Eötvös Loránd University

Corvinus University of Budapest

MSc Thesis

Empirical Portfolio Selection based on Option Implied Measures

Author:

Miklós Milán Vancsa

Supervisor:

Milán Csaba Badics

December 27, 2018

Acknowledgements

Writing this thesis would not have been possible without the guidance of Milan Badics. His proposals and academic knowledge was crucial to analyse the topic of this thesis in great depth. I would also like to thank to Gabor Neszveda, who made it possible for me to use data from the OptionMetrics database.

I would also like to thank my family, who always supported me during my studies, and my girlfriend, whose patience was essential to me during my master studies.

Contents

List of Figures

List of Tables

1 Introduction

Predicting returns on the stock market and efficient portfolio selection have been age-old problems of the financial academic literature. Recently a new area of research have emerged, which tries to connect the evolution of market returns of a given asset to the option market of the respective assets. This information spillover was first analysed in great details by Carr and Wu (2008) and Bali and Hovakimiam (2009). During my thesis I will introduce numerous academic papers, which were written in recent years about this topic. These papers in general define some kinds of measures, which try to grasp different information from the option market. These pieces of information mainly try to contain some features of the risk neutral distribution implied by the market price of traded call and put options on the underlying asset. Researchers suspect that these pieces of option-implied information has significant explanatory and predictive power for the future returns of the underlying asset. The basis of this hypothesis is coming from the fact that measures like the Black-Scholes implied volatility surface is containing market expectations about the future evolution of some features (like the volatility) for the option horizon.

In my thesis, I try to analyse the connection of various option-implied measures with the returns on equity indices and individual equities particularly in the case of portfolio selection. To analyse the topic of connection between option-implied measures and equity returns, I have studied two US equity indices and 24 individual equities. I have two research question. First, I want to analyze whether these variables are priced on the market and have statistical connection with underlying returs. In order to answer this I will use standard methods like expectation hypothesis regression and time series regressions. My main research questions are whether there is an information spillover from option markets to equity markets, and whether this phenomenon is helpful in portfolio selection?

To answer these research questions first I define the option-implied measures used in the thesis and then calculate them for my samples. The implied measures are calculated using MATLAB. Then I use various statistical models to carry out a rolling window forecast from these variables, and these predictions will be the basis of the portfolio selection. The predictions and the selection are carried out in Python.

The structure of my thesis is the following. Already in the Introduction chapter I will give a short description for the implied volatility calculated from the classical Black-Scholes model, which is the traditional concept of implied volatility. The next chapter is dealing with the model-free way of calculating implied moments from call and put prices available in the market. In that chapter I introduce the method of Breeden and Litzenberger (1978) and the method of Bakshi et al (2003), which will be the method I actually use to calculate model-free implied moments. After that I define the risk premia, which is the difference between realized return moments and the implied moments from the previous chapter. Then I try to describe the empirical features of these risk premia. The last chapter of the main part of this thesis is answering my empirical research questions. Lastly I carry out some robustness checks and draw my conclusions.

1.1 Black-Scholes Implied Volatility

In their seminal papers Black and Scholes (1973) and Merton (1973) introduced a completely new philosophy of derivative security pricing and paved path to many nowadays industry-standard financial mathematical models. The derivation of the Black-Scholes-Merton model is not in the scope of this thesis, I will only report some assumptions and results which are relevant for this work here. Interested readers are recommended to check more details about the model from for example Shreve (2008). The model aims to price simple European call (or put) options on an underlying (originally equities) and assumes that the price process of this underlying follows a Geometric Brownian Motion with a constant drift and volatility paramater, so

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t), \tag{1}$$

where $S(t)$ denotes the price of the underlying and $W(t)$ is a one-dimensional Wiener-process. If we assume that the risk-free rate r is constant and the underlying pays no cash-flow during the term of the option contract (for example an equity does not pay dividends), then these assumptions result in the following call and put prices with K strike price and $T-t$ maturity.

$$c = S(t)N(d_1) - \exp(-r(T-t))KN(d_2) \tag{2}$$

$$p = \exp(-r(T-t))KN(-d_2) - S(t)N(-d_1),$$

where N is the cumulative distribution function of the standard normal distribution, and,

$$d_1 = \frac{\log\left(\frac{S(t)}{K}\right) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \qquad d_2 = d_1 - \sigma\sqrt{T-t}, \tag{3}$$

The most important aim of this section is to shortly introduce the Black-Scholes implied volatility smile and implied volatility surface, which are derived from the classical Black-Scholes option-pricing model of Black and Scholes (1973). The volatility smile and surface gives a basic intuition behind the option-implied measures, which is the basic of this thesis.

The Black-Scholes model assumes that the volatility parameter is constant, so it does not depend on the time to maturity or the exercise price. We can test this assumption by re-calculating the volatility parameter from option prices available in the market, so we can see what volatility parameter was "used" by the market to price the given derivative contract. This is possible since the partial "derivative" of the option price with respect to σ, which is called the vega of the option, can be calculated as

$$vega = S(t)\phi(d1)\sqrt{T-t} = K\exp(-r(T-t))\phi(d2)\sqrt{T-t}, \tag{4}$$

where ϕ is the density function of the standard normal distribution. We can see that the value of vega is always positive, which means that the option price is strictly monotonically increasing with the value of the volatility parameter. This means that the option pricing function can be inverted. So, it is possible to calculate the volatility parameter with which the Black-Scholes model would give back the option price observable in the market. This volatility parameter is called the implied volatility. If the assumption of consant volatility is correct, then the implied volality should be independent from the other variables and parameters of the model.

However, we know that in general the assumption of constant volatility is not satisfied in reality, since the implied volatility of OTM and ITM options are significantly higher than the implied volatility of ATM options. This phenomenon is called the implied volatility smile. This smile or smirk like behaviour can be experienced in many asset classes with slight differences. For equity options we experience that for call options ITM options have higher implied volatilities, than OTM options (also there are significantly less available deep OTM options than deep ITM options on the market), but for example for foreign exchange options the implied volatility is in some cases more symmetric as a function of the strike price around the ATM options.

There is also another empirical phenomenon related to the implied volatility. In general we can experience that the convexity of the smile gets smaller as we go for longer time to maturity, this means that the implied volatility depends on both the strike price and the time to maturity. This phenomenon is called the implied volatility surface. The existence of the volatility surface is often connected to stylized facts of the financial markets like the heavy-tails of the return distribution, the clustering of the volatility process, or the presence of jumps in the asset price process.

In this introductory subchapter I have shortly introduced the Black-Scholes model and the implied volatility calculated from this model. This should give a basic intuition of the option implied information. To resolve the problem of constant volatility many different types of models have emerged, for example the local volatility models like Dupire (1994), the stochastic volatility models like Heston (1993) or the jump-diffusion models like Merton (1976). However, if we calculate option implied information based on these methods, then they would be heavily based on the model specification, which would make it hard to decide whether a connection between option-implied data and stock returns are created by real market forces or only by the specification of a given model.

2 Implied moments

In this section, I will introduce the concept of risk neutral moments, which is crucial for my research. I have already defined the implied volatility in the introductory part of this work. However, it is easy to see that the value of that implied volatility is fundamentally determined by the structure of the Black-Scholes model. Anderson et al (2000) has pointed out that if we infer a measure from a misspecified model, then we can create spurious connection between the measure and our data, which is not driven by the market, but by the model itself. Also, theoretically it is commonly argued that the cause of the volatility smile is the negatively skewed and heavy-tailed return distribution. However, even if our models are well-specified, it is not clear that the relation is a general property or only driven by a specific modelling choice (for example Bakshi et al (2003) argues that parametrization can create artificial dependence on the third or fourth moment). These reasons make it very important to have a model-free way to calculate the risk neutral moments in order to get a good understanding of the connection between option-implied information and underlying returns.

In this chapter I will introduce two model-free methods to recover risk neutral moments from market data. After that I will calculate the second, third and fourth moments based on the second, more general method.

2.1 Risk neutral density function

A natural way to calculate the risk neutral moments is to first calculate the risk neutral density function and then calculate the moments by the definition of the moments of a random variable. So, to calculate the risk neutral moments in a model-free way, we only need a model-free way to calculate the risk neutral density function. Breeden and Litzenberger (1978) proposed a method to calculate the risk neutral density function from options traded on the market. In this section, I will introduce this methodology to recover risk neutral density function from call option prices available in the market based on Breeden and Litzenberger (1978) and Márkus (2017).

Let $\left(\Omega, \mathscr{F}, \mathbb{P}\right)$ be a probability space with filtration $\mathbb{F} = (\mathscr{F})_{t\in[0,T]}$ and let $f(.)$ be the payoff function of a European-type security. The first and second fundamental law of derivative pricing leads to the conclusion that the price of a security is the discounted conditional expected value with respect to the risk neutral measure, interested readers can find these theorems for example in Márkus (2017).

$$P(t) = B_t E^{\mathbb{Q}}(f(S(T))|\mathscr{F}_t) = B_0 \exp(r(T-t)) E^{\mathbb{Q}}(f(S(T))|\mathscr{F}_t) \tag{5}$$

where $P(t)$ denotes the price of the security at time t, $\mathbb{Q}$ is the risk neutral measure, r is the risk-free rate, which is assumed to be constant here, B_t is the price of a bond, which is assumed to evolve according to the deterministic $dB_t = rB_t dt$ differential equation, and $S(T)$ is the terminal value of the underlying. In a common financial modelling case, we would define a stochastic process, which could describe the dynamics of the underlying price, then this would define the distribution of the terminal value of the asset price. For example, the Black-Scholes model assumes that the stock price process follows a Geometric Brownian Motion, which implies a log-normal distribution for the asset price. However, now we reverse this approach, and we would like to find the risk neutral distribution of the asset price with which the pricing formula above gives back the price of the security observed in the market.

Definition based on Márkus(2017). *Let's assume that $B_0 = 1$. If a risk neutral distribution of the asset price process can be determined, with which the pricing formula gives back the market price of the security with $f(.)$ payoff, then the pricing formula can be expressed as*

$$B_t E^{\mathbb{Q}}(f(S(T))|\mathscr{F}_t) = \exp(r(T-t)) \int_0^{\infty} p(S(t), t, T, S(T)) f(S(T)) dS(T) \tag{6}$$

This distribution is called the implied distribution, and $p(S(t), t, T, S(T))$ is the implied density function.

So to infer the risk neutral moments, we want to calculate this implied density function. Let $\tau = T - t$ be the time to maturity and $S(t)$ be the asset price at time t. If this security is a European call option, then the price can be expressed as

$$C(S(T), t, K, T) = \exp(-r\tau) \int_K^{\infty} (S(T) - K) \cdot p(S(t), t, T, S(T)) dS(T) = \tag{7}$$

$$\exp(-r\tau) \cdot \int_0^{\infty} \mathbb{1}_{(S(T) \geq K)} \cdot (S(T) - K) \cdot p(S(t), t, T, S(T)) dS(T)$$

where $\mathbb{1}_{(S(T) \geq K)}$ is the indicator function, and K is the exercise price of the option. We multiply both sides of the equation with $\exp(r\tau)$ and we take the derivative of both sides with respect to the exercise price. The order of the integral and the derivative on the left-hand side can be exchanged according to the Leibniz-rule.

$$\exp(r\tau) \frac{\partial C}{\partial K}(C(S(t), t, T, K) = \int_0^{\infty} \frac{\partial (S(T) - K)}{\partial K} \cdot \mathbb{1}_{(S(T) \geq K)} \cdot p(S(t), t, T, S(T)) dS(T) +$$

$$\int_0^{\infty} \frac{\partial \mathbb{1}_{(S(T) \geq K)}}{\partial K} \cdot (S(T) - K) \cdot p(S(t), t, T, S(T)) dS(T) \tag{8}$$

The derivative of the indicator function is the Dirac-delta function, so the equation takes the form of

$$\exp(r\tau)\frac{\partial C}{\partial K}(S(t),t,T,K) =$$

$$\int_0^{\infty} -\mathbb{1}_{(S(T)\geq K)} \cdot p(S(t),t,T,S(T))dS(T) + \int_0^{\infty} \delta_K(S(T)) \cdot p(S(t),t,T,S(T))dS(T) =$$

$$\int_K^{\infty} -p(S(t),t,T,S(T))dS(T) + p(S(t),t,T,K)\cdot(K-K) =$$

$$-\int_K^{\infty} p(S(t),t,T,S(T))dS(T) = -1 + \int_{-\infty}^{K} p(S(t),t,T,S(T))dS(T) \tag{9}$$

Where we used the fact that by definition the Dirac-delta is a function, which satisfies that $\int_{-\infty}^{\infty} f(s)\delta_t(s) = f(t)$ for any infinitely differentiable function and that the integral of a density function on the real line is 1. Now, we take the derivative with respect to the exercise price again.

$$\exp(r\tau)\frac{\partial^2 C}{\partial K^2} = p(S(t),t,K,T) \tag{10}$$

This is called the Breeden-Litzenberger formula. So, according to this method the price of a European type security with payoff function $f(.)$ is

$$P(t) = \int_0^{\infty} \frac{\partial^2 C}{\partial K^2} f(K)dK \tag{11}$$

This concludes the calculation of the risk neutral density function, which was the main aim of this section. The method of Breeden and Litzenberger (1978), introduced here, can be used to calculate the risk neutral moments with the definition of the moments of a given random variable. However, this method has some practical problems in most realistic cases, therefore in the next section I am going to introduce another methodology to calculate the risk neutral moments without having to calculate the risk neutral density function first.

2.2 Spanning derivative securities

In this section, I am going to introduce a method to calculate the risk neutral moments without having to calculate the appropriate density function before. This method is mainly based on the general derivative pricing methodology of Bakshi and Madan (2000). Bakshi and Madan (2000) argues that a pricing method based on the state-price density does not always give closed form or analytically tractable solution in case of a general stochastic dynamics setting. For most of the realistic derivative pricing examples the exercise region of a derivative security depends on a process for which the risk neutral density function is mathematically intractable or cannot even be calculated. A good example for this phenomenon would be American options, whose exact risk neutral densities cannot be easily characterized. The source of the problem is the lack of analyticity of the payoff function. This motivated Bakshi and Madan (2000) to introduce a spanning methodology based on the characteristic function of the state price density. The main advantage of spanning payoff via characteristic function is that the function (in most cases) is infinitely differentiable. This property makes for example the boundary condition of the Black-Scholes differential equation (in terms of the characteristic function) differentiable and mathematically tractable as shown by Bakshi and Madan (2000). I will calculate a simple example of replicating a general payoff structure, which is a direct consequence of the general result of spanning derivative securities, based on the assumptions of Bakshi and Madan (2000). However, the general methodology is not directly related to this thesis, interested readers are advised to read more details about the framework of Bakshi and Madan (2000) in the original article.

The problem with using the previous method to calculate the risk neutral moments is that we have to calculate the risk neutral density first. If the risk neutral density is intractable or cannot be calculated, then it would be impossible to calculate risk-neutral moments. To overcome this difficulty Bakshi et al (2003) proposed another model-free method to compute the risk-neutral moments. This method is based on the general option pricing methodology of Bakshi and Madan (2000) and use the fact that any payoff with bounded expectations can be spanned using positions through strike prices. The authors aim to represent features of the risk neutral distribution with option prices available in the market. They recover the risk neutral variance, skewness and kurtosis in a model-free manner with only OTM call and put options. In the next theorem, I will present a simplified case of replicating a general derivative security with linear combination of options plus underlying and risk-free bond positions.

Theorem based on Bakshi et al (2003). *Let $S(T)$ be the terminal value of the underlying asset, and let f be a payoff function integrable with respect to the risk neutral density. Assume that f is twice continuously differentiable, then for any fixed $S(t)$*

$$f(S(T)) = f(S(t)) + (S(T) - S(t))f_S(S(t)) \tag{12}$$

$$+ \int_{S(t)}^{\infty} f_{SS}(K)(S(T) - K)^+ dK + \int_0^{S(t)} f_{SS}(K)(K - S(T))^+ dK,$$

where f_S denotes the first order derivative of the payoff function.

Proof. The basic idea of the proof comes from Carr and Madan (2001). The fundamental theorem of calculus for any fixed $S(t)$ is the basis for the calculations below

$$f(S(T)) = f(S(t)) + \int_{S(t)}^{S(T)} f_S(u)du = f(S(t)) + \mathbb{1}_{S(T)>S(t)} \int_{S(t)}^{S(T)} f_S(u)du$$

$$-\mathbb{1}_{S(T)<S(t)} \int_{S(T)}^{S(t)} f_S(u)du = f(S(t)) + \mathbb{1}_{S(T)>S(t)} \int_{S(t)}^{S(T)} \Big[f_S(S(t)) + \int_{S(t)}^{u} f_{SS}(K)dK\Big]du$$

$$-\mathbb{1}_{S(T)<S(t)} \int_{S(T)}^{S(t)} \Big[f_S(S(t)) - \int_{u}^{S(t)} f_{SS}(K)dK\Big]du \tag{13}$$

Now, integrating out the parts, which do not depend on u, and applying the Fubini theorem results in

$$
\begin{aligned}
f(S(T)) &= f(S(t)) + \mathbb{1}_{S(T)>S(t)} f_S(S(t))(S(T)-S(t)) - \mathbb{1}_{S(T)<S(t)} f_S(S(t))(S(t)-S(T)) \\
&+\mathbb{1}_{S(T)>S(t)} \int_{S(t)}^{S(T)} \int_{K}^{S(T)} f_{SS}(K) du dK + \mathbb{1}_{S(T)<S(t)} \int_{S(T)}^{S(t)} \int_{S(T)}^{K} f_{SS}(K) du dK \\
&= f(S(t)) + f_S(S(t))(S(T)-S(t)) + \mathbb{1}_{S(T)>S(t)} \int_{S(t)}^{S(T)} \int_{K}^{S(T)} f_{SS}(K) du dK \\
&+\mathbb{1}_{S(T)<S(t)} \int_{S(T)}^{S(t)} \int_{S(T)}^{K} f_{SS}(K) du dK
\end{aligned} \tag{14}
$$

Now, calculating the first integrals and rearranging the indicator functions gives the replicating equation and concludes the proof

$$
\begin{aligned}
f(S(T)) &= f(S(t)) + f_S(S(t))(S(T)-S(t)) + \mathbb{1}_{S(T)>S(t)} \int_{S(t)}^{S(T)} f_{SS}(K)(S(T)-K) dK \\
&+\mathbb{1}_{S(T)<S(t)} \int_{S(T)}^{S(t)} f_{SS}(K)(K-S(T)) dK \\
&= f(S(t)) + f_S(S(t))(S(T)-S(t)) + \int_{S(t)}^{\infty} f_{SS}(K)(S(T)-K)^+ dK \\
&+ \int_{0}^{S(t)} f_{SS}(K)(K-S(T))^+ dK
\end{aligned} \tag{15}
$$

□

We know from no-arbitrage pricing theory that the price of a security is the risk the discounted risk neutral conditional expectation with respect to the risk neutral measure (Shreve (2008) or Márkus (2017)). Based on the theorem above, and the fact that the discounted underlying price is a martingale under the risk neutral measure, the price at time t is

$$\exp(-r(T-t))E^{\mathbb{Q}}(f(S(T)|\mathscr{F}_t)) = \exp(-r(T-t))f(S(t)) + S(t)f_S(S(t))$$
$$- \exp(-r(T-t))f_S(S(t))S(t) + \int_{S(t)}^{\infty} f_{SS}(K)C(t,T,K)dK + \int_{0}^{S(t)} f_{SS}(K)P(t,T,K)dK$$
$$= \exp(-r(T-t))(f(S(t)) - f_S(S(t))S(t)) + S(t)f_S(S(t)) + \int_{S(t)}^{\infty} f_{SS}(K)C(t,T,K)dK$$
$$+ \int_{0}^{S(t)} f_{SS}(K)P(t,T,K)dK \tag{16}$$

Where we have used that the price of a call is $C(t,T,K) = \exp(-r(T-t))E^{\mathbb{Q}}((S(T)-K)^+|\mathscr{F}_t)$ and the price of the put option is $P(t,T,K) = \exp(-r(T-t))E^{\mathbb{Q}}((K-S(T))^+|\mathscr{F}_t)$. Economically the formula shows that the payoff $f(S(T))$ can be replicated with $f(S(t)) - f_S(S(t))S(t)$ position in a zero-coupon bond, $f_S(S(t))$ position in the underlying and $f_{SS}(K)$ positions in a linear combination of OTM call and put options.

In this section, I have introduced a model-free method to replicate payoff functions that are twice differentiable based on the general result of Bakshi and Madan (2000). In the next section, I am going to use this result to calculate the risk neutral moments without first having to explicitly calculate the risk neutral density function. After that, I will present an alternative definition of these moments, which are commonly used in the academic literature based on Carr and Wu (2008) and Kozhan et al (2013).

2.3 Implied variance

Now, I will calculate the second moment of risk neutral distribution based on the formula from the previous section, and I will shortly introduce a market instrument called variance swap based on Carr and Wu (2008). I note here to avoid any misconceptions that I will refer to the risk neutral second moment as the risk neutral variance. The commonly used assumptions behind this it that the mean of log returns is zero (for example from Cont (2001)). The notion of implied variance is in close relation to the variance swap contract. Variance swap is a contract which gives clean exposure to the behaviour of market volatility. The trading parties of the variance swap agree to exchange the realized variance of an underlying over the lifetime of the swap for a pre-agreed strike. Therefore, the payoff of the contract is

$$Payoff = Notional \cdot [RV_{t,T} - SW_{t,T}] \tag{17}$$

The floating leg of the swap is defined based on Bollerslev et al (2009) as

$$RV_{t,T} = \frac{252}{T} \sum_{k=1}^{n} \log\left(\frac{S_k^2}{S_{k-1}^2}\right) \tag{18}$$

There are other ways to define the annualized realized variance in the literature, but we will use this one during this thesis. The main reasons to use this realized variance definition is explained in the next chapter about realized moments. According to the no-arbitrage argument at inception the value of a swap is naturally zero, so the floating leg will be

$$SW_{t,T} = E^{\mathbb{Q}}(RV_{t,T}) \tag{19}$$

The implied variance is referred to as the market's risk neutral expectations of return variation. This is the reason why the fixed leg of the variance swap is the implied variance. From the definition of the realized variance, we can read the payoff function that should be used to create the replicating portfolio of the previous section. Now, to calculate the implied variance the payoff function will be

$$f(S(T)) = \left(\log\left(\frac{S(T)}{S(t)}\right)\right)^2 = r^2(t,T) \tag{20}$$

In the next theorem, I will calculate the price of this volatility contract defined by this payoff function based on Bakshi et al (2003).

Theorem. *The price of the volatility contract, which has the payoff structure* $f(S(T)) = r^2(t,T)$ *is*

$$V(t,T) = \exp(r(T-t))E^{\mathbb{Q}}(RV_{r,T}) = \int_{S(t)}^{\infty} \frac{2(1-\log\left(\frac{K}{S(t)}\right))}{K^2} C(t,T,K)dK + \int_{0}^{S(t)} \frac{2(1+\log\left(\frac{S(t)}{K}\right))}{K^2} P(t,T,K)dK \tag{21}$$

Proof. Let's collect all parts of the replicating formula

$$f_S(K) = 2\log\left(\frac{K}{S(t)}\right)\frac{1}{K} \tag{22}$$

$$f_{SS}(K) = \frac{2(1-\log\left(\frac{K}{S(t)}\right))}{K^2} \tag{23}$$

The expressions before the integrals are evaluated at $f_S(S(t))$, so the logarithmic return will be zero, which makes the whole analytical part zero. Finally, the replication of the volatility contract is obtained by plugging in the second derivatives to the integrals.

$$V(t,T) = \int_{S(t)}^{\infty} \frac{2(1-\log\left(\frac{K}{S(t)}\right))}{K^2} C(t,T,K)dK + \int_{0}^{S(t)} \frac{2(1+\log\left(\frac{S(t)}{K}\right))}{K^2} P(t,T,K)dK \tag{24}$$

□

Economically, the significance of the above theorem is that we do not need risk-free bonds or underlying positions to replicate the volatility contract. It means that taking long positions in a linear combination of OTM call and put options is enough to recover the implied variance from the market. This also means that the fixed leg of a variance swap would be the annualized linear combination of OTM options. The weighting function attached to options with different strikes is $\frac{2(1-\log\left(\frac{K}{S(t)}\right))}{K^2}$, which gives far higher weights to deep OTM options, than near ATM options. The intuition behind this structure is that when we fit an implied volatility curve in practise we can experience a parabolic shape in the space of moneyness and implied volatility.

2.4 Implied skewness

Now, I will calculate the third moment of risk neutral distribution based on the formula from Bakshi et al (2003). Kozhan et al (2013) introduces a trading strategy that is the purest bet on market skewness, called skew swap. The structure of this skew swap is similar to the market convention of the variance swap introduced in the previous section.

The payoff function used for calculating the risk neutral skewness is

$$f(S(T)) = \left(\log\left(\frac{S(T)}{S(t)}\right)\right)^3 = r^3(t,T) \tag{25}$$

In the next theorem, I will calculate the price of this skewness contract defined by this payoff based on Bakshi et al (2003)

Theorem. *The price of the skewness contract, which has the payoff structure* $f(S(T)) = r^3(t,T)$ *is*

$$\begin{aligned} W(t,T) = &\int_{S(t)}^{\infty} \frac{6\log\left(\frac{K}{S(t)}\right) - 3\left(\log\left(\frac{K}{S(t)}\right)\right)^2}{K^2} C(t,T,K)dK \\ &- \int_0^{S(t)} \frac{6\log\left(\frac{S(t)}{K}\right) + 3\left(\log\left(\frac{S(t)}{K}\right)\right)^2}{K^2} P(t,T,K)dK \end{aligned} \tag{26}$$

Proof. Let's collect all parts of the replicating formula

$$f_S(K) = 3\left(\log\left(\frac{K}{S(t)}\right)\right)^2 \frac{1}{K} \tag{27}$$

$$f_{SS}(K) = \frac{6\log\left(\frac{K}{S(t)}\right) - 3\left(\log\left(\frac{K}{S(t)}\right)\right)^2}{K^2} \tag{28}$$

The expressions before the integrals are evaluated at $f_S(S(t))$, so the logarithmic return will be zero, which makes the whole analytical part zero. Finally, the replication of the skewness contract is obtained by plugging in the second derivatives to the integrals.

$$W(t,T) = \int_{S(t)}^{\infty} \frac{6\log\left(\frac{K}{S(t)}\right) - 3\left(\log\left(\frac{K}{S(t)}\right)\right)^2}{K^2} C(t,T,K)dK$$
$$- \int_{0}^{S(t)} \frac{6\log\left(\frac{S(t)}{K}\right) + 3\left(\log\left(\frac{S(t)}{K}\right)\right)^2}{K^2} P(t,T,K)dK \tag{29}$$

□

Economically, the significance of the above theorem is that we do not need risk-free bonds or underlying positions to replicate the skewness contract. We can replicate the implied skewness by taking long position in a linear combination of OTM call options, and taking short positions in a linear combination of OTM put options. This contract allows a clean play on skewness since if the return distributions have negative skewness, then the OTM put options are priced at a premium compared to OTM call options.

2.5 Implied kurtosis

Now, I will calculate the fourth moment of risk neutral distribution based on the formula from Bakshi et al (2003). The implied kurtosis has much less literature compared to the previous two measures. The reason for this is that most empirical results show that it has no effect on the underlying returns. Two articles of the most important literature examining the implied kurtosis are Amaya et al (2015) and Harris and Qiao (2018), who calculated the measure based on Bakshi et al (2003) and showed that the so-called kurtosis risk premium has no relation to stock returns. I will also test these results in my work, so here I will introduce the concept of implied kurtosis first.

The payoff function used to calculate the implied kurtosis is $f(S(T)) = r^4(t,T)$. The next theorem calculates the implied kurtosis, like with the other two moments in the previous sections.

Theorem. *The price of the kurtosis contract, which has the payoff structure* $f(S(T)) = r^4(t,T)$ *is*

$$X(t,T) = \int_{S(t)}^{\infty} \frac{12\Big(\log\Big(\frac{K}{S(t)}\Big)\Big)^2 - 4\Big(\log\Big(\frac{K}{S(t)}\Big)\Big)^3}{K^2} C(t,T,K)dK$$
$$+ \int_{0}^{S(t)} \frac{12\Big(\log\Big(\frac{S(t)}{K}\Big)\Big)^2 + 4\Big(\log\Big(\frac{S(t)}{K}\Big)\Big)^3}{K^2} P(t,T,K)dK \tag{30}$$

Proof. Let's collect all parts of the replicating formula

$$f_S(K) = 4\Big(\log\Big(\frac{K}{S(t)}\Big)\Big)^3 \frac{1}{K} \tag{31}$$

$$f_{SS}(K) = \frac{12\Big(\log\Big(\frac{K}{S(t)}\Big)\Big)^2 - 4\Big(\log\Big(\frac{K}{S(t)}\Big)\Big)^3}{K^2} \tag{32}$$

The expressions before the integrals are evaluated at $f_S(S(t))$, so the logarithmic return will be zero, which makes the whole analytical part zero. Finally, the replication of the kurtosis contract is obtained by plugging in the second derivatives to the integrals.

$$X(t,T) = \int_{S(t)}^{\infty} \frac{12\Big(\log\Big(\frac{K}{S(t)}\Big)\Big)^2 - 4\Big(\log\Big(\frac{K}{S(t)}\Big)\Big)^3}{K^2} C(t,T,K)dK$$
$$+ \int_{0}^{S(t)} \frac{12\Big(\log\Big(\frac{S(t)}{K}\Big)\Big)^2 + 4\Big(\log\Big(\frac{S(t)}{K}\Big)\Big)^3}{K^2} P(t,T,K)dK \tag{33}$$

□

This means that quartic payoff can be replicated by long positions in a linear combination of OTM call and put options. In these sections I have calculated three implied moments, where the payoff of the respective contracts were defined by the appropriate polynomial of the logarithmic return of the underlying. However, there are ways to define these implied measures in the literature. I will examine these alternative definitions in the next section.

2.6 Alternative formulations

In the academic literature on empirical asset pricing other definitions for the risk neutral moments have appeared mainly for the risk neutral variance and the risk neutral skewness in Carr and Wu (2008) and Kozhan et al (2013). In this section, I will argue, why I have used the definitions presented in the previous sections. I will illustrate the main difference based on the implied variance measure. The main difference stems from the definition of the payoff function used to represent a given implied moment. Kozhan et al (2013) define the payoff function as $f(S(T)) = -2\log\left(\frac{S(T)}{S(t)}\right)$.This results in the following replication equation.

$$V(t,T) = \int_{S(t)}^{\infty} \frac{2}{K^2} C(t,T,K)dK + \int_{0}^{S(t)} \frac{2}{K^2} P(t,T,K)dK \tag{34}$$

Carr and Wu (2008) arrives to the same result with a slightly different methodology. This method to calculate the implied variance is the same as the one used by the Chicago Board of Options Exchange (CBOE) to calculate the VIX index, which is the most well-known indicator of expected return variance. Here, I will show that this definition of implied variance contains information about the higher moments, not just the second moments. This part is based on Du and Kapadia (2012), Chow et al (2014) and Chow et al (2016), who criticize the formulation of the VIX index for the same reason.

First, let $R_T = \frac{S(T)-S(t)}{S(t)}$ be the arithmetic return and $r_T = \log\left(\frac{S(T)}{S(t)}\right)$ be the logarithmic return of the underlying. Let's examine the first order Taylor series expansion of $\log(S(T))$ with integral remainder.

$$\log(S(T)) = \log(S(t)) + \frac{S(T)-S(t)}{S(t)} + \int_{S(t)}^{\infty} \frac{-1}{K^2}(S(T)-K)^+ dK + \int_{0}^{S(t)} \frac{-1}{K^2}(K-S(T))^+ dK \tag{35}$$

Rearranging the terms results in

$$R_T - r_T = \int_{S(t)}^{\infty} \frac{1}{K^2}(S(T) - K)^+ dK + \int_0^{S(t)} \frac{1}{K^2}(K - S(T))^+ dK \tag{36}$$

Taking the risk neutral expectation of the both sides results in

$$\begin{aligned} E^{\mathbb{Q}}(R_T) - E^{\mathbb{Q}}(r_T) = \exp(r(T-t))\Big(& \int_{S(t)}^{\infty} \frac{1}{K^2} C(t,T,K) dK \\ & + \int_0^{S(t)} \frac{1}{K^2} P(t,T,K) dK \Big) \end{aligned} \tag{37}$$

Now, we want to connect this equation to the risk neutral moments, to achieve this we apply the Taylor series expansion of the exponential function.

$$\begin{aligned} (1 + R_T) = \frac{S(T)}{S(t)} = \exp\Big(\log\Big(\frac{S(T)}{S(t)}\Big)\Big) = & \sum_{k=0}^{N} \frac{1}{k!}\Big(\log\Big(\frac{S(T)}{S(t)}\Big)\Big)^k \\ & + o\Big(log\Big(\frac{S(T)}{S(t)}\Big)\Big)^N \end{aligned} \tag{38}$$

The first and the second parts of the sum on the right side is 1 and $r_T = log\Big(\frac{S(T)}{S(t)}\Big)$ respectively. If we take these parts to the other side and take the risk neutral expected value, we get the same expression as in the previous equation

$$E^{\mathbb{Q}}(R_T) - E^{\mathbb{Q}}(r_T) = \sum_{k=2}^{N} \frac{1}{k!} E^{\mathbb{Q}}(r_T)^k + o\Big(E^{\mathbb{Q}}(r_T)^N\Big) \tag{39}$$

So to combine these equations, and factor out parts of the sum, we can see that

$$\begin{aligned} & \exp(r(T-t))\Big(\int_{S(t)}^{\infty} \frac{1}{K^2} C(t,T,K) dK + \int_0^{S(t)} \frac{1}{K^2} P(t,T,K) dK\Big) \\ & = \frac{1}{2} E^{\mathbb{Q}}(r_T^2) + \frac{1}{6} E^{\mathbb{Q}}(r_T^3) + \frac{1}{24} E^{\mathbb{Q}}(r_T^4) + o\Big(E^{\mathbb{Q}}(r_T^4)\Big) \\ & = \frac{1}{2} VIX^2 \end{aligned} \tag{40}$$

The main consequence of this calculation is that this replicating strategy contains information about not only the second moment of the risk neutral distribution, but also about higher moments. This can cause a considerable bias for this definition of the implied variance since it is documented that skewness has a prominent role in financial markets. For example if the price of the underlying is on a downward trend, then demand for put options increase and their price becomes relatively higher, than the price of call options. Hence, the value of the volatlity contract, which can be replicated by long positions in options, increases with the put premia. However, the skew contract can be replicated by long positions in call options and short positions in put options, therefore in this case its value decreases. If both effects are incorporated in our indicator, then we get a biased measurment of the change of expected volatility in the market. Chow et al (2014) show that the definitions based on the result of Bakshi et al (2003) is unbiased, that's why I have decided to use that definition. Carr and Wu (2008) and Du and Kapadia (2012) shows that the difference between the two described measures is proportional to the jump intensity. This means that this method would create a biased estimate for the whole family of Levy-type models. Carr and Wu (2008) also show that the two measure are identical, if the underlying follows a diffusion process. However, Du and Kapadia (2012) explicitly calculates the difference for many popular financial models, like the Merton jump diffusion model (Merton, 1976) to show the dependence of the difference on the jump intensity.

Naturally, we do not have available data for continous options neither in my database nor on the real market, therefore we have to apply a discretization scheme. I chose to use a method commonly used in the literature and is presented in a detailed way in Kozhan et al (2013). Using this discretization scheme the summation is carried out along the available strike prices, the implied skewness would be calculated as

$$IS_{t,T} = 2\sum_{K_i<S(t)}\left(\frac{6\log\left(\frac{K_i}{S(t)}\right) - 3\left(\log\left(\frac{K}{S(t)}\right)\right)^2}{K_i^2}C(t,T,K_i)\Delta I(K_i)\right)$$
$$+\sum_{K_i>S(t)}\left(\frac{6\log\left(\frac{S(t)}{K_i}\right) + 3\left(\log\left(\frac{S(t)}{K_i}\right)\right)^2}{K_i^2}P(t,T,K_i)\Delta I(K_i)\right) \quad (41)$$

,where $I(K_i)$ are the integrators defined as

$$\Delta I(K_i) = \begin{cases} \dfrac{K_{i+1} - K_{i-1}}{2}, & \text{if } 0 \leq i \leq N \\ 0, & \text{otherwise} \end{cases} \tag{42}$$

This concludes the section, but we will return to the results in the chapter about Other Option-Implied Factors, since there will be factors that can be defined from these results.

3 Moment Risk Premia

In this chapter I am going to introduce the concept of moment risk premia based on the definitions of the respective implied moments. In the Empirical Literature chapter, I will give a detailed introduction about most of the academic results, which have shown that it is not the level of a given implied moment, but rather the respective risk premia, which has strong predictive power for the underlying returns. This is the reason, why this chapter is dealing with the concept of moment risk premia. In the chapter first I will introduce the variance risk premium, then the skew risk premium and the kurtosis risk premium. The end of the chapter is dealing with the empirical properties of these risk premia.

3.1 Variance Risk Premium

In this section I will introduce the concept of the variance risk premium. I will show the definition of the variance risk premium and also give some commonly used explanations of the existence of the premium. The variance risk premium is the difference between the realized variance and the implied variance $VRP = RV - IV$ (Carr and Wu, 2008). In the previous chapter I have introduced the variance swap market instrument. Carr and Wu (2008) was the first to quantify the variance risk premium with the help of the variance swap. Their aim was to find the source of the variance risk premium. Carr and Wu (2008) quantifies the VRP using a pricing kernel. I will show their argument now. We have seen that the fixed leg of the variance swap is

$$SW_{t,T} = E^{\mathbb{Q}}(RV_{t,T}) \tag{43}$$

Using the theory of measure change, we can see that for a pricing kernel $M_{t,T}$

$$SW_{t,T} = E^{\mathbb{Q}}(RV_{t,T}) = \frac{E^{\mathbb{P}}(M_{t,T}RV_{t,T})}{E^{\mathbb{P}}(M_{t,T})}, \tag{44}$$

where $\mathbb{P}$ denotes the physical probability measure. Duffie (1992) showed that no-arbitrage argument guarantees that there exists at least one such pricing kernel. Let's introduce $m_{t,T} = \frac{M_{t,T}}{E^{\mathbb{P}}(M_{t,T})}$, and substitute it into the equation (44) $SW_{t,T} = E^{\mathbb{P}}(m_{t,T}RV_{t,T})$. The expected value of $m_{t,T}$ is 1, so using the well-known $Cov(X,Y) = E(XY) - E(X)E(Y)$ formula for the covariance of two random variables X and Y, we can decompose the fixed leg of the variance swap as

$$SW_{t,T} = E^{\mathbb{P}}(m_{t,T}RV_{t,T}) = E^{\mathbb{P}}(RV_{t,T}) + Cov^{\mathbb{P}}(m_{t,T}, RV_{t,T}), \tag{45}$$

The first term is the time-series conditional mean of the realized variance, so the negative of the covariance term defines the average return variance risk premium as this is the difference between the average realized variance and the implied variance. However, if we only follow this methodology, then we would only get an estimate of the average profit and loss of a strategy handling large variety of variance swap positions. We rather calculate the individual variance risk premium for every equity in our data, but the calculation above gives a basis for understanding the existence of the variance risk premium.

The only thing remained to fully define the variance risk premium is to define the calculation of the realized variance. I will use the result of Andersen et al (2001), Andersen et al (2003) and Barndorff-Nielsen and Shephard (2002), who show that summing squared high-frequency returns we get an unbiased estimator of the integrated variance. These papers show mathematically that by increasing the sampling frequency, the sum of squared returns is converging to the integrated variance with probability 1. However, they also show that in reality microstructural effects are causing large noise in small frequency data, which makes the econometric analysis hard. Most papers advise to use five-minute return data. Following this logic, I construct the realized variance, as the most common method in the literature, as the sum of squared returns for a given time frequency $\frac{1}{n}$

$$RV_{t,T} = \sum_{i=1}^{n} r_i^2 \tag{46}$$

The existence of the variance risk premium is often explained by excess demand for protection compared to the supply of protection. The negative sign before the covariance is explained as the buyer of the variance swap, who pays the fixed leg, is willing to pay a premium in order to be protected against the increase of volatility in the market. This conclusion is also supported by Bakshi and Madan (2000), who express the variance risk premium as a function of the risk aversion parameter of a representative trading agent scenario.

3.2 Skewness Risk Premium

In this section I will introduce the concept of the skew risk premium. I will show the definition of the skew risk premium and give some commonly used explanations of the existence of the premium. Kozhan et al (2013) defines a trading strategy that is behaving for the market skewness just as the variance swap behaves to the variance, therefore it offers a clean exposure on the evolution of the skewness on the market. They refer to this trading strategy as the skew swap. The fixed leg of the contract is the implied skewness, while the floating leg is the realized skewness during the lifetime of the contract.

$$Payoff = Notional[RS_{t,T} - IS_{t,T}], \tag{47}$$

where IS is the implied skewness calculated in the previous chapter. The calculation of the realized skewness is carried out by the results of Amaya et al (2015), who use

$$RS_{t,T} = \sqrt{n} \sum_{i=1}^{n} r_i^3 \tag{48}$$

as the measure of realized skewness. This measure tries to capture the cubic variation process, which is zero for a pure diffusion process. If the cubic variation is significantly different from zero, then it is highly probable that there is stochastic volatility or a jump component in the underlying process. Amaya et al (2015) notes that the multiplier before the sum is needed to have the implied and the realized skewness on the same scale.

The skew risk premium is defined as the difference between the implied skewness and the realized skewness $SRP_{t,T} = RS_{t,T} - IS_{t,T}$. Kozhan et al (2013) also showed there is a significant difference between their implied skewness and realized skewness, so they define the skew risk premium the same way as the variance risk premium is defined.

Kozhan et al (2013) also showed that the difference is positive, which means that usually investors expect more negative skewness, than what is realized in the market previously and are willing to pay a premium to be insured against high negative skewness.

3.3 Kurtosis Risk Premium

In this section I will introduce the concept of the kurtosis risk premium. The implied kurtosis and kurtosis risk premium is far less documented, than the previous two measures. The reason for this is that many empirical research have come to the conclusion that no moments that are higher the third moment has predictive power in case of equity returns. This is true for both the level of implied moments and the respective risk premia. However, as I will mention in the Empirical Literature chapter, there are studies, which show that measures proxying the implied kurtosis does have effect on the dynamics of future returns.

The kurtosis risk premium is defined in the same way as the two previous risk premia, so we take the difference between the realized kurtosis and implied kurtosis

$$KRP_{t,T} = RK_{t,T} - IK_{t,T} \tag{49}$$

The realized kurtosis is defined folloing Amaya et al (2015)

$$RK_{t,T} = n \sum_{i=1}^{n} r_i^4 \tag{50}$$

Amaya et al (2015) also notes that the n scaling factor is needed to make sure that the measure has an appropriate magnitude.

In the previous three sections I have introduced three moment risk premia. For my empirical analysis I have calculated them for monthly maturity and used this to present the statistical properties of the risk premia in the next section and the empirical analysis in Chapter 7. However, in the last chapter I will perform a robustness check with two different risk premia maturity to check whether the results are sensible to a given variable construction.

3.4 Empirical Properties

In this section, I will show some empirical properties of the implied moments and the moment risk premia. Most importantly I will collect some empirical regularities mentioned separately in some of the academic research papers on option implied moments and empirical asset pricing. I will also test these empirical facts in my own database. There are various stylized facts regarding the empirical behaviour of stock returns. Probably the most important article in this topis is Cont (2001), who collects and tests the previously examined stylized facts like volatility clustering, leverage effect, gain and loss asymmetry or the conditional heavy tails, together in a consistent way. However, according to my knowledge none has tried to achieve something similar for this option-implied information. Naturally, there would be much more possible empirical property to analyse, than the ones I have, but this could be a complete topic of a separate work. In this chapter I will examine whether the implied variance, skewness and kurtosis are priced and whether these are time-varying or not for the equities and equity indices, which are present in my dataset?

3.4.1 Are the Moment Risk Premia Priced?

I will state that a given moment risk premia is priced on average, if the sample averages of a given daily annualized implied moment is significantly different from the respective realized moment. For example the variance risk premium is considered as priced risk premium, if the fixed leg of the variance swap contract is significantly different from the floating leg of the instrument.

In the appendix Table 10 presents the required statistical properties of the three implied risk premia for all equities and equity indices. The first columns contain the average value of the respective monthly risk premia, the second columns contain their standard deviation, while the third value is computed in order to decide, whether the risk premia is priced or not. Since most of the premia series are exercising significant serial autocorrelation, the t-values are computed using the methodology of Newey and West (1986), which can account for heteroscedasticity and autocovariance. We can conclude from Table 10 that both for S&P500 index (SPX) and the S&P100 index (OEX) the VRP is significantly negative with large absolute t-values, which means the investors are willing to pay premium in order to be defended against the changing volatility in the market. This is consistent with the result of Carr and Wu (2008), who shows that for the S&P500 index, the S&P100 index, the Dow Jones Industrial Index and for the NASDAQ-100 index the VRP is significantly negative with strong t-values. The sign and the significance of the VRP for individual equity premia are not so unambiguous, which is also consistent with the result of Carr and Wu (2008). However, we can see that for the majority of the equities the VRP is priced and is negative. So, we can conclude that on average the investors are willing to pay this price for the protection even for individual equities.

Cont (2001) establishes that in general the skewness of the return distribution is negative, which implies that large negative stock moves are more probable, than large positive moves. This finding is supported in my sample as both the realized and the implied skewness is negative in most cases. In the table the values are displayed as percentages in order to be comparable. For the indices we can see that the implied skewness is bigger in absolute terms, than the realized measure, which makes the risk premium positive. This result is consistent with the result of Kozhan et al (2013). For individual equities the sign of the risk premium is more ambiguous, more equities seem to be positive, therefore the market expects higher skewness on the market, than the realized one. It is also not unambigous, whether the skewness risk premium is priced on average. However, this does not mean that the skewness premium has no effect for the return generating process as we will see later. We can also conclude from Table 10 that the properties of the kurtosis risk premium are quite similar to that of the skewness risk premium.

3.4.2 Are the Moment Risk Premia time-varying?

In this section I will examine whether these moment risk premia are time-varying or not and also describe some time series properties of the evolution of the premia.

Based on Carr and Wu (2008) I will carry out this analysis by fitting an expectation hypothesis regression for all three of the moment risk premia. This time series regression takes the following form for the variance risk premium

$$RV_{t,T} = a + bIV_{t,T} + u \tag{51}$$

The appendix contains the expectation hypothesis regression results for all three moment risk premia for every equity in my dataset. In order to check the hypothesis that a moment risk premium is time-varying, we have to analyse the null hypothesis for the expectation hypothesis regression that $a = 0$ and b=1. If $a = 0$ and $b = 1$, then the given variance risk premium is zero and constant in time. If the slope is not significantly different from one, then the realized moment and the option-implied moment are changing synchronously, which means the respective premium calculated from them is constant in time. The tables in the appendix contain the parameter estimates for the a and b coefficients, the Newey-West t-statistics, and the 95% confidence interval for the beta, from which we can conclude whether the coefficient is significantly different from 1 with 5% significance level.

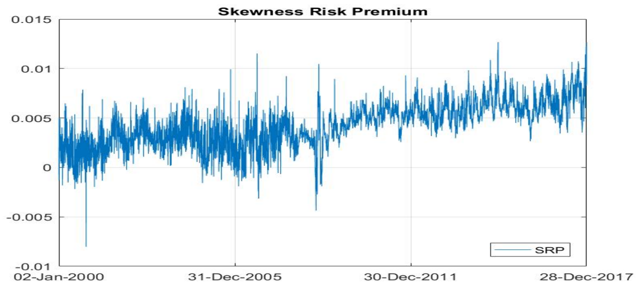
If we analyse Table 11, we can see that for that variance risk premium both indexes are strongly time-varying and different from zero, which coincides with the result of Carr and Wu (2008), who also showed this feature for the Dow Jones Industrial Index and the NASDAQ-100 index. If we analyse the result for individual equities, we can see that for all of them we can reject both null hypotheses, so we conclude that the VRP is time-varying for every element in my sample. To illustrate this property I plot the evolution of the variance risk premium for the S&P500 index below.

About the skew risk premium, we can conclude from Table 12 in the Appendix that both indeces the skew risk premium is time-varying, confirming the results of Kozhan et al (2013). However, the results for the individual equities are not as unambiguous as for the VRP. We can conclude that in total 20 out of the 26 elements from our sample is time-varying, for the remaining 6 we cannot reject the null hypothesis that the SRP is constant in time. Although, the S&P500 has time-varying skew risk premia

Figure 1: SPX VRP series

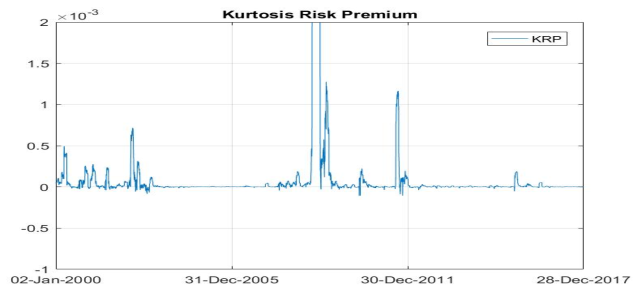
plot below for the S&P500 SRP illustrates the difference between the VRP and the SRP.

Figure 2: SPX SRP series



Lastly, we have to analyse the kurtosis risk premium. We can draw the least meaningful conclusion from this regression model about the general dynamics of the risk premia. The reason for this is that exactly half of the sample has time-varying KRP, while the other half has constant KRP.

Figure 3: SPX KRP series



Concluding this section, we can regard the VRP and SRP time-varying for most cases. However, the conclusion about KRP is not clear. These results are strongly in line with the results of Carr and Wu (2008) about the variance risk premium, Kozhan et al (2013) and Harris and Qiao (2018). We can also see that all three moment risk premia reacted to the financial crisis of 2008 (although it seems that it affected the SRP less than the two other). There are multiple other possible empirical properties that can be analyzed, for example the correlation between the premia, which was the main topic of Kozhan et al (2013), or the time series properties of the risk premia. However, as I mentioned in the beginning of the chapter, the whole empirical analysis of the individual series could be a topic of a completely different work.

4 Other Option-Implied Factors

In this chapter I am going to introduce some other measures, which are derived from the market data of options and their volatility surfaces. These measures are documented to have individual predictive power for the underlying returns. However, according to my knowledge they have never been examined together. Some of these measures try to quantify the same feature of the risk neutral distribution.

In the empirical part of the thesis, I will try to be consistent in choosing, which type of measures are used together. Most importantly, I will try to examine, which of these measures contain independent explanatory or predictive power from the moment risk premia defined previously. Also some of these measures rather give a clear view, whether higher order moments should be examined in the predictive framework.

4.1 Upside and Downside Variance Risk Premia

Feunou et al (2017) explored the so-called implied semivariances, and the semivariance premia. The basis of their argument is that numerous studies, such as Bonomo et al (2010) and Rossi and Timmermann (2015) have shown that there is an asymmetry in risk-return trade-off between responses to positive and negative market scenarios. Intuitively, investors tend to hedge against "bad uncertainty" or downward shocks and not hedge against "good uncertainty" or positive shocks, since there are significantly more long positions on equity markets.

Feunou et al (2017) was inspired by these findings to introduce the upside and downside variance risk premia. The authors argue that variance risk premium cannot give a clear view of the effect of the option implied measures to stock return, since they mix these two opposing features. In order to overcome this, they distribute the positive and negative parts of the variance risk premium into two separate variables. Feunou et al (2017) notes that the findings in the beginning of this section would imply that the variance risk premium is mainly driven by the downward variance risk premium, while the effect of the upside variance risk premium is more subdued. In this section, I will introduce the upside and downside realized and implied moments and the resulting variance risk premia.

At first, I will shortly introduce this decomposition method with the payoff function of the volatility contract, which would define the implied semivariances needed to calculate the semivariance risk premia. In the second chapter, we have seen that the security with twice-continuously differentiable payoff function can be spanned by positions in bond, underlying and OTM call and put options as

$$\begin{aligned} f(S(T)) &= f(S(t)) + (S(T) - S(t)) f_S(S(t)) \\ &+ \int_{S(t)}^{\infty} f_{SS}(K)(S(T)-K)^+ dK + \int_0^{S(t)} f_{SS}(K)(K - S(T))^+ dK, \end{aligned} \tag{52}$$

We have also seen that the arbitrage-free price of this claim is

$$\begin{aligned} &\exp(-rT) E^{\mathbb{Q}}(f(S(T)|\mathscr{F}_t)) \\ &= \exp(-rT)(f(S(t)) - f_S(S(t))S(t)) + S(t) f_S(S(t)) + \int_{S(t)}^{\infty} f_{SS}(K) C(t,T,K) dK \\ &+ \int_0^{S(t)} f_{SS}(K) P(t,T,K) dK \end{aligned} \tag{53}$$

Feunou et al (2017) defines the payoff used for the upside and downside implied variance as

$$f^U(S(T)) = \begin{cases} \left(\log\left(\frac{S(T)}{S(t)}\right)\right)^2, & \text{if } S(T) \geq S(t) \\ 0, & \text{otherwise} \end{cases} \tag{54}$$

$$f^D(S(T)) = \begin{cases} 0, & \text{if } S(T) > S(t) \\ \left(\log\left(\frac{S(T)}{S(t)}\right)\right)^2, & \text{otherwise} \end{cases} \tag{55}$$

The first order derivatives can be calculated as previously, but again they will not be present in the final replicating formula, because if we plug-in $S(t)$, then the logarithmic returns will be zero. The respective second order derivatives are

$$f^U_{SS}(K) = \begin{cases} \frac{2(1-\log\left(\frac{K}{S(t)}\right))}{K^2}, & \text{if } S(T) \geq S(t) \\ 0, & \text{otherwise} \end{cases} \tag{56}$$

$$f^D_{SS}(K) = \begin{cases} 0, & \text{if } S(T) > S(t) \\ \frac{2(1+\log\left(\frac{S(t)}{K}\right))}{K^2}, & \text{otherwise} \end{cases} \tag{57}$$

However, we can see that both $f^D(S(T))$ and $f^U(S(T))$ are not twice-continuously differentiable at $S(t)$, but this discontinuity is not creating infinite integrals as shown by Da Fonseca and Xu (2017), so we will use this definition. This concludes the calculation of the replicating strategy.

$$E^{\mathbb{Q}}(f^D(S(T))) = \int_0^{S(t)} \frac{2(1+\log\left(\frac{S(t)}{K}\right))}{K^2} dK \tag{58}$$

$$E^{\mathbb{Q}}(f^U(S(T))) = \int_{S(t)}^{\infty} \frac{2(1-\log\left(\frac{K}{S(t)}\right))}{K^2} dK \tag{59}$$

We can see from these results that $V(S(T)) = E^{\mathbb{Q}}(f(S(T))) = E^{\mathbb{Q}}(f^D(S(T))) + E^{\mathbb{Q}}(f^U(S(T)))$.

Now, that we have defined an upper and a lower implied variance, we turn to the calculation of the upper and lower realized variance. The commonly accepted methodology to calculate realized semi-variances was developed by Barndorff-Nielsen et al (2008), who decomposed the realized variance as $RV = RV^+ + RV^-$, where the RV^- is the downside realized variance, which is defined as

$$RV^-_{t,T} = \sum_{i=1}^{n} r_i^2 \mathbb{1}_{r_i<0}, \tag{60}$$

while the RV^+ is the upside realized variance, which is defined as

$$RV^+_{t,T} = \sum_{i=1}^{n} r_i^2 \mathbb{1}_{r_i>0} \tag{61}$$

Barndorff-Nielsen et al (2008) shows that in this way if the process defining the dynamics of the underlying have both diffusion and jump components, then for example the downside realized variance converges to half of the quadratic variation plus the sum of negative jumps, which is exactly what we want here.

The implied and realized semivariances define the two risk premia, which we will use in the empirical part of the thesis. The upper variance risk premium is the difference between the upper realized and implied variances $UVRP_{t,T} = RV^+_{t,T} - IV^+_{t,T}$, while the downside variance risk premium is the difference between downside realized and implied variances $DVRP_{t,T} = RV^-_{t,T} - IV^-_{t,T}$.

In this section, I have introduced two semivariance measures, which are commonly used in forecasting underlying returns with option implied data. In the empirical part of the thesis, I will explore the connection of these measures to other factors in forecasting index and equity returns.

4.2 Volatility of volatility

Park (2013) argues that most academic papers have assigned the tail risk to jump processes. However, there can be two sources with crucial influence on the left tail of the return distribution: the jump variation and the volatility of stochastic volatility.

In a stochastic volatility setting, Park (2013) also shows that the skewness or kurtosis of the return distribution are proportional to the volatility of stochastic volatility. The authors try to capture this phenomenon by linking the volatility of volatility to the tail risk using their own so-called VVIX index. Park (2013) assumes that the VIX index is a measure of the implied volatility of the S&P500, which can be problematic as we have seen before. We ignore this problem here in order to be able to define a measure that is an indicator of the volatility of volatility. In the market it is possible to trade with options on the VIX index and it is possible to download the volatility surface of VIX option data. From this point it is possible to carry out the calculation of implied variance as we have seen before. In this way, we get a measure, which is an implied volatility of an implied volatility measure. The VVIX index is the second moment of the VIX return distributions. The calculation of the second moment of the risk neutral distribution is carried out by the method of Bakshi et al (2003). Only difference is that here we want to calculate a volatility not just a second moment. Below, I will calculate the implied second moment, like before, and also the risk neutral expected value to get the real variance, which would then determine the volatility, we are looking for.

Theorem. *We will use $Var_{t+1} = E^{\mathbb{Q}}(r^2) - (E^{\mathbb{Q}}(r))^2$ formula to calculate the variance. The second moment is calculated as before*

$$E^{\mathbb{Q}}(r^2) = \exp(r(T-t))\Big(\int_{S(t)}^{\infty} \frac{2(1-\log\Big(\frac{K}{S(t)}\Big))}{K^2} C(t,T,K)dK + \int_{0}^{S(t)} \frac{2(1+\log\Big(\frac{S(t)}{K}\Big))}{K^2} P(t,T,K)dK\Big), \tag{62}$$

where now the $S(t)$ *is the time* t *value of the VIX index. The expected value is*

$$\mu_{t,T} = \exp(r(T-t)) - 1 - \frac{\exp(r(T-t))}{2} IV(T-t) - \frac{\exp(r(T-t))}{6} IS(T-t) - \frac{\exp(r(T-t))}{24} IK(T-t) \quad (63)$$

So, the Volatility of volatility is calculated as

$$VV_{t,T} = \frac{E^{\mathbb{Q}}(r^2) - \mu_{t,T}}{T-t} \quad (64)$$

The proof of this result follows from the implied moments calculations and also can be found in Bakshi et al (2003). Next to the theoretical arguments Park (2013) also provides empirical evidence that the VVIX index serves as an indicator for the tail risk. First, he shows the VVIX is consistent with stock market decline, an increase in volatility and a dry-up of market liquidity, which are usual features of market crashes. He also shows that spikes in the VVIX index coincides with crisis scenarios. The empirical investigation shows that the VVIX index has strong predictive power for equity returns. Park also distributes the effect of the VVIX index into integrated volatility of volatility and volatility jumps. He shows that most of the predictive power of the VVIX index can be attributed to the integrated volatility of volatility.

A good feature of this measure is that since it a second moment of a return distribution, it is not affected highly by the largely missing deep OTM options. This phenomenon is strongly affecting other, mostly jump-induced tail risk measures.

Unfortunately, this measure can only be calculated for the S&P500 since the VIX is a measure of the implied volatility of this index. My goal in this thesis is to select my portfolio based on individual measures, so I won't be able to use this in my empirical analysis, but I wanted to introduce this measure in this work as a possibility.

4.3 Tail Risk Premium

In this section I will introduce a tail risk measure, called the tail risk premium or TRP in short. One of the most important stylized facts of financial markets is that return distributions has heavier tails, than the normal distribution. It means that a measure, which can incorporate information about the tails of the risk neutral distribution might have explanatory or predictive power regarding future returns of the underlying. In this section, I will return to the calculations performed in the Alternative formulations section from the Implied moments chapter.

There are several different measures in the academic literature, which try to incorporate this phenomenon, defined by for example Bollerslev et al (2015). I try to be consistent with the previous calculations, therefore I use a tail risk measure derived from the Alternative formulations section. I have shown there that the definition for the implied variance used by Kozhan et al (2013) contains information about the higher moments of the risk neutral distribution. Following those results Du and Kapadia (2012) and Chow et al (2016) define a risk measure, which is based on the difference between the implied tail ($IT_{t,T}$) and the realized tail ($RT_{t,T}$) of the return distribution. The implied tail was first defined by Du and Kapadia (2012) as the difference between the squared VIX and the replicating portfolio of the variance contract. They argue that this measure captures the tail of a distribution, since it only contains information about the third and higher moments as we can see from the calculation below.

$$IT_t = VIX_t^2 - IV_t = \int_{S(t)}^{\infty} \frac{2}{K^2} C(t,T,K) dK + \int_0^{S(t)} \frac{2}{K^2} P(t,T,K) dK$$

$$- \int_{S(t)}^{\infty} \frac{2(1 - \log\left(\frac{K}{S(t)}\right))}{K^2} C(t,T,K) dK + \int_0^{S(t)} \frac{2(1 + \log\left(\frac{S(t)}{K}\right))}{K^2} P(t,T,K) dK$$

$$= \int_{S(t)}^{\infty} \frac{2\log\left(\frac{K}{S(t)}\right)}{K^2} C(t,T,K) dK + \int_0^{S(t)} \frac{-2\log\left(\frac{S(t)}{K}\right)}{K^2} P(t,T,K) dK$$

$$= \sum_{n=3}^{\infty} \frac{2}{n!} E^{\mathbb{Q}}(r_{t+1}^n) \tag{65}$$

We can see that this measure contains information about the third and the fourth implied moments, which are already modelled by the implied skewness (skewness premium) and implied kurtosis (kurtosis premium). This measure can be used to decide, whether higher order moments are in statistically significant relationship with the underlying returns, or not. It is well-documented that the risk neutral skewness possesses an important role in the financial market and many articles have shown that it has a strong relationship with stock returns. (Du and Kapadia (2012) and Chow et al (2017))

The realized tail defined by Chow et (2016) as the difference between the so-called polynomial variation and the realized variance. The polynomial variation is also defined following the results from the Alternative formulations as two times the difference between the effective returns and the logarithmic returns, so $PV_{t,T} = 2(R_{t,T} - r_{t,T})$ and $RT_{t,T} = PV_{t,T} - RV_{t,T}$. The measure I will use is the difference between the realized and the implied tails, called the tail risk premium, $TRP_{t,T} = RT_{t,T} - IT_{t,T}$.

In this section, I have defined a measure that tries to model the tail risk of underlying return distribution. There are other very popular tail risk measures, one of the most widely used ones are the jump-based measures like the signed jump measure defined by Patton and Sheppard (2015). I wanted to use these types of measures as well, but these can only be calculated using high-frequency data, which was unfortunately unavailable for me when I was writing this thesis.

4.4 Term structure of Implied Moment Measures

Lou and Zhang (2016) showed that their model-free forward variance measure has strong predictive power for equity returns in short term in 1- , 3-, and 6-month horizon. The authors use the same methodology, which is used by CBOE to construct the VIX index with a particular maturity. As an example from Lou and Zhang (2016): on October 31, 2008, the VIX with 126 business days is defined as

$$VIX^2_{t,126} = \left[T_1\sigma_1^2\frac{N_{126}-N_{T_1}}{N_{T_2}-N_{T_1}} + T_2\sigma_2^2\frac{N_{T_2}-N_{126}}{N_{T_2}-N_{T_1}}\right]\frac{N_{252}}{N_{126}} \tag{66}$$

In this formula T_1 and T_2, which are the number of business days to March 21, 2009 and June 20, 2009, which are expiry dates of traded SPX options, the N_{T_i} denotes the number of business days to a given date. Following this logic the $[T_1, T_2]$ forward VIX^2 term at time t is defined as

$$VIX^2_{t,(T_1,T_2)} = VIX^2_{t,T_2-t}\frac{N_{T_2}}{N_{T_2}-N_{T_1}} - VIX^2_{t,T_1}\frac{N_{T_1}}{N_{T_2}-N_{T_1}} \tag{67}$$

The forward variance in the period τ_1-ahead at time t in the period $[t+\tau_1, t+\tau_1+\tau_2]$ is defined by Lou and Zhang (2016) as the appropriate forward VIX term $VIX^2_{t,(t+\tau_1,t+\tau_1+\tau_2)}$. Lou and Zhang (2016) use three different forward variance the 3-, 6-, 9-month forward variances as $FV_{t,(t,t+3m)}$, $FV_{t,(t+3m,t+6m)}$ and $FV_{t,(t+6m,t+9m)}$, and use these measures in their forecasting framework.

Borochin et al (2018) defines a skew term spread from, which is the difference between the long-term skewness and the short-term skewness. The intuition behind this spread is that they observe positive predictability with short-term skewness, which they regard to informed-trading demand, but negative predictability with long-term skewness caused by the so-called skewness preference (for a detailed discussion of the informed-trading demand and the skewness preference and their relationship with skewness the reader is advised to read the original paper, Borochin et al (2018)). The skew term spread is simply defined as the difference between 12-month risk neutral skewness and the 1-month risk neutral skewness. The empirical results of Borochin et al (2018) show that the skew term spread has strong predictive power, and portfolios with low skew term spread significantly outperform portfolios with high spread in their sample.

The empirical results of these measures are encouraging. Both measures try to get incorporate information about the term structure of the variance or the skewness. This would imply that the term structure of the variance or the skewness has predictive power for the underlying return. However, the definitions of them are rather ad hoc, because they only use chosen points of the term structure of the implied measures and disregards other parts. In this section I will describe a methodology to create variables, which incorporate the whole term structure of the variance and the skewness.

4.4.1 Diebold - Li Framework

In this section I will shortly introduce a statistical method, which helps me to construct variables that represent the term structure of risk neutral variance and risk neutral skewness in their entirety. The definition of the variables is based on the econometric framework of Diebold and Li (2006). The problem with forecasting yield curves (just like with volatility surfaces) is that they are multi-dimensional time series, which makes the forecasting troublesome in most cases. Diebold and Li introduced a methodology with which these time series can be approximated with only 3 components with high precision. The idea of fitting the Diebold-Li framework was introduced by Simity (2018) in his thesis.

Diebold and Li aimed to create an econometric model that can be effective in forecasting the evolution of yield curves. The authors work around the problem of dimensionality with using the exponential component framework Nelson and Siegel (1987). Diebold and Li model the term structure of yield curve as

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\Big(\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau}\Big) + \beta_{3,t}\Big(\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau} - \exp(-\lambda_t\tau)\Big) \tag{68}$$

The $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ are the dynamic factors. The authors argue that because of the definitions of the loadings, the three components can be interpreted as the level, the slope and the curvature of the yield curve respectively. I will illustrate this interpretation in the next section. The τ denotes the maturity and the λ_t parameter governs the rate of the exponential decay, so a very low λ_t would imply a slow decay, therefore would better capture the structure of the curve at long maturities. This parameter can either be fixed before fitting the model or optimized to achieve the best fit.

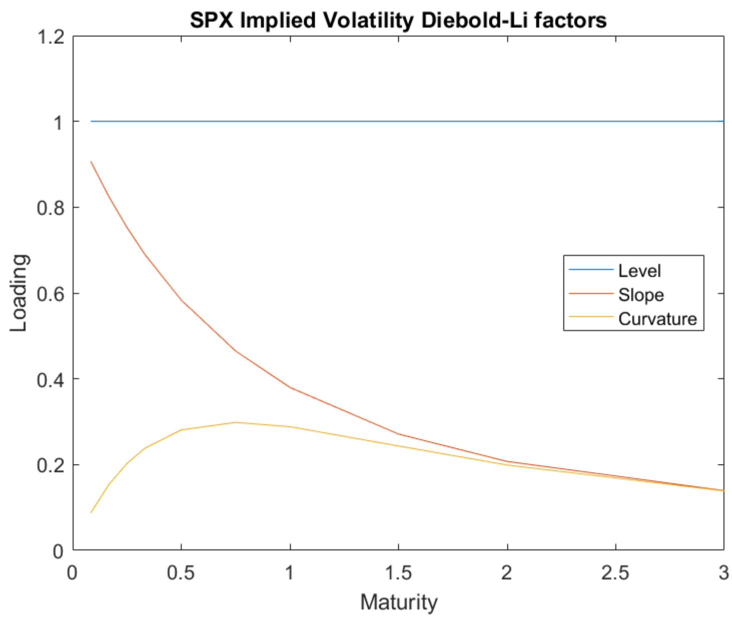
An alternative methodology for decreasing the dimensionality of the data would be the Principal Component Analysis (PCA). The main difference between PCA and the Nelson-Siegel framework is that here we impose a structure on the factor loadings. In short, this means that while PCA calculates different components and loadings for each time series, the Nelson-Siegel framework uses pre-determined component loadings. The advantage of the Nelson-Siegel approach is that the factors have clear interpretation (level, slope and curvature), which can be problematic in the case of PCA, so the use of the Nelson-Siegel approach makes the economic interpretation lot easier.

4.4.2 Fitting Diebold - Li Framework to Implied Moments

In this section, I will describe the results of fitting the describe framework to individual implied variance and implied skewness series. The decay parameter of the the Diebold-Li model can be optimized together with the regression coefficients by nonlinear least squares algorithm to achieve the highest R^2 value, which would mean the best possible fitted model or also can be set before fitting a model. In order to avoid higher computational time, I used fixed λ_t parameter. A way to fix this parameter is described in the original article by Diebold and Li (2006). The decay value determines the maturity at which the curvature loading achieves its maximum. They picked the value that maximizes the curvature at approximately one-third of the maximum available maturity. In my dataset I could compute the model-free implied volatility and implied skewness for at most 3 years of maturity, so the picked lambda value should maximize the curvature around 1 year.

The Diebold-Li model was fitted to the model-free implied volatility and implied skewness for every equity or equity indices in the dataset. To illustrate the model, I have plotted loadings for the implied volatility of the S&P500 index against the available maturities.

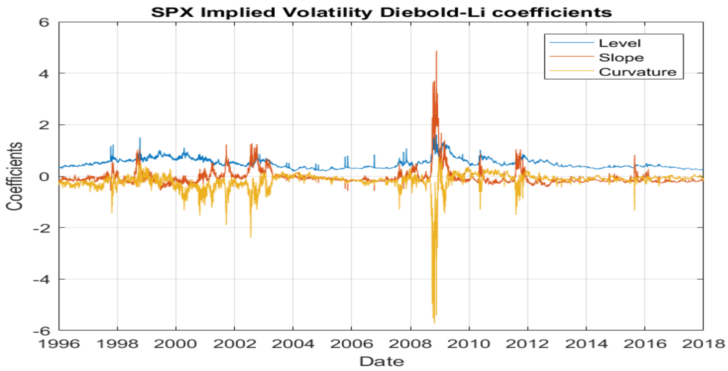
Figure 4: Diebold-Li loadings



In general the model created a very good fit for both implied moments with R^2 values ranging from 0.9 to 0.98, therefore we can conclude that these three factors are able to model the term structure of the implied volatility and the implied skewness. I will use these factors instead of the quite ad-hoc factors of Lou and Zhang (2016) for the implied volatility and Borochin et al (2018) for the implied skewness.

There are some general time series properties for the Diebold-Li coefficients for these implied moments. These properties are illustrated in the two plots below and two tables in the Appendix.

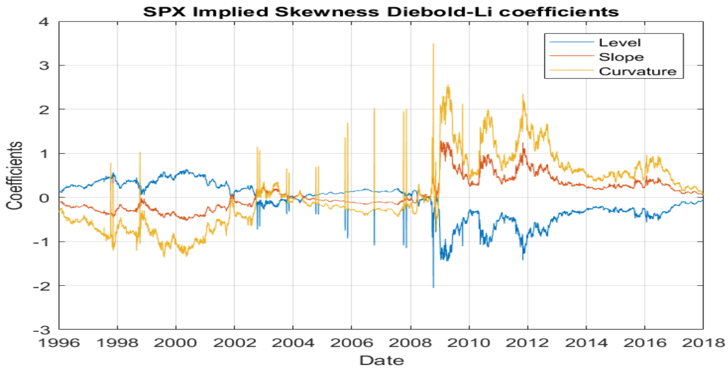
Figure 5: SPX VRP factor series



The plot shows that on average the mean of the slope and the curvature are around zero, but the variances of the series are time-varying. These features are valid in general in my dataset and are illustrated in Table 14 in the Appendix. The table also shows that for every equity or equity indices the coefficients are significantly different from zero. It is also important to note that the coefficients are in general highly autocorrelated, therefore we will have to be careful when we are using these factors in the empirical analysis, so in Table 14 the t-values are calculated by the method of Newey-West (1986), where the authors presented a method, which enables us to calculcate a heteroscedasticity- and autocorrelation-consistent covariance matrix.

The next plot shows the evolution of the Diebold-Li coefficients for the implied skewness of the S&P500 index

Figure 6: SPX SRP factor series



The plot shows that the coefficients of the implied skewness are significantly less stable in time, than the coefficients of the implied volatility. Also, in general the coefficients are less autocorrelated. Table 15 shows the same properties for the implied skewness as Table 14 for the implied volatility. In general the t values are significantly smaller here, which indicates that in general the coefficients here are less likely to be significant.

In this section, I have presented the two articles which tried to incorporate the term structure of implied moments in their analysis. I have argued that their methods are not able the grasp the entirety of the term structure. Therefore, I have presented a methodology created by Diebold and Li, which is often used to model the term structure of yield curves and fitted this method to the implied volatility and implied skewness. In general the methodology achieved a very good fit, so we can conclude that these factors model the term structure well, therefore I will use them as proxies for the term structure in the empirical modelling section.

5 Empirical literature

In this chapter, I will summarize the most important empirical results regarding the measures already introduced in the thesis. These articles are mainly focused on the empirical connection between measures derived from the call and put options or the implied volatility surface and the time series or cross-section of underlying returns.

After Carr and Wu (2008) first investigated the existence of the variance risk premium and characterized the empirical properties of the premium, many empirical researchers have turned their attention to the possibility of forecasting underlying returns with the variance risk premium. The intuition behind this idea is twofold. For once, as I have already mentioned, the variance risk premium is in close relation with the risk appetite of the investors. The risk appetite has great effect on the demand of the investors, which affects the evolution of the underlying price. The other channel which connects the variance risk premium to the returns, is the connection between the option market and the underlying market. The option-implied variance is the market expectation of the variance for the lifetime of an option. This forward-looking nature of this measure could possibly create an informational channel from the option market to the underlying market, which might make it a good predictor for underlying returns.

Bollerslev et al (2009) was the first to investigate this hypothesis closely. They analysed the returns of the S&P500 index and used simple linear regression to predict the next period returns. They have carried out this procedure for multiple periods and come to the conclusion that the predictive power of the variance risk premium is strong for short horizon, especially for 4- to 6-month horizon. They argue that the reason behind this result is the fact that the variance risk premium is simply a proxy for market uncertainty. Later, Bollerslev et al (2014) showed that this phenomenon cannot be explained by the bias caused by the finite sample, and they also showed that the result is not specific to the S&P500 index, rather it is satisfied for many international equity markets. Another branch of this literature is concerned about the cross-sectional variation of underlying returns. The methodologies applied by this branch are based on from one side the philosophy of the classical Fama-French factor model (Fama and French, 1993) and Fama-Machbeth regressions (Fama and Macbeth, 1973), on the other side the methodology is more focused on portfolio selection mainly by the sorting or the double-sorting methods from Bali et al (2016). The articles in this area of research are not focused on prediction of underlying returns, but rather on finding the cause of the difference in their sample between the actual returns in a cross-section with the help of different factors. Bali and Hovakimian (2009) examined the S&P500 index and all of its constituents in their article, where they both apply a factor model regression and a sorting methodology to examine whether the variance risk premium and the difference between call and put implied variance can explain the return differences. They show that both variables have significant predictive power, and they also show that they contain different information about the underlying returns.

Following these results many researchers have examined whether this strong relationship is specific to the equity markets or not. Chevallier and Sévi (2014) has examined the predictability of the WTI oil futures market and shown that the relationship between the variance risk premium and underlying return is particularly strong in this market even after controlling for market specific factors. Chevallier (2013) have come to same conclusion about CO_2 markets. Bams et al (2017) analysed this relationship on the equity, crude oil and gold markets and found the existence for all of them. However, they showed that only the equity uncertainty is priced market-wide, the crude oil uncertainty is only priced on oil-markets, and the gold uncertainty is only asset specific. Prokopczuk and Simen (2014) analysed a profitability of a commodity market portfolio, where the short positions are selected based on the variance risk premium. They show that the short variance swap positions have Sharpe ratio of around 40%.

Some articles have started to examine the effects of decomposing the variance risk premium on the precision of the predictive models. Feunou et al (2017) decomposed the variance risk premium to the downside variance risk premium and the upside variance risk premium. They have come to the conclusion that the downside variance risk premium is the main component of the variance risk premium, and there is a significant relationship between the downside measure and equity premium. These results were also confirmed by Da-Fonseca and Xu (2017).

Other articles have tried to examine the relationship between underlying returns and higher moment risk premia. Kozhan et al (2013) tried to characterize the skew risk premium. They show some stylized facts about the skew risk premium and conclude that the skew risk premium and variance risk premium is closely interrelated. Harris and Qiao (2018) analysed the predictive power of all three previously described moment risk premia together. They conclude that the variance risk premium and the skew risk premium are significantly related to the equity returns, but the kurtosis risk premium is not, but if they control for firm specific variables, then only the skew risk premium remains significantly related to equity returns. However, Park et al (2018) showed that the implied volatility convexity, which is a proxy for the implied kurtosis level, is a good portfolio selection criterion. They have shown that the portfolio created by the equities, which are in the highest quantile based on the implied convexity, significantly outperform portfolios with lower implied convexity.

The academic literature has also dealt with the concept of tail risk. The tail risk is associated with tail of the risk neutral distribution. They commonly argue that the predictability of underlying return is caused by the risk neutral higher moments, which researchers try to capture with tail risk measures. There are two commonly accepted causes of the irregularity of higher moments, the volatility of volatility and the jump component. We have to note that these two components are not exclusive, for example Bates (1996) developed a stochastic volatility model, which is extended by a compound Poisson process to create a jump component. Many empirical articles have tried to attribute the tail risk to a jump component or a volatility of volatility and examined the relationship between the tail risk and underlying premium. Park (2013) argues that it is the volatility of volatility, which is driving the heavy-tail of the risk neutral distribution and find that his VVIX index is a good predictor of S&P500 returns. However, Patton and Sheppard (2015) argues that the main driver of market volatility is negative return jumps. The cause of the tail risk is not decided, but the empirical results are closely related to higher moment risk premia result, which suggest that it is skew risk premium rather than variance risk premium that is driving the underlying returns.

Finally, Lou and Zhang (2016) defined the forward implied volatility, which is derived from two points from the term structure of the implied volatility. They show that this forward variance measure is a good predictor for stock market return. Borochin et al (2016) also analysed the term structure of the implied skewness. Their skew spread is simply the difference between two points from the term structure. They show that portfolios created from the lowest decile of the skew spread significantly outperform other portfolios with higher skew spread. These findings show that the term structure of implied measure might have significant information for underlying returns.

In this chapter, I have introduced some of the most important academic literature concerning the relationship between information derived from option prices available on the market and the return of the underlying. The most important problem with the literature is that as far as I know most of the research papers have not carried out a comprehensive analysis of all the defined measure to understand, which has independent predictive or explanatory power from the other measures. The table below tries to summarize the main articles mentioned in this chapter.

Table 1: Main Empirical Articles

Article	**Year**	**Underlying**	**Measure**
Bali and Hovakimian	2009	Equity	VRP, Kurtosis
Bams et al	2017	Equity oil and gold	VRP
Bollerslev et al	2014	Index	VRP
Bollerslev et al	2009	International Indices	VRP
Borochin et al	2018	Equity	IS Term Structure
Chevallier	2013	CO_2	VRP
Chevallier and Sevi	2014	WTI	VRP
Chow et al	2014	Equity	VRP, TRP
Feunou et al	2017	Equity	VRP, UVRP, DVRP
Harris and Qiao	2018	Equity	VRP, SRP, KRP
Kozhan et al	2013	Equity	VRP, SRP
Luo and Zhang	2017	Indices	IV Term Structure
Park	2013	Equity	Vol of Vol
Park et al	2018	Equity	Kurtosis
Prokopczuk and Simen	2014	Commodity	VRP

6 Data and Methodology

In this thesis, I have used data and carried out every calculation for the period between January 1, 2000 and December 29, 2017. I have used smoothed Black-Scholes implied volatility surfaces data from the OptionMetrics database. The smoothing technique applied by OptionMetrics takes the raw Black-Scholes implied volatilities calculated from actual options traded on the market and apply a Gaussian-type kernel smoothing algorithm to calculate Black-Scholes implied volatility for fixed time to maturity and delta grids for every day. Most of the options traded on the market are American-type options. In order to calculate the raw implied volatility surfaces OptionMetrics apply a tree-based algorithm, then applies the same kernel smoothing technique as for European-type options. There is a detailed explanation on the exact methodologies in the handbook of OptionMetrics (OptionMetrics, 2018). From the smoothed implied volatility grid, call and put options are re-calculated using the Black-Scholes formula from the first chapter, which are then used for the final calculation of implied measures. In order to calculate the option prices I needed risk free interest rate curves and daily stock prices. The curves were downloaded from OptionMetrics (2018), while index and stock prices were obtained from Bloomberg (2018). I also have to note that in most literature where the aim is to select equities, the variables include in general traditional factors (like size, liquidity or book-to-value, etc.) for individual equities. Unfortunately, I did not have data for these, but it would be a good improvement to include these factors next to my option-implied measures. The appendix contains the tickers and the full name of the equities and indices used in this work. The number of samples were selected to keep the size of the data manageable. Also the tickers were selected from different industries, and I also wanted them to be liquid enough and be part of the S&P500 since 2000.

As I have mentioned in the introduction of realized measures, in the academia it is the standard to calculate the daily realized measure by summing 5-minute returns on the appropriate power. Unfortunately, it is very hard to get intraday data with university privileges, therefore I had to resort to daily prices. Following the methodology of Kozhan et al (2013) and Harris and Qiao (2018) I took the daily logarithmic return as a proxy variable and raised it to the appropriate power. This can cause some differences for my results, but I think this was the best possible option to use. I have considered other methodology choices, like summing for daily return and then rolling over a day, but in this way I would use a given return multiple times, which would introduce spurious autocorrelation into my realized measure, therefore I decided to use the simple daily returns. It can and will be a question of future research to re-calculate every statistical model and investment decisions, when I might be able to use intraday stock data for equities.

The empirical goal of this thesis is twofold. First, I am trying to model the connection between the defined measures and the daily logarithmic returns. For this objective, my method is based on Bollerslev et al (2009), who tries to predict the underlying return for multiple horizons with linear regressions where the independent variable is the variance risk premium. The main methodology is the dealing with the question of portfolio selection based on option implied measures. The methodology used in this work is based on Gu et al (2018). In the original article the authors apply various statistical and machine learning techniques by which they try to predict the evolution of the cross sectional variation of stock returns based on their explanatory independent variables. They use more traditional factor model variables such as market premium, size, firm-level information (for example: earnings-to-price), industry variables, liquidity proxy variables, momentum-trading specifice variables or price trends.

From this point, they use linear regression, regularized linear regression (lasso and ridge regression), principal component regression, generalized linear models, regression trees, and regression random forests and neural networks to predict the return for their test period. I have to note that their strategy is not dynamic, they make the learning, the validation, and the test processes on fixed time intervals. Their first conclusion is that applying non-linear predictive models, especially shallow neural networks, create significantly higher out-of-sample R^2 compared to linear models. After that, they form portfolios of the sample based on the predicted values from the different predictors. They sort stocks in their sample into deciles and then form a zero cost portfolio from the highest and the lowest deciles. They again conclude that non-linear models perform better, than linear models in the portfolio selection based on comparing Sharpe-ratios. They also noted that in both cases momentum-trading information and liquidity variables are the most important covariates.

In this work, I will use the same methodology to select my porfolios, except for the fact that I will try to use a dynamic strategy. I will define the models which will be used to predict the next period stock returns in the portfolio selection section, then I will present the results for the portfolio selection processes. The next chapter is dealing with presenting the empirical results for both empirical goals.

7 Empirical Results

In this chapter, I will present the results from the empirical framework introduced in the previous chapter. First, I will show the result on individual stock and index level, then I will present the results of the quantile-portfolios based on profitability and Sharpe-ratio level.

7.1 Analysing variables

In this section, I try to analyse the connection between the defined option-implied variables and the underlying returns. In the main part, I will show the results for the S&P500, but the appendix contains results for the individual equities as well. The main goal of this section is to analyse whether the previously defined variables have statistical explanatory power for the dynamics of equity returns or the result of the portfolio selection is just a mere luck. One of the most important aim here is to analyse whether the Diebold-Li coefficients contain any independent explanatory power for next day returns. It is also an important aim of this section to check whether our built intuitions about the measures are correct, and check whether my results are in line with previous research papers. The main text contain the results for the S&P500,

but I will state the general conclusions as well.

The analysis will be carried out by fitting linear regressions on a selected group of the defined factors and the underlying return series. The methodology here is the same as the one used by Bollerslev et al (2009), who after calculating the variance risk premium for the S&P500 ran the following linear regression.

$$\frac{1}{h}\sum_{i=1}^{h} r_{t+i} = \beta_1(h) + \beta_0(h) VRP_{t,t+h} + u_{t,t+h}, \tag{69}$$

where h denotes the forecasting horizon. I will carry out the same analysis for the variance risk premium and other option implied factors for one day horizon. I will start with the risk premia calculated directly from option-implied moments. Table 2 contains the results of seven regression models. The table shows the regression coefficients for the models and below them the respective Newey-West t-statistics in grey. The last line contains the adjusted R^2 values for every fitted model. We can see that the adjusted R^2 values are very small, but this is common in this research area for example Bollerslev et al (2009).

Table 2: Regression results for the S&P500 index

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
VRP	0.5842					0.60139	0.6628	
	2.5143					2.5579	2.2283	
SRP		-12.5155				-2.30318	-2.45545	
		-2.8343				-0.8506	-0.981	
KRP			-33.5891				-32.53025	
			-0.4998				-0.2398	
DVRP				1.611				
				1.8979				
UVRP				-0.6142				
				-0.7139				
DSRP					-41.41506			
					-4.1664			
USRP					7.90448			
					1.8471			
TRP								-4.2018
								-2.9472
R-squared	0.037	0.0083	0.007273	0.021	0.00173	0.083	0.074	0.0064

The first three columns contain results for univariate regression for the moment risk premia. The first column shows that the variance risk premia is positively related to the underlying returns and this result is strongly significant. This means that when the difference between the realized moments and option-implied moments decreases, then the price tends to increase. This result confirms many previous academic results for example Bollerslev et al (2009) and Bams et al (2017) and reinforces the perception that the variance risk premium can be interpreted as a measure of market uncertainty. The parameter of the skewness and kurtosis are significantly higher in absolute terms, which is caused by the fact that the SRP and the KRP are in general very small in absolute terms. The sign of the skewness risk premium coefficient is different from the coefficient of the VRP, which is in line with general results from the academic literature. The SRP has significant explanatory power indicated by the relatively large negative t-value, which shows that the skew risk premium still contains some information about the uncertainty of the dynamics of underlying returns. However, the adjusted R^2 value here is significantly lower, than for the first model, so it provides less information about market returns, than the VRP regression. Although the kurtosis risk premia has a parameter, which is large in absolute terms, it's coefficient is not significantly different from zero, which is a similar conclusion as the one drawn by Harris and Qiao (2018).

If we analyse the results from the semivariance and semiskewness measures, then we can come to a similar conclusion as Feunou et al (2017) and Da Fonseca and Xu (2017). First, in Model 4 we decompose the variance risk premium into downside and upside variance risk premium. We can observe that the downside variance risk premium is dominating the effect of the VRP, while the coefficients of the upside variance risk premium is significantly lower in absolute terms and not even significantly different from zero. This confirms the existence of the asymmetry in risk-return trade-off between responses to positive and negative market scenarios introduced in section 4.1. A very similar conclusion can be drawn for the upside and the downside skew risk premia. The last column contains the result of the univariate regression, where the independent variable is the tail risk premium. The result is closest to the results for the skew risk premium, which indicates the established empirical property that the higher order moments and higher order risk premia are dominated by the third moment and third moment premia.

Now, we turn to the analysis of the Diebold-Li coefficients. Table 3 contains the regression results for the models containing either one specific factor or all Diebold-Li factors next to the variance risk premium.

Table 3: Regression results for the S&P500 index

		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	VRP	0.5790	0.5731						
		1.9974	2.3468						
	Level	0.1006		0.5470					
		0.4402		1.2374					
Variance	Slope	-0.1246			-0.1102				
		-0.8026			-1.1744				
	Curvature	0.0113				0.0740			
		0.0872				0.9118			
	Level		0.5486				0.2466		
			0.7367				1.0564		
Skewness	Slope		0.4342					-0.0718	
			0.4075					-0.7625	
	Curvature		0.1007						0.0203
			0.3360						0.2446
	R-squared	3.90%	3.30%	0.34%	0.21%	0.08%	0.26%	0.08%	0.01%

The first column contains the level, the slope and the curvature of the term structure of the implied variance. The Diebold-Li factors are highly correlated with each other, the empirical observation underlying this correlation is the fact that when markets become turbulent the general level of the implied volatility surface and the curvature of the surface increases at the same time. The columns show that none of the factor coefficients are statistically significantly different from zero. This relevant multicollinearity can make it harder to distinguish the effects of different components from each other. This make it intuitive to try out models where the factors are included individually. Unfortunately, the coefficients of the Diebold-Li factors are not significant in these cases as well. We can come to the same conclusion about the term structure of the implied skewness as well. I have also run these models where the variance risk premium and individual Diebold-Li factors are also included, but the results showed no relevant differences. These results imply that I have to reject my hypothesis that the term structure of the implied variance and the implied skewness has significant information about the dynamics of underlying returns. However, both Luo and Zhang (2017) for the implied variance and Borochin et al (2018) showed the significance of their term structure measure for the equity return for significantly longer time horizons like 1 month, 3 months or 6 months horizons. I will briefly touch on the relationship between the term structure of implied moments and the equity returns in the Robustness Check chapter of this thesis.

The conclusions drawn about the measures are true in general for the equities in my dataset. The appendix contains two tables about the regression coefficients of implied moments. The coefficients of DVRP and the UVRP are estimated together, while the coefficients of USRP and the DSRP are estimated together and with controlling for the VRP. We can see that for most of the equities (with 4 exceptions) the coeffient of the VRP is positive. The same similarity can be observed for the SRP. About the kurtosis, we cannot come to a clear conclusion about the sign of the moment risk premia, but these values are in most cases not significant. The sign and the magnitude of the other measures are quite similar to the presented table about the S&P500. For example we can see that unfortunately the estimates of Diebold-Li coefficients are not significant. In cases when there are differences from these conclusions we can observe structural differences (meaning that not just one parameter is different but most of them are), good example for this are the NKE, KO or even DUK in some cases. It is also important to note that the result for the OEX completely agrees with the result for the SPX, which is in line with previous academic research papers that were mainly dealing with indices and not with individual equities, which gives me confidence in my results.

To conclude the first part of this empirical analysis, we can see that for SPX and OEX are in line with the intuition about the defined measures and are in line with previous academic research. About the individual equities we cannot come to such a clear conclusion since there are some exceptions, but this has already been shown by previous research papers for example Carr and Wu (2008).

7.2 Portfolio selection

The last subsection would enable us to select those variables only which has shown significant results. However, as it will be shown later, some statistical models are able to "filter out" insignificant variables, and I want to test out this ability as well, so I leave every defined factor in the prediction models. I wanted to check the ability of the defined factors for portfolio selection with a dynamically adjusted strategy. This would make the result closer to what could happen in a real life trading scenario. However, this makes the forecasting framework very heavy computationally. The portfolio selection is carried out in a rolling-window manner. This means that the positions are selected by return predictions from statistical models, which are fitted to yearly data and then rolled-over one day after it forecasts the returns for the next day. This process is carried out throughout the whole period of the dataset for each individual equity. After this procedure, for every model I select the tickers belonging to the highest quantile of the forecasted returns every day. The equities for the lowest quantile are also selected in order to decide on the effectiveness of our portfolio selection. Then next day realized returns of the selected items are aggregated and analysed.

In this subchapter, first I will introduce the statistical models used in this rolling-window framework. Then I will present the results from selecting portfolios from the highest and lowest forecasted return quantiles. Lastly, I will try to check these results in a simplified real world market scenario as well.

7.2.1 Statistical models

In this section, I will shortly describe the statistical models used in the portfolio selection procedure. The introduction here will not be formally detailed, since the methodology of these statistical models is not in the main scope of this thesis. However, the basics and the estimators will be presented, and interested readers are advised to read for example James et al (2013) and Hastie et al (2009), on which this introduction is based, for detailed discussions regarding the models. In this analysis I will use linear regression, ridge regression, decision tree regression, random forest regression and support vector regression in order to forecast the next day returns.

The first model used is the simple linear regression. In this framework I fit the linear function on year of return data and option implied data, and then I predict the next day return. The process is then repeated for the yearly data again, but the data is rolled over by one day. The linear regression used here is defined as

$$r_{t+1} = \beta_{0,t} + \sum_{i=1}^{n} \beta_{i,t} x_{i,t} + \epsilon_t, \tag{70}$$

where n is the number of independent variables. The coefficients of the regression model is calculated using ordinary least squares method.

The second statistical method used in this thesis is the ridge regression model, which is a modified version of the previous linear regression. The ridge model belongs to the family shrinkage methods for regression models, and imposes a penalty on the size of the coefficients. This means that instead of minimalizing the sum of squared error as in the usual least squares method, the penalized sum of squared errors is minimalized as can be observed below.

$$\beta_t = \arg\min_{\beta} \Big(\sum_{i=1}^{N} (r_{i,t+1} - \beta_{0,t} - \sum_{j=1}^{n} \beta_{j,t} x_{j,t})^2 + \lambda \sum_{j=1}^{n} \beta_{j,t}^2 \Big) \tag{71}$$

The λ is called the penalization parameter, which is greater than 0, and controls the shrinkage of the model. The higher this parameter, the more the coefficients are driven towards zero. One of the most important advantages of the ridge regression is that it can, to a certain extent, select out insignificant independent variables. The penalization parameter is not estimated during the estimation of the model, it has to be selected in advance. In the machine learning literature, these values are referred to as hyperparmeters, which can be optimized before the prediction. I have carried out this hyperparameter optimization by selecting the last month of every window as validation set, and checked the forecasting error for the pre-defined hyperparameter set. In the end the best model is selected based on the mean squared error on the validation set and only this best model is used to predict the next day return. For the ridge regression the λ parameter is selected here from the set [0.01, 0.1, 0.2, 0.3, ... 0.9, 1,2]. This hyperparameter optimization significantly increases the computational time, but gives a clearer view on the investment performance of a given model.

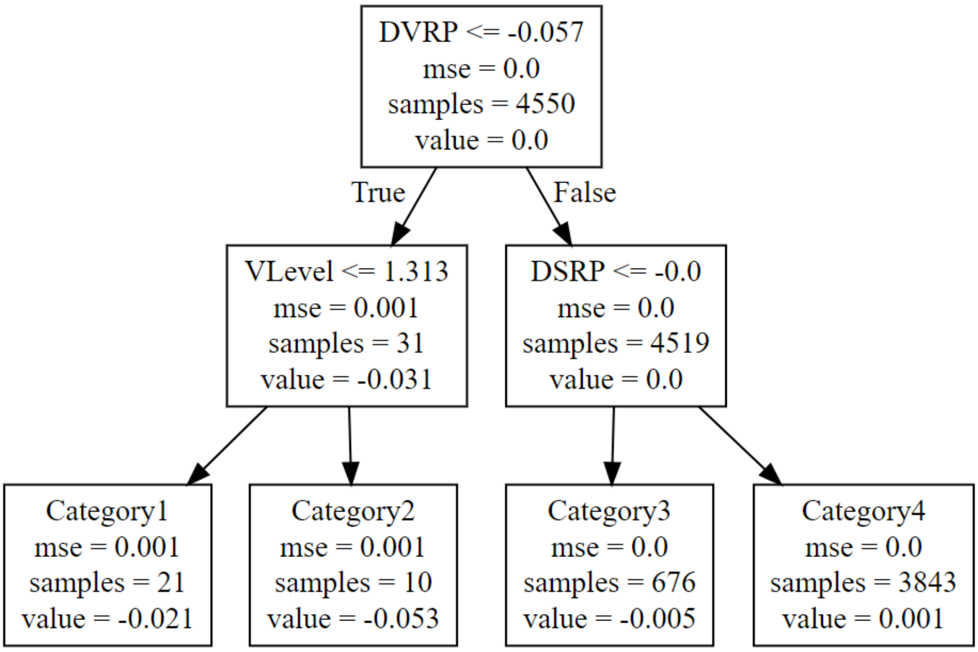
The next model used here is closer to the field of machine learning, than to traditional econometrics. The model is called decision tree regression, and the main goal of the model is to capture the non-linear connection between the independent variable and the dependent variable, while also incorporating the interactions between the independent variables. In this thesis, I have used CART (Classification and Regression Trees) regression trees during the prediction process. The fitting procedure of these regression trees can be divided into two steps. First, the space of independent variable is divided into K number of non-overlapping regions R_1, R_2 ... R_K. After the splitting, for every observation falling into the same R_i region we make the same prediction, which is basically the mean of the value of the dependent variables of those observations that fall into R_i. In theory, the regions could be created in any shape, but in this case the regions are selected as high-dimensional rectangles. The rectangles are selected in order to minimize the residual sum of squares (RSS)

$$\min \sum_{k=1}^{K} \sum_{i \in R_i} (r_{t+1} - \hat{r}_{t+1}) \tag{72}$$

However, it would be very demanding to consider every possible K steps, so instead a recursive binary splitting algorithm is applied, where we begin with the whole dataset and then at every step we split the space into two, therefore two new branches are created. This branching is then repeated for the two separate branches. After the splitting is finished the remaining groups are called leaves and are used for the prediction.

The amount of branching performed is called the depth of the tree. The maximum depth of a tree can be pre-defined in order to avoid overfitting, This depth can be treated as a hyperparameter just like the penalization term of the ridge regression, so it can be optimized by the train-validation splitting method. I have carried out the same optimization as for the ridge regression and the depth of the regression trees are chosen from [2, 3, 4, 5]. The figure below is just an illustration of how this method works in practise. The regressor there is fitted on the whole time series of the S&P500 returns and factors with maximum depth of 2.

Figure 7: Decision tree illustration



The next model is really close to the decision tree regressors and is called random forest regression. The random forest model is one of the so-called ensemble models. The ensemble models are in general composed of multiple learning algorithms at once in order to improve the predictive performance. In the case of the random forest the multiple learning algorithms are differently initialized decision trees on different subset of predictors. After the pre-defined number of trees are fitted and they performed their individual predictions, the predictions are averaged and this will be the prediction of the random forest model.

The number of trees used in the random forest model is a hyperparameter. In practise 100, 1000 or even more trees can be used for a given problem. However, in order to create a dynamically adjusted strategy with random forest with optimized hyperparameters takes considerable computational time and power. Therefore I had to fix the number of trees at 15 and only the depth of the tree is optimized at every step on the same set as the individual trees. Although this is a great simplification, the running time still takes approximately 10 hours.

The last model of this thesis is the support vector regressor (SVR), which is based on the support vector machines, which is one of the most popular classification models. The quantitative background of the support vector regressor is more sophisticated, than the previous models and it would be really hard to summarize in this limited space. Therefore I recommend the reader to check the building blocks of the support vector regression from Hastie et al (2009). Here, I only note that I have used epsilon-SVR (so I have no control over the number of support vectors in the model, but I have control over the margin bound) and I have used Radial Basis Function (RBF) as the transforming kernel function. I have to note that for every model, except for the linear regression, the hyperparameter space can be significantly bigger, so the scope of the optimization should be increased. This can increase the performance of these models in theory.

In this section, I have defined the models that will be evaluated in the next section. To finish this section it is worth noting that if we turn to the topic of machine learning here, it would be natural to think about using Neural Networks, Deep Neural Networks, Recurrent Neural Networks, Convolutional Neural Networks or Long Short-Term Memory Networks, which are the most popular forecasting methods nowadays. Interested readers can learn more about these Deep Learning techniques from for example Goodfellow et al (2017). The main problem with using these techniques in my thesis is that the sophisticated learning procedure of these models takes considerable time and computational power even for one dataset. Here I would have to carry out a fit over 4000 times, which was impossible for me with the available time and available computational power. This should be a topic of a later research, when I will be able to obtain the computational power of many powerful GPU systems.

7.2.2 Quantile results

In this section, I report and analyse the results for every introduced statisctical model. The main indicators for the comparison are the difference between highest quantile cumulated returns and the lowest quantile cumulated returns, and the Sharpe ratios of the two quantiles.

Table 4 below presents the portfolio selection results. The table contains annualized average returns and annualized Sharpe-ratios. The first section of the table corresponds to the highest quantile, where the first row reports the time-series average of the cumulated returns of the selected positions, while the third line reports the Newey-West t-values for the null hypothesis that the average return of the highest quantile portfolios is zero. The line between them reports the Sharpe-ratio for the selected portfolio. The Sharpe-ratio is calculated following Sharpe (1966) as

$$SR = \frac{r_p - r_{rf}}{\sigma_p}, \tag{73}$$

where r_p is the portfolio return r_{rf} is the risk free return, and σ_p is the standard deviation of the portfolios excess return. In order to be consistent, I have used the same riskfree rates here from OptionMetrics (2018) database as previously in the calculation of the implied moments. The section under this contains the same three lines for the lowest quantile portfolio.

Table 4: Portfolio Selection Results

		Linear	**Ridge**	**Dec. Tree**	**R. Forest**	**SVR**
Highest	**Average return**	0.208	0.197	0.170	0.214	0.032
	Sharpe ratio	0.122	0.105	0.095	0.123	0.013
	t-value	9.114	8.233	7.353	9.145	1.394
Lowest	**Average return**	-0.248	-0.170	-0.181	-0.260	0.008
	Sharpe ratio	-0.134	-0.101	-0.106	-0.147	-0.002
	t-value	-8.482	-6.419	-6.671	-9.350	0.300
	Difference	0.457	0.367	0.352	0.474	0.025
	Sharpe ratio	0.322	0.272	0.261	0.334	0.010
	t-value	21.354	17.671	17.362	20.241	1.182

Table 4 shows these variables for every previously defined statistical models. First, we can observe that the support vector regression with the radial basis kernel performs significantly worse than any other introduced model. This can imply that the radial basis kernel transformation is unnecessary and even deteriorates our ability to predict next day returns. We can see that amongst all models the Random Forest regressor performed the best in both the difference in average returns and in the Sharpe-ratios. This is not surprising, since many previous applications have shown that the random forest is performing quite strongly in short-term time-series predictions, sometimes even stronger then the state-of-the-art deep learning frameworks.

What is more surprising is that the linear regression is performing significantly better than the decision tree regression and the ridge regression. This is surprising since a linear model is able to beat a model that in theory should be able to incorporate non-linear information as well, also in theory ridge regression should be better in selecting important variables and handle the problem of multicorrelation between explanatory variables. I do not have a distinct answer to explain any of these results. In my opinion less than yearly data for training and 1 month data for validating is too few to harness the advantages that these two models have over the linear regression. It is also possible that the relative good performance comes from the fact that the hyperparameter optimization should be increased as I have noted in the previous chapter.

In this subchapter, I have shown the result of the portfolio selection based on the option-implied factors. The result are promising since for every model the higher quantile perform better in average returns and Sharpe-ratios as well (although the difference for the SVR is not significant). Among the well-performing models the random forest showed the best results, since the average return and the Sharpe ratio were the highest for the high quantile and the lowest for the low quantile there. The results of the previous section are very encouraging, because it implies that our option implied factors contain useful information for portfolio selection. However, it should be tested whether in a real world scenario it can be profitable. One potential boundary for this profitable use of a strategy based on option implied factors is the cost of creating such a portfolio.

7.2.3 Investing results

In the last section of the main part of my thesis, I try to carry out a naive testing of the above mentioned real world trading scenario with an investing strategy based on the option-implied measures. For this I have downloaded bid-ask prices from Bloomberg (2018) and have assumed a transaction cost of 0.5 percent. I investigate two trading strategies. The first strategy is called Long strategy, and the portfolio here consists of long positions in equities belonging to the highest quantile of predicted returns. The second strategy is called Long-Short portfolio, which completes the previous portfolio with short positions on the lowest quantile equities. I apply the transaction cost when there are differences in the elements in the respective quantiles. Naturally, these strategies are not realistic since the frequency of change in the portfolio is unrealistically high, which means that I suspect that most of the profit is eaten up by price differences, the bid-ask spread and the transaction cost. The Table 5 below summarizes the results for the Long strategy, for every model the first line contains the results without transaction costs and the results in the second line results are exposed to transaction costs.

Table 5: Long Strategy Results

	End Wealth	Annualized Return	Max	Min	Average
Linear	1146.807	0.86%	1234.187	590.078	966.090
	1135.339	0.79%	1213.470	578.276	941.496
Ridge	1169.976	0.98%	1557.586	916.144	1202.959
	1134.877	0.79%	1526.434	879.498	1174.892
Dec. Tree	1127.729	0.75%	1357.937	728.258	1073.823
	1105.368	0.63%	1344.358	705.897	1046.361
Rand For.	860.322	-0.94%	1240.389	654.421	964.412
	834.512	-1.13%	1203.564	641.333	931.606
SVR	1101.267	0.60%	1345.237	827.602	1009.270
	1090.254	0.54%	1331.784	811.050	1003.867

The results confirm my previous concerns that this naive strategy is really costly, since the yearly returns are very low. What is suprising that the random forest, which was performing the best previously, is actually has the worst result on average. This is caused by the huge variety of equities in the portfolio, which makes this strategy the most costly. The Table 6 summarizes the same results for the Long-Short portfolio. The low average and lowest values are caused by the huge systematic losses during the financial crisis.

Table 6: Long-Short Strategy Results

	End Wealth	Annualized Return	Max	Min	Average
Linear	1237.740	1.33%	1552.085	524.560	1075.259
	1225.362	1.27%	1505.522	514.069	1047.893
Ridge	1714.555	3.37%	1956.431	893.708	1317.933
	1663.118	3.18%	1878.174	875.382	1287.145
Dec. Tree	1415.773	2.17%	1679.459	562.424	1078.520
	1401.616	2.11%	1595.486	551.176	1050.795
Rand For.	925.759	-0.48%	1460.318	459.923	931.374
	888.729	-0.74%	1401.906	441.526	899.665
SVR	1174.827	1.01%	1345.237	827.602	1009.270
	1102.564	0.61%	1331.784	811.050	1003.867

The results here are a bit more encouraging, but the yearly returns are still pretty low. One thing worth mentioning is that the Ridge regression performs significantly better than the others in both strategies (especially for the Long-Short ones), which can be caused by the models ability to select important variables. As I mentioned before the aim of this section was not to define a highly profitable portfolio, but to give a naive check whether the models are profitable with the presence of bid-ask spread and transaction costs. The results show that the profit is very low (especially on 17 years horizon), but it indicates for me that it should be possible to define profitable strategies from the defined indicators. One possible idea would be to forecast returns for the next month, and hold the portfolio constant monthly. This could make the costs significantly lower.

8 Robustness Check

In order to give an even more solid base to my results and validate the findings further, I will carry out various robustness checks. The aim of robustness check is to examine how sensitive my results are to specific data definitions or model settings. Basically, I will recompute the analysis of the previous chapters with slightly different settings and check the sensitivity of the results to these changes.

8.1 Portfolio Selection

It is not trivial to select the upper and lower quantiles for the long-short portfolio. It is valid to ask, whether the good selection results seen in the main text is only valid for the higher and lower quantiles or can be observed in other selection strategies as well. Here I will carry out the same portfolio selection for the top and bottom third and top and bottom fifth of the return prediction. Table 7 below shows the results for the fifth portfolio.

Table 7: Fifth portfolio

		Linear	**Ridge**	**Dec. Tree**	**R. Forest**	**SVR**
Highest	**Average return**	0.161	0.144	0.117	0.164	0.009
	Sharpe ratio	0.044	0.028	0.008	0.044	-0.008
	t-value	9.422	7.812	6.677	9.204	0.511
Lowest	**Average return**	-0.011	-0.006	-0.008	-0.011	0.000
	Sharpe ratio	-0.213	-0.174	-0.194	-0.234	-0.076
	t-value	-9.2447	-5.9484	-7.5256	-10.2293	0.2325
	Difference	0.0195	0.0135	0.0141	0.0200	0.0002
	Sharpe ratio	0.2168	0.1290	0.1431	0.2258	-0.0941
	t-value	22.495	19.887	16.821	21.106	1.843

We can observe basically the same ordering of the statistical models (however the highest Sharpe values are significantly smaller). Also the results again show that the measures can be good for equity selection. Table 8 below shows the results for the third portfolio.

Table 8: Third portfolio

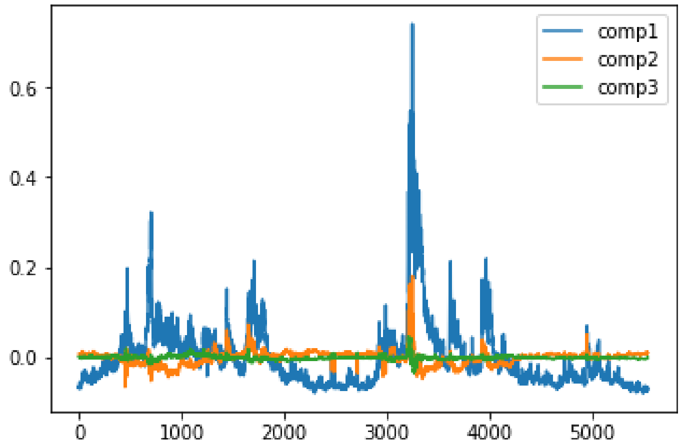
		Linear	Ridge	Dec. Tree	R. Forest	SVR
Highest	**Average return**	0.161	0.144	0.117	0.164	0.009
	Sharpe ratio	0.841	0.531	0.153	0.846	-0.161
	t-value	8.601	8.281	6.351	8.600	1.574
Lowest	**Average return**	-0.212	-0.114	-0.152	-0.218	0.005
	Sharpe ratio	-0.213	-0.174	-0.194	-0.234	-0.076
	t-value	-7.376	-5.993	-5.309	-8.097	0.145
	Difference	0.373	0.258	0.269	0.383	0.004
	Sharpe ratio	0.217	0.129	0.143	0.226	-0.094
	t-value	22.495	19.887	16.821	21.106	1.843

This section reinforces my results that the option implied information has some potential to help selecting equities for my portfolio.

8.2 Principal Component Analysis

As I have noted previously, it is possible to model the total term structure of the implied moments with principal component analysis instead of the Diebold-Li framework. This is what I will carry out in this section of the Robustness check. So I will analyse whether the components have explanatory power for next day returns. I have carried out the principal component analysis on the term structure of implied variance and skewness. I have consistently taken the first three components out of the PCA components to make it comparable to the Diebold-Li results. The plot blow shows that for the SPX the first component dominates the other two components.

Figure 8: PCA for the SPX IV



The Table 18 contains the same regression results as I have reported for the Diebold-Li factors. We can see the same result that unfortunately the term structure of the implied moments has no explanatory power for the next day equity returns.

Unfortunately, I did not have enough time to finish, what I have planned for this chapter, since there are possibilities to carry out other robustness checks as well. For example it would be valid to check other regression models for the variable analysis section to get other perspective on the connection between explanatory variables, or it would also be useful to check the results of both the explanatory regression and the portfolio selection with different forecasting horizons. The second check would be useful for the term structure in particular, since previous articles showed its significance to longer forecasting horizons. For the portfolio selection, it would be interesting to check the results for less frequent portfolio rebalancing. However, as I have mentioned I did not have the time to perform these, but in future when I will extend this research, I should do even more robustness checks.

9 Conclusion

In this chapter, I try to summarize the results of my thesis and also draw some conclusions about the research questions introduced in the first chapter. In short, my main question was whether information calculated from option prices available on the market has predictive power and the ability to help portfolio creation. In order to answer this question during my work I have defined measures calculated from market prices of call and put options available in the OptionMetrics database. These measures were calculated for two US equity indices and 24 individual equities, so I have to repeat that my conclusions and answers can only be regarded for these observations.

In this thesis, first I have provided a theoretical background of calculating option-implied information from option prices. In this part, the method of Breeden and Litzenberger (1978) to calculate the entirety of the implied density functions, and the method of Bakshi et al (2003) to calculate moments of the implied distribution were introduced. Since the first method has some limitations, I have decided to use the "model-free" method of Bakshi et al (2003) and have explicitly calculated the second, third and fourth implied moments from this method. Next, I have introduced the concept of moment risk premia for the three implied moments and tried to describe them empirically. Following this, I have defined other option implied variables which try to grasp information from the implied distribution which might be ignored by the three moment risk premia measures and have previously showed encouraging results in some academic research papers. For my thesis I have used a different methodology to test the term structure of the implied volatility and implied skewness. Previously this was proxied by the difference between random points from the term structure of the implied moments. For my thesis I have introduced the concept of Diebold-Li, which tries to compress the total information of the term structure of implied variance and implied skewness with 3-3 factors. This should, in theory, give a better view on the connection between the term structure of implied moments and the underlying return. After this, I have introduced previous academic research literature and data and the methodology of the empirical analyses of this work.

In the empirical results chapter, I have answered my research questions for the sample available for me during writing this thesis. First, I have performed a traditional analysis based on Bollerslev et al (2009) to check the statistical predictive power for the next day return. We have seen that the results are in line with our intuition about the moment risk premia and are in line with previous research. About the other measures we have seen that the tail risk premium is quite in line with the skewness risk premium, the effect of VRP is dominated by the DVRP. About the term structure, we have seen that for once the predictive power is coming from the Level and the Slope factors. However, they do not seem to be statistically significant for next day returns. For the main research question we have seen that there is potential in using the defined measure for portfolio selection, since the selected higher quantile portfolio had strong positive average return and has significantly outperformed the lower quantile portfolio. However, we have seen that a naive strategy based on these observations lead to a profitable investment, but the return is very low, since it is eaten up by price changes, bid-ask spread and transaction costs. There I have provided some ideas which could make real world investing results better.

I think that the results of this work are encouraging for future research. During the thesis I have noted some potential future improvements. One of the most important steps to increase the academic value of this work would be to carry out the same analysis for every individual equity in the S&P500 index, and also to other indices. This would obviously make the results more meaningful and would give a clear view on what we can use these measure for on the market. Also it would be important to get high-frequency data to better model the daily realized volatility. The dynamical strategy could be enhanced by more state-of-art statistical models like the deep neural networks, for this I would need to have significantly better computational power. Lastly, I would like to test out the performance of better strategies on more realistic market scenarios. The results of this work are encouraging me to continue this work in these directions.

References

Amaya, D., Christoffersen, P., Jacobs, K. and Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns?, *Journal of Financial Economics* **118**(1): 135–167.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001). The distribution of realized stock return volatility, *Journal of financial economics* **61**(1): 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003). Modeling and forecasting realized volatility, *Econometrica* **71**(2): 579–625.

Anderson, E., Hansen, L. P. and Sargent, T. (2000). Robustness, detection and the price of risk, *Manuscript, Stanford* **19**.

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

Bakshi, G. and Madan, D. (2000). Spanning and derivative-security valuation, *Journal of financial economics* **55**(2): 205–238.

Bali, T., Engle, R. F. and Murray, S. (2016). *Empirical Asset Pricing*, John Wiley and Sons, Inc.

Bali, T. G. and Hovakimian, A. (2009). Volatility spreads and expected stock returns, *Management Science* **55**(11): 1797–1812.

Bams, D., Blanchard, G., Honarvar, I. and Lehnert, T. (2017). Does oil and gold price uncertainty matter for the stock market?, *Journal of Empirical Finance* **44**: 270–285.

Barndorff-Nielsen, O. E., Kinnebrock, S. and Shephard, N. (2008). Measuring downside risk-realised semivariance, *SSRN id-1262194* .

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(2): 253–280.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options, *The Review of Financial Studies* **9**(1): 69–107.

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

Bloomberg (2018). Bloomberg Terminal of the Corvinus University of Budapest.

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

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

Bonomo, M., Garcia, R., Meddahi, N. and Tédongap, R. (2010). Generalized disappointment aversion, long-run volatility risk, and asset prices, *The Review of Financial Studies* **24**(1): 82–122.

Borochin, P., Chang, H. and Wu, Y. (2018). The information content of the term structure of risk-neutral skewness, *SSRN id-2971057* .

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

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

Carr, P. and Wu, L. (2008). Variance risk premiums, *The Review of Financial Studies* **22**(3): 1311–1341.

Chevallier, J. (2013). Variance risk-premia in CO2 markets, *Economic Modelling* **31**: 598–605.

Chevallier, J., Sévi, B. et al. (2014). A fear index to predict oil futures returns, *Technical report*.

Chow, V., Jiang, W. and Li, J. V. (2014). Does VIX truly measure return volatility?, *SSRN id-2489345* .

Chow, V., Jiang, W. and Li, J. V. (2016). Vix decomposition, the price of fear and stock return predictability, *SSRN id-2747169* .

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues, *Taylor & Francis* .

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

Diebold, F. X. and Li, C. (2006). Forecasting the term structure of government bond yields, *Journal of econometrics* **130**(2): 337–364.

Du, J. and Kapadia, N. (2012). Tail and volatility indices from option prices, *Unpublished working paper. University of Massachusetts, Amhurst* .

Duffie, D. (1992). *Dynamic Asset Pricing Theory*, Princeton, NJ: Princeton University Press.

Dupire, B. et al. (1994). Pricing with a smile, *Risk* **7**(1): 18–20.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds, *Journal of financial economics* **33**(1): 3–56.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests, *Journal of political economy* **81**(3): 607–636.

Feunou, B., Jahan-Parvar, M. R. and Okou, C. (2017). Downside variance risk premium, *Journal of Financial Econometrics* **16**(3): 341–383.

Goodfellow, I., Bengio, Y., Courville, A. and Francis, B. (2017). *Deep Learning*, Adaptive Computation and Machine Learning Series.

Gu, S., Kelly, B. T. and Xiu, D. (2018). Empirical asset pricing via machine learning, *SSRN id-3159577* .

Harris, R. D. and Qiao, F. (2018). Moment risk premia and the cross-section of stock returns, *SSRN id-2845138* .

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning*, Springer Series in Statistics.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The review of financial studies* **6**(2): 327–343.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R*, Springer Texts in Statistics.

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

Luo, X. and Zhang, J. E. (2017). Expected stock returns and forward variance, *Journal of Financial Markets* **34**: 95–117.

Merton, R. C. (1973). Theory of rational option pricing, *The Bell Journal of economics and management science* pp. 141–183.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous, *Journal of financial economics* **3**(1-2): 125–144.

Márkus, L. (2017). *Eszközár Folyamatok Modellezése, valamint Európai és Amerikai opciók árazása*, ELTE egyetemi jegyzet.

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves, *Journal of business* pp. 473–489.

Newey, W. K. and West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

OptionMetrics (2018). Optionmetrics volatility surface database.

Park, H.-h., Kim, B. and Shim, H. (2018). A smiling bear in the equity options market and the cross-section of stock returns, *SSRN id-2632763* .

Park, Y.-H. (2013). Volatility of volatility and tail risk premiums, *SSRN id-2341734* .

Patton, A. J. and Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility, *Review of Economics and Statistics* **97**(3): 683–697.

Prokopczuk, M. and Wese Simen, C. (2014). Variance risk premia in commodity markets, *SSRN id-2195691* .

Rossi, A. G. and Timmermann, A. (2015). Modeling covariance risk in Merton's icapm, *The Review of Financial Studies* **28**(5): 1428–1461.

Sharpe, W. F. (1966). Mutual fund performance, *The Journal of business* **39**(1): 119–138.

Shreve, S. (2008). *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer.

Simity, M. (2018). Predicting index returns using risk aversion measures from the volatility surface, *Corvinus University of Budapest Master Thesis* .

10 Appendix

Table 9: Tickers

Ticker	Company Name	Industry
AMD	Advanced Micro Devices Inc	Information Technology
BA	Boeing Company	Industrials
BAC	Bank of America Corp	Financials
CMS	CMS Energy	Utilities
DIS	The Walt Disney Company	Communication Services
DUK	Duke Energy	Utilities
F	Ford Motor	Consumer Discretionary
FDX	FedEx Corporation	Industrials
GE	General Electrics	Industrials
GPS	Gap Inc	Consumer Discretionary
HOG	Harley-Davidson	Consumer Discretionary
IBM	International Business Machines	Information Technology
INTC	Intel Corp.	Information Technology
JNJ	Johnson & Johnson	Health Care
KO	Coca-Cola Company	Consumer Staples
MCD	McDonald's Corp.	Consumer Discretionary
MSFT	Microsoft Corp.	Information Technology
NKE	Nike	Consumer Discretionary
ORCL	Oracle Corp.	Information Technology
PFE	Pfizer Inc.	Health Care
PG	Procter & Gamble	Consumer Staples
VZ	Verizon Communications	Communication Services
WFC	Wells Fargo	Financials
WMT	Walmart	Consumer Staples
OEX	S&P 100	Index
SPX	S&P 500	Index

10.1 Empirical Properties tables

Table 10: Empirical Properties of Implied Moments

	VRP			SRP			KRP		
	Mean	Std	t	Mean	Std	t	Mean	Std	t
AMD	-2.713	2.733	-6.197	-1.3%	0.448	-0.457	3.06%	0.174	2.823
BA	-0.785	0.653	-8.600	-0.2%	0.048	-0.662	0.10%	0.008	1.878
BAC	-0.946	2.812	-2.789	-0.6%	0.255	-0.421	1.89%	0.119	1.405
CMS	-1.258	1.798	-7.939	-1.5%	0.198	-1.257	0.15%	0.059	0.423
DIS	-0.748	0.727	-7.750	0.0%	0.049	0.068	0.11%	0.008	2.242
DUK	-0.019	0.204	-1.136	-0.1%	0.024	-0.969	0.08%	0.006	2.142
F	-1.509	2.343	-4.567	-2.1%	0.400	-0.920	1.42%	0.179	1.326
FDX	-0.702	0.714	-5.977	0.2%	0.031	1.130	0.06%	0.004	2.635
GE	-0.666	1.177	-4.580	0.4%	0.034	1.912	0.00%	0.010	-0.091
GPS	-1.338	1.341	-7.013	-0.4%	0.124	-0.584	0.39%	0.024	2.782
HOG	-1.073	1.525	-4.779	0.7%	0.078	1.469	0.15%	0.015	2.463
IBM	-0.558	0.686	-5.892	0.2%	0.037	0.862	0.07%	0.005	2.463
INTC	-0.807	1.019	-5.215	-0.5%	0.090	-0.783	0.31%	0.019	2.410
JNJ	-0.222	0.312	-5.866	0.0%	0.020	0.138	0.04%	0.004	1.703
KO	-0.003	0.163	-0.131	0.0%	0.013	0.263	0.06%	0.002	2.809
MCD	-0.469	0.465	-7.481	0.1%	0.016	0.829	0.02%	0.002	1.791
MSFT	-0.333	0.564	-4.230	0.1%	0.050	0.413	0.20%	0.008	2.782
NKE	0.069	0.261	1.444	-0.1%	0.072	-0.334	0.32%	0.014	2.848
ORCL	-0.914	1.543	-3.453	-0.1%	0.220	-0.077	0.99%	0.065	2.116
PFE	-0.500	0.489	-8.576	0.0%	0.018	0.364	0.02%	0.003	2.030
PG	-0.137	0.317	-3.346	-1.3%	0.191	-1.028	0.49%	0.067	1.090
SPX	-0.378	0.363	-8.865	0.1%	0.006	3.066	0.00%	0.001	0.136
VZ	-0.320	0.498	-5.811	-0.1%	0.021	-1.233	0.02%	0.003	1.284
WFC	-0.631	2.112	-2.510	2.4%	0.187	1.773	0.69%	0.060	1.994
WMT	-0.470	0.552	-5.525	0.2%	0.016	2.434	0.03%	0.002	2.433
OEX	-0.365	0.398	-7.461	0.1%	0.006	4.142	0.00%	0.001	-0.615

10.2 Time-varying regression models

Table 11: VRP Expectation Hypothesis Regression

	a	t	b	Confidence int.		R squared
AMD	0.0085	4.2345	0.1162	0.0469	0.1855	0.1236
BA	-0.0016	-2.3949	0.1390	0.1163	0.1618	0.7176
BAC	0.0009	0.6486	0.1772	0.1490	0.2054	0.7759
CMS	-0.0003	-0.2206	0.0790	0.0522	0.1058	0.4746
DIS	-0.0027	-4.4966	0.1562	0.1359	0.1764	0.7275
DUK	0.0085	4.2345	0.1162	0.0469	0.1855	0.1236
F	-0.0002	-0.1274	0.1184	0.0918	0.1451	0.5590
FDX	-0.0015	-2.3551	0.1372	0.1157	0.1586	0.7056
GE	0.0010	1.3351	0.1044	0.0825	0.1262	0.6542
GPS	0.0003	0.2021	0.1074	0.0789	0.1359	0.4966
HOG	-0.0014	-1.5387	0.1266	0.1101	0.1431	0.7278
IBM	0.0002	0.5742	0.1071	0.0928	0.1214	0.5694
INTC	-0.0010	-0.7435	0.1524	0.1236	0.1812	0.5821
JNJ	0.0005	0.7694	0.1115	0.0591	0.1639	0.2844
KO	0.0051	8.2492	-0.1148	-0.1636	-0.0660	0.0413
MCD	-0.0006	-1.9489	0.1181	0.1010	0.1351	0.6483
MSFT	0.0064	3.3600	0.0660	-0.0071	0.1391	0.0678
NKE	0.0097	9.1633	-0.0808	-0.1270	-0.0346	0.0124
ORCL	0.0018	1.0208	0.1235	0.0963	0.1508	0.5526
PFE	-0.0013	-2.4280	0.1359	0.1121	0.1596	0.5838
PG	0.0038	2.4905	0.0433	-0.0513	0.1379	0.0078
SPX	-0.0016	-3.7215	0.1548	0.1229	0.1867	0.7739
VZ	0.0010	2.2432	0.1409	0.1115	0.1702	0.6518
WFC	0.0022	2.3364	0.1553	0.1259	0.1846	0.7434
WMT	-0.0009	-2.3983	0.1299	0.1098	0.1500	0.7491
OEX	-0.0013	-3.0688	0.1414	0.1095	0.1734	0.7528

Table 12: SRP Expectation Hypothesis Regression

	a	**t**	**b**	**Confidence int.**		**R squared**
AMD	0.0005	1.7088	-0.0317	-0.2111	0.1477	0.0004
BA	-0.0003	-1.4453	-0.7601	-2.5473	1.0271	0.0061
BAC	0.0007	1.1714	3.1321	-0.4585	6.7228	0.1132
CMS	-0.0017	-1.4676	0.6541	0.0677	1.2404	0.0125
DIS	0.0000	-0.1762	-0.0870	-0.3051	0.1310	0.0005
DUK	0.0005	1.7088	-0.0317	-0.2111	0.1477	0.0004
F	-0.0031	-1.3366	-1.9688	-3.5937	-0.3439	0.0941
FDX	0.0003	3.0136	2.2809	0.6583	3.9034	0.0759
GE	0.0000	0.2648	0.0829	-0.6594	0.8252	0.0015
GPS	-0.0012	-1.7848	-0.2843	-1.8395	1.2708	0.0008
HOG	0.0005	1.5793	0.8801	-1.1609	2.9211	0.0346
IBM	0.0000	-0.2029	0.1128	-0.6615	0.8871	0.0002
INTC	-0.0007	-1.4752	0.3947	-0.5920	1.3815	0.0019
JNJ	0.0001	0.9222	2.3256	-2.7290	7.3803	0.0363
KO	0.0000	0.3007	0.2709	-0.1774	0.7193	0.0004
MCD	0.0000	0.1635	0.6613	-0.6362	1.9587	0.0148
MSFT	-0.0003	-1.1149	-2.7846	-4.9022	-0.6670	0.0284
NKE	-0.0003	-0.7542	-1.1241	-4.8862	2.6380	0.0001
ORCL	0.0003	1.1655	0.6625	-2.8102	4.1352	0.0037
PFE	-0.0001	-1.2802	0.1133	-0.0721	0.2987	0.0017
PG	-0.0014	-1.2172	-0.4420	-1.4096	0.5256	0.0001
SPX	0.0001	1.7914	1.0016	0.0505	1.9527	0.1486
VZ	0.0001	0.9556	-0.0545	-0.2814	0.1724	0.0012
WFC	0.0013	1.5520	-0.4177	-1.0690	0.2337	0.0195
WMT	0.0000	0.0083	-0.8139	-1.7478	0.1201	0.0245
OEX	0.0000	1.1487	0.5224	-0.2620	1.3068	0.0616

Table 13: KRP Expectation Hypothesis Regression

	a	t	b	Confidence int.		R squared
AMD	0.0027	2.9622	0.5851	-0.1065	1.2766	0.0248
BA	0.0000	-0.0359	7.3114	3.1129	11.5100	0.2875
BAC	0.0045	2.2026	10.3389	8.1462	12.5316	0.6152
CMS	0.0023	1.8501	2.4347	-0.2929	5.1622	0.0423
DIS	0.0001	0.8143	9.1873	5.3448	13.0299	0.2807
DUK	0.0027	2.9622	0.5851	-0.1065	1.2766	0.0248
F	0.0069	1.2173	6.7691	3.9429	9.5952	0.0682
FDX	0.0002	2.1976	4.6536	2.2808	7.0263	0.2392
GE	0.0006	3.2721	2.0622	1.1402	2.9843	0.2961
GPS	0.0019	4.0263	4.6407	1.0700	8.2113	0.0971
HOG	0.0010	2.5040	3.8009	2.2431	5.3586	0.4504
IBM	0.0003	3.2621	3.0696	1.6819	4.4573	0.0864
INTC	0.0003	1.2458	9.8348	4.2308	15.4387	0.2279
JNJ	0.0000	0.0212	13.5435	3.2895	23.7975	0.1790
KO	0.0002	3.4461	0.0890	-3.3205	3.4985	0.0000
MCD	0.0001	2.3458	4.0144	1.3937	6.6352	0.2097
MSFT	0.0007	2.9373	8.3660	5.0231	11.7088	0.1125
NKE	0.0014	3.1626	-21.5362	-37.4213	-5.6511	0.0027
ORCL	0.0007	1.8193	6.5504	2.9888	10.1120	0.3447
PFE	0.0002	3.5185	1.4961	-0.1685	3.1608	0.0971
PG	0.0026	1.2326	-3.9410	-16.8676	8.9856	0.0001
SPX	0.0000	-2.5755	6.9949	5.5063	8.4835	0.6220
VZ	0.0000	-0.4425	6.1872	3.7098	8.6646	0.4047
WFC	0.0037	2.4992	4.7431	3.4799	6.0064	0.3690
WMT	0.0000	1.5768	5.0528	3.5344	6.5712	0.4642
OEX	0.0000	-3.0019	5.9292	4.3394	7.5191	0.6092

10.3 Diebold-Li coefficients

Table 14: Diebold-Li Implied volatility coefficients

	Level			**Slope**			**Curvature**		
Ticker	Mean	Std	t	Mean	Std	t	Mean	Std	t
AMD	2.001	0.727	24.339	1.304	2.705	3.414	0.865	2.593	4.036
BA	0.857	0.412	10.668	0.098	0.519	1.577	-0.329	0.617	-6.446
BAC	0.858	0.835	7.025	0.445	3.142	1.088	-0.712	2.721	-3.749
CMS	0.924	0.876	10.336	0.515	1.384	4.244	-0.897	2.285	-10.213
DIS	0.755	0.439	9.210	0.146	0.582	2.233	-0.271	0.571	-9.081
DUK	0.357	0.471	4.293	-0.267	0.495	-3.123	-0.251	0.544	-5.706
F	1.285	0.839	16.371	0.510	2.622	1.696	-0.127	1.921	-0.924
FDX	0.752	0.426	8.302	0.113	0.518	1.590	-0.223	0.445	-8.762
GE	0.664	0.707	5.358	0.157	0.792	1.978	-0.379	1.084	-6.202
GPS	1.097	0.708	8.345	0.525	1.003	4.409	-0.342	1.037	-7.012
HOG	0.927	0.608	8.962	0.392	1.385	1.984	-0.346	1.216	-5.210
IBM	0.587	0.398	6.644	0.104	0.476	2.049	-0.256	0.572	-6.292
INTC	0.831	0.580	7.007	0.211	0.731	2.223	-0.295	0.647	-8.625
JNJ	0.341	0.183	11.539	-0.047	0.226	-2.169	-0.222	0.275	-12.779
KO	0.171	0.141	7.838	-0.092	0.121	-6.593	-0.218	0.228	-11.518
MCD	0.527	0.348	7.659	0.042	0.271	1.613	-0.219	0.363	-10.810
MSFT	0.548	0.342	9.077	-0.075	0.407	-1.542	-0.229	0.498	-7.401
NKE	0.077	0.185	1.607	-0.014	0.064	-3.205	-0.065	0.170	-5.325
ORCL	0.872	0.704	5.404	0.332	1.215	1.825	-0.237	1.057	-3.473
PFE	0.605	0.422	7.410	0.019	0.300	0.737	-0.249	0.404	-12.513
PG	0.295	0.200	9.295	-0.078	0.220	-3.413	-0.226	0.311	-9.339
SPX	0.492	0.195	14.972	-0.046	0.341	-1.217	-0.187	0.357	-6.570
VZ	0.526	0.350	8.860	-0.109	0.450	-2.176	0.140	0.605	3.607
WFC	0.567	0.783	5.032	0.295	1.997	1.163	-0.394	2.060	-2.684
WMT	0.588	0.459	5.270	-0.003	0.290	-0.128	-0.264	0.377	-11.160
OEX	0.413	0.239	9.724	0.017	0.349	0.414	-0.140	0.369	-4.734

Table 15: Diebold-Li Implied skewness coefficients

	Level			**Slope**			**Curvature**		
Ticker	Mean	Std	t	Mean	Std	t	Mean	Std	t
AMD	2.861	2.534	6.067	-2.437	2.276	-5.778	-5.507	4.528	-6.982
BA	0.233	0.553	1.677	-0.178	0.480	-1.495	-0.569	0.995	-2.238
BAC	0.428	1.402	2.205	-0.331	1.246	-2.017	-1.080	2.518	-2.765
CMS	0.085	1.124	0.927	-0.037	1.058	-0.409	-0.443	1.690	-4.011
DIS	0.287	0.651	1.748	-0.231	0.565	-1.655	-0.633	1.176	-2.142
DUK	-0.515	2.181	-1.836	0.449	1.936	1.729	0.934	3.781	2.154
F	0.700	2.056	2.060	-0.577	1.875	-1.880	-1.480	3.451	-2.784
FDX	0.376	0.766	1.676	-0.300	0.665	-1.544	-0.842	1.385	-2.128
GE	-0.094	0.790	-0.802	0.103	0.716	0.958	0.007	1.187	0.041
GPS	0.436	1.146	1.639	-0.334	1.011	-1.435	-1.087	2.065	-2.332
HOG	0.270	0.883	1.543	-0.195	0.779	-1.266	-0.768	1.504	-2.587
IBM	0.261	0.570	1.668	-0.207	0.489	-1.571	-0.598	1.071	-1.965
INTC	0.321	0.819	3.002	-0.252	0.709	-2.898	-0.762	1.523	-2.769
JNJ	0.148	0.284	2.682	-0.113	0.236	-2.531	-0.359	0.570	-3.042
KO	0.279	0.387	3.965	-0.226	0.325	-3.933	-0.596	0.767	-4.052
MCD	0.235	0.449	2.530	-0.189	0.386	-2.377	-0.518	0.829	-2.990
MSFT	0.254	0.758	1.820	-0.207	0.667	-1.754	-0.546	1.346	-2.064
NKE	0.042	0.141	1.988	-0.031	0.117	-1.832	-0.105	0.286	-2.325
ORCL	0.545	1.246	1.464	-0.448	1.091	-1.406	-1.165	2.298	-1.650
PFE	0.163	0.533	2.141	-0.125	0.460	-1.974	-0.400	0.976	-2.478
PG	0.200	0.414	2.879	-0.158	0.347	-2.852	-0.453	0.820	-2.921
SPX	-0.088	0.412	-1.011	0.086	0.352	1.171	0.092	0.777	0.539
VZ	0.328	0.474	3.557	-0.284	0.421	-3.418	-0.560	0.770	-4.326
WFC	0.084	1.079	1.117	-0.037	0.991	-0.545	-0.443	1.358	-3.012
WMT	0.327	0.677	1.757	-0.265	0.580	-1.684	-0.702	1.266	-1.969
OEX	0.039	0.254	0.779	-0.019	0.219	-0.439	-0.170	0.469	-1.808

10.4 Regression coefficients

Table 16: Explanatory regression coefficients A

	VRP	SRP	KRP	DVRP	UVRP	TRP	DSRP	USRP
AMD	0.222	3.240	0.785	-0.611	-0.591	-2.060	-2.111	0.414
	2.178	0.876	0.104	-1.192	-1.468	-0.841	-1.314	2.663
BA	0.533	-3.673	51.564	1.436	-0.097	-23.668	-5.129	3.091
	2.421	-0.215	1.019	1.181	-0.180	-2.969	-2.146	2.440
BAC	0.208	-4.770	12.363	0.431	0.218	-1.168	0.883	-0.340
	1.534	-0.465	1.496	0.241	1.600	-0.518	1.077	-0.345
CMS	0.210	-0.584	10.037	0.672	0.323	-4.586	-0.882	0.217
	2.042	-1.531	1.694	0.607	1.715	-0.759	-0.558	2.149
DIS	0.495	-9.730	72.075	0.831	0.237	-9.614	2.144	0.770
	2.247	-0.813	1.315	0.533	0.480	-1.675	0.674	1.692
DUK	-1.786	-0.312	-62.058	-3.888	0.663	-4.968	-1.623	0.327
	-2.442	-1.785	-0.044	-0.666	1.268	-1.431	-0.325	2.524
F	0.115	0.447	3.127	0.241	0.181	0.380	0.482	0.026
	0.922	0.249	0.806	0.338	0.797	0.155	0.205	0.735
FDX	0.435	-42.097	170.298	0.565	0.326	-16.624	4.411	3.567
	2.057	-3.776	1.792	0.655	0.548	-1.557	1.387	0.876
GE	0.286	-22.825	14.792	0.973	0.244	-14.246	-2.291	0.410
	1.909	-1.843	0.819	0.383	0.757	-2.261	-0.304	1.341
GPS	0.198	-2.073	-0.281	0.610	-0.064	-4.354	-1.872	1.071
	1.716	-0.337	-0.017	0.723	-0.154	-1.712	-1.586	1.634
HOG	0.230	-18.697	14.593	0.739	0.663	-7.351	-2.618	0.975
	1.619	-1.587	1.475	0.521	2.383	-1.652	-0.569	1.298
IBM	0.161	-13.753	-3.288	0.904	0.306	-17.163	-6.634	2.295
	0.934	-0.795	-0.070	1.748	0.576	-2.886	-1.424	1.774
INTC	0.536	-10.707	70.742	0.906	-0.052	-8.489	2.934	0.578
	2.845	-2.048	1.269	0.941	-0.143	-1.822	0.644	2.085
JNJ	0.252	0.580	-160.489	1.233	0.072	-29.096	-7.745	1.069
	1.011	1.906	0.025	1.551	0.123	-2.409	-0.675	1.810
KO	-0.884	-171.383	-440.997	-4.414	-0.169	-35.295	5.341	-14.843
	-1.476	-2.685	-1.633	-0.854	-0.280	-3.011	2.337	-2.449
MCD	0.312	-9.751	59.826	-0.640	-0.070	-16.228	-0.126	1.189
	1.415	-0.401	0.855	-0.022	-0.083	-1.354	-0.446	1.036
MSFT	-0.021	-17.904	-119.122	-0.035	-0.153	-11.397	-2.434	0.982
	-0.098	-1.419	-1.578	-0.637	-0.329	-3.586	-0.487	0.042
NKE	-2.554	-0.626	-83.915	-6.692	0.411	-20.695	-0.175	1.377
	-2.185	-0.071	-1.617	-0.061	0.947	-1.760	-0.680	2.459
ORCL	0.275	-3.069	22.393	-0.844	-0.020	-5.566	-3.042	1.458
	2.323	-0.280	1.025	-1.272	-0.072	-1.401	-1.610	2.769
PFE	0.244	-11.191	-4.751	-0.506	-0.163	-37.051	-0.218	0.095
	0.982	-0.716	-0.163	-0.060	-0.234	-2.786	-0.066	0.940
PG	0.043	2.563	-7.909	-0.271	0.111	-28.528	-1.652	0.239
	0.133	2.881	-3.573	-2.033	0.746	-1.448	-0.968	0.297
VZ	0.167	7.538	-15.608	-0.173	0.404	-13.819	-5.939	5.224
	0.803	0.526	-0.254	-1.505	0.527	-2.361	-2.262	0.564
WFC	0.071	0.203	-1.404	-0.824	0.110	-1.352	-4.827	0.868
	0.432	0.045	-0.198	-1.370	1.263	-0.680	-1.041	1.249
WMT	0.033	-12.737	-69.831	0.176	0.203	-26.745	-2.428	3.508
	0.158	-0.522	-0.741	0.386	0.303	-2.154	-0.927	0.279
OEX	0.518	-11.818	121.939	1.600	0.257	-50.357	-5.204	-7.305
	2.475	-3.571	1.603	0.671	0.516	-4.122	-0.973	-1.920

Table 17: Explanatory regression coefficients B

	Vlevel	Vslope	Vcurvature	Slevel	Sslope	Scurvature
AMD	0.269	-0.018	-0.149	-0.055	0.070	0.019
	1.285	-0.297	-2.763	-0.851	0.932	0.581
BA	-0.123	0.030	0.052	-0.030	0.039	0.009
	-0.569	0.294	0.591	-0.268	0.304	0.143
BAC	0.199	0.034	-0.036	0.042	-0.042	-0.037
	2.065	1.734	-1.351	0.818	-0.754	-1.180
CMS	0.122	0.031	-0.009	-0.011	0.020	-0.022
	2.231	1.161	-0.377	-0.383	0.650	-0.868
DIS	0.200	-0.035	0.044	-0.020	0.024	0.009
	1.070	-0.349	0.384	-0.188	0.198	0.142
DUK	0.373	-0.328	-0.411	0.118	-0.127	-0.075
	1.227	-1.108	-1.692	1.277	-1.222	-1.407
F	0.139	0.062	0.058	0.014	-0.020	-0.001
	1.329	2.389	1.433	0.348	-0.464	-0.028
FDX	0.150	0.035	0.078	0.061	-0.067	-0.037
	0.959	0.343	0.679	0.867	-0.851	-0.912
GE	0.087	0.001	0.024	-0.023	0.024	0.020
	1.141	0.015	0.570	-0.444	0.425	0.617
GPS	0.119	-0.011	0.115	0.053	-0.059	-0.031
	0.903	-0.163	2.113	0.800	-0.785	-0.818
HOG	0.080	0.070	-0.068	-0.057	0.064	0.024
	0.650	1.765	-1.390	-0.796	0.814	0.518
IBM	0.035	0.016	-0.018	-0.038	0.042	0.024
	0.222	0.154	-0.248	-0.448	0.434	0.501
INTC	0.034	-0.077	0.124	0.037	-0.041	-0.020
	0.205	-0.954	1.494	0.426	-0.418	-0.410
JNJ	0.085	0.011	-0.028	0.056	-0.058	-0.038
	0.326	0.089	-0.216	0.404	-0.357	-0.535
KO	-0.189	0.225	0.036	-0.075	0.094	0.033
	-0.514	0.577	0.156	-0.602	0.643	0.505
MCD	0.498	0.092	0.046	0.096	-0.111	-0.053
	2.798	0.489	0.388	0.836	-0.835	-0.827
MSFT	0.184	-0.110	-0.066	0.000	0.004	-0.003
	0.774	-0.752	-0.534	0.003	0.046	-0.066
NKE	1.342	-1.362	-0.955	0.405	-0.471	-0.212
	0.461	-0.413	-0.639	0.524	-0.501	-0.576
ORCL	0.100	0.026	0.030	0.020	-0.017	-0.016
	0.827	0.614	0.614	0.360	-0.283	-0.543
PFE	-0.009	0.158	-0.017	-0.080	0.091	0.044
	-0.058	1.160	-0.127	-0.905	0.893	0.872
PG	0.234	-0.105	-0.058	0.083	-0.094	-0.048
	0.864	-0.770	-0.442	0.951	-0.922	-1.021
VZ	0.181	0.088	-0.062	-0.008	0.003	0.021
	1.065	0.997	-0.807	-0.069	0.021	0.288
WFC	0.046	0.013	0.004	-0.033	0.040	0.005
	0.726	0.809	0.378	-0.802	0.880	0.133
WMT	0.016	-0.138	0.073	-0.007	0.006	0.007
	0.139	-0.952	0.741	-0.116	0.084	0.203
OEX	0.172	-0.036	0.063	-0.239	0.279	0.127
	0.870	-0.415	0.745	-1.683	1.719	1.554

Table 18: PCA regressions

	VRP1	**VRP2**	**VRP3**	**SRP1**	**SRP2**	**SRP3**
AMD	-7.01E-05	-1.144	1.636	-7.01E-05	-1.961	8.308
	-0.47	-2.670	0.628	-0.47	-1.683	1.740
BA	2.99E-05	-2.326	-6.332	2.99E-05	3.036	-50.496
	0.44	-2.585	-1.244	0.44	1.133	-2.583
BAC	-2.60E-05	-0.691	1.484	-2.60E-05	-1.930	-12.139
	-0.35	-1.139	0.636	-0.35	-0.600	-1.069
CMS	1.20E-04	-1.103	-0.299	1.20E-04	-1.411	-17.346
	1.54	-2.365	-0.159	1.54	-0.922	-2.498
DIS	1.02E-04	-1.941	-5.423	1.02E-04	-1.326	-30.627
	1.44	-2.375	-1.181	1.44	-0.554	-1.527
DUK	2.59E-05	-3.592	-11.608	2.59E-05	-3.378	-7.033
	0.27	-1.004	-2.125	0.27	-2.079	-0.716
F	1.51E-04	-1.131	-4.413	1.51E-04	-2.580	1.351
	1.81	-1.659	-1.271	1.81	-1.218	0.160
FDX	2.30E-06	-1.276	-6.689	2.30E-06	-0.148	-16.623
	0.04	-1.467	-1.190	0.04	-0.072	-0.910
GE	6.23E-05	-1.261	0.806	6.23E-05	-2.682	-26.135
	0.97	-1.907	0.194	0.97	-1.052	-3.079
GPS	7.64E-05	-0.779	1.605	7.64E-05	-2.313	-5.171
	0.78	-1.637	0.528	0.78	-1.128	-1.001
HOG	-2.82E-05	-1.156	-7.732	-2.82E-05	-3.369	-14.822
	-0.38	-1.674	-2.326	-0.38	-1.243	-1.300
IBM	9.15E-06	-0.900	-3.364	9.15E-06	-1.024	-12.726
	0.14	-1.239	-0.738	0.14	-0.418	-0.714
INTC	8.02E-05	-1.623	-3.460	8.02E-05	3.077	-2.220
	0.87	-2.607	-0.621	0.87	1.254	-0.212
JNJ	9.53E-06	-1.583	-1.566	9.53E-06	1.327	-21.380
	0.21	-1.504	-0.311	0.21	0.320	-1.026
KO	4.05E-05	-1.449	18.653	4.05E-05	-5.251	-29.236
	0.78	-0.979	1.709	0.78	-1.659	-0.810
MCD	5.71E-05	-1.795	-1.673	5.71E-05	0.351	-43.897
	1.03	-2.162	-0.342	1.03	0.120	-2.310
MSFT	1.17E-04	-0.975	-0.645	1.17E-04	1.267	-20.722
	1.50	-1.222	-0.161	1.50	0.734	-2.066
NKE	6.12E-05	11.704	2.532	6.12E-05	-6.735	116.701
	0.80	1.820	0.049	0.80	-0.717	2.087
ORCL	7.96E-05	-1.149	-2.940	7.96E-05	-1.711	-14.611
	0.81	-2.495	-0.762	0.81	-1.072	-1.711
PFE	3.23E-05	-0.750	-2.585	3.23E-05	-1.071	-6.842
	0.59	-0.854	-0.483	0.59	-0.410	-0.308
PG	1.86E-05	-1.347	4.440	1.86E-05	-3.640	-8.546
	0.30	-1.189	0.558	0.30	-0.900	-0.576
VZ	4.97E-05	-1.160	-1.040	4.97E-05	0.900	-9.655
	0.83	-1.201	-0.299	0.83	0.364	-0.443
WFC	9.64E-06	-0.477	-3.149	9.64E-06	-0.716	5.533
	0.18	-0.712	-0.943	0.18	-0.233	0.609
WMT	1.76E-05	-0.600	0.353	1.76E-05	-0.957	-4.399
	0.29	-0.805	0.068	0.29	-0.384	-0.311
OEX	6.03E-05	-2.133	-6.236	6.03E-05	4.334	-46.254
	1.57	-2.105	-1.132	1.57	1.325	-2.156