journal of Banking & Finance 95*, 44–63.

Harris, R. D. F. and F. Qiao (2018). Moment risk premia and the cross-section of stock returns. Working paper, University of Exeter and Tsinghua University.

Harrison, J. M. and D. M. Kreps (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory 20*(3), 381–408.

Held, M., J. Kapraun, M. Omachel, and J. Thimme (2020). Up- and downside variance risk premia in global equity markets. *Journal of Banking & Finance 118*, 105875.

Hill, R. W. and P. W. Holland (1977). Two robust alternatives to least-squares regression. *Journal of the American Statistical Association 72*(360a), 828–833.

Hinkley, D. V. (1975). On power transformations to symmetry. *Biometrika 62*(1), 101–111.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies 5*(3), 357–386.

Huang, T. and J. Li (2019). Semivariance premium and expected stock returns. Working paper, International Business School Suzhou.

Israelov, R. and L. N. Nielsen (2015). Covered calls uncovered. *Financial Analysts Journal 71*(6), 44–57.

Jiang, G. J. and Y. S. Tian (2005). The model-free implied volatility and its information content. *Review of Financial Studies 18*(4), 1305–1342.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–292.

Khrashchevskyi, I. (2020). Macroeconomic risk and higher-moment risk premia. Working paper, Stockholm Business School, Stockholm.

Kilic, M. and I. Shaliastovich (2019). Good and bad variance premia and expected returns. *Management Science 65*(6), 2522–2544.

Kim, T. H. and H. White (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters 1*(1), 56–73.

Kozhan, R., A. Neuberger, and P. Schneider (2013). The skew risk premium in the equity index market. *Review of Financial Studies 26*(9), 2174–2203.

Kumar, A. (2009). Who gambles in the stock market? *Journal of Finance 64*(4), 1889–1933.

Lin, H., C. Wu, and G. Zhou (2018). Forecasting corporate bond returns with a large set of predictors: An iterated combination approach. *Management Science 64*(9), 4218–4238.

Lin, Y., T. Lehnert, and C. Wolff (2019). Skewness risk premium: Theory and empirical evidence. *International Review of Financial Analysis 63*, 174–185.

Londono, J. M. and N. R. Xu (2020). Variance risk premium components and international stock return predictability. Working paper, Boston College, Carroll School of Management.

Lu, Z. and S. Murray (2019). Bear beta. *Journal of Financial Economics 131*(3), 736–760.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business 36*(4), 394–419.

Mayhew, S. (2002). Competition, market structure, and bid-ask spreads in stock option markets. *Journal of Finance 57*(2), 931–958.

Moors, J. J. A. (1988). A quantile alternative for kurtosis. *The Statistician 37*(1), 25–32.

Neuberger, A. (2012). Realized skewness. *Review of Financial Studies 25*(11), 3423–3455.

Neuberger, A. and R. Payne (2021). The skewness of the stock market over long horizons. *Review of Financial Studies 34*(3), 1572–1616.

Neumann, M., M. Prokopczuk, and C. Wese Simen (2016). Jump and variance risk premia in the S&P 500. *Journal of Banking & Finance 69*, 72–83.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Newey, W. K. and K. D. West (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies 61*(4), 631–653.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica 32*(1–2), 122–136.

Rapach, D. E., J. K. Strauss, and G. Zhou (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies 23*(2), 821–862.

Rapach, D. E. and G. Zhou (2013). Forecasting stock returns. In *Handbook of Economic Forecasting*, pp. 328–383. Elsevier.

Ruan, X. and J. E. Zhang (2019). Moment spreads in the energy market. *Energy Economics 81*, 598–609.

Rubinstein, M. E. (1973). The fundamental theorem of parameter-preference security valuation. *Journal of Financial and Quantitative Analysis 8*(1), 61–69.

Schneider, P. (2015). Generalized risk premia. *Journal of Financial Economics 116*(3), 487–504.

Schneider, P., C. Wagner, and J. Zechner (2020). Low-risk anomalies? *Journal of Finance 75*(5), 2673–2718.

Schürhoff, N. and A. Ziegler (2011). Variance risk, financial intermediation, and the cross-section of expected option returns. Working paper, University of Lausanne, Lausanne.

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

Simon, D. P. (2017). Trading the VIX futures roll and volatility premiums with VIX options. *Journal of Futures Markets 37*(2), 184–208.

Simon, D. P. and J. Campasano (2014). The VIX futures basis: Evidence and trading strategies. *Journal of Derivatives 21*(3), 54–69.

Stock, J. H. and M. W. Watson (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting 23*(6), 405–430.

Todorov, V. (2010). Variance risk-premium dynamics: The role of jumps. *Review of Financial Studies 23*(1), 345–383.

Ungar, J. and M. T. Moran (2009). The cash-secured PutWrite strategy and performance of related benchmark indexes. *Journal of Alternative Investments 11*(4), 43–56.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

Whaley, R. E. (2002). Return and risk of CBOE Buy Write monthly index. *Journal of Derivatives 10*(2), 35–42.

Woebbeking, F. (2021). Cryptocurrency volatility markets. *Digital Finance*, forthcoming.

Statement of contribution

to each paper of the cumulative dissertation

To the three papers of the cumulative dissertation, I personally contributed as follows:

1. to the paper *Quantile Risk Premiums* co-authored by Felix Brinkmann and Olaf Korn: significant contributions to the conceptualization, empirical analysis, and writing, in total: 80%,

2. to the paper *Decomposed Higher-Moment Risk Premiums and Market Return Predictability* (single-author paper): conceptualization, empirical analysis, and writing, in total: 100%, and

3. to the paper *How to Harvest Variance Risk Premiums for the Long-term Investor?* co-authored by Olaf Korn and Gabriel J. Power: significant contributions to the conceptualization, empirical analysis, and writing, in total: 80%.

Göttingen, 20.09.2021

Place, Date

Julian Dörries

Ph.D. program in Economics

Declaration for admission to the doctoral examination

I confirm

1. that the dissertation "Moment Risk Premiums in Option Markets: On Measurement, Structure, and Investment Implications" that I submitted was produced independently without assistance from external parties, and not contrary to high scientific standards and integrity,

2. that I have adhered to the examination regulations, including upholding a high degree of scientific integrity, which includes the strict and proper use of citations so that the inclusion of other ideas in the dissertation are clearly distinguished,

3. that in the process of completing this doctoral thesis, no intermediaries were compensated to assist me neither with the admissions or preparation processes, and in this process,

   - no remuneration or equivalent compensation were provided

   - no services were engaged that may contradict the purpose of producing a doctoral thesis

4. that I have not submitted this dissertation or parts of this dissertation elsewhere.

I am aware that false claims (and the discovery of those false claims now, and in the future) with regards to the declaration for admission to the doctoral examination can lead to the invalidation or revoking of the doctoral degree.

Göttingen, 20.09.2021

Place, Date

Julian Dörries