Essays on Asset Pricing Using Option-Implied Information

KAGKADIS, ANASTASIOS

Essays on Asset Pricing Using Option-Implied Information

by
Anastasios Kagkadis

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Durham University Business School
November 2014

To my grandparents.

Abstract

The forward-looking nature of the options market makes it an ideal environment for investigating the determinants and the information content of investors' expectations about the future. Therefore, this thesis explores the interrelations arising between the macroeconomic and stock market environment, and the S&P 500 index options market.

First, we examine how investors' sentiment driven by macroeconomic fundamentals and investors' erroneous beliefs impact the risk-neutral skewness. Our findings reveal that the macroeconomic fundamentals component of investor sentiment is the main driving force of risk-neutral skewness throughout the whole sample period, while the error in investors' beliefs has limited explanatory power and only during the earlier years examined. Moreover, we show that the fundamentals component of investor sentiment affects differently the prices of call and put options. Second, we extend the concept of risk-neutral skewness by creating measures of forward skewness and gauge their predictive ability for a wide range of macroeconomic variables, asset prices, as well as systemic risk, crash risk, and uncertainty variables. Overall, we document that forward skewness encapsulates important information about future macroeconomic and financial market conditions for horizons up to one year ahead over and above forward variance. Third, we propose a novel measure of dispersion in expectations that is derived from the dispersion of options' trading volume across strike prices. We show that dispersion consistently forecasts negative excess market returns, for horizons up to two years ahead, exhibiting a predictive ability comparable to that of the variance risk premium and outperforming all other variables considered.

This thesis contributes to the asset pricing and macro-finance literature by unravelling the determinants of the pricing kernel, showing that the call and put options markets are segmented and revealing that option prices and trading volume have significant forecasting ability for many aspects of the macroeconomic and financial environment. In that respect our findings are of particular interest not only to academics but also to investors and policy makers.

Acknowledgements

First and foremost, I would like to express my deep gratitude and appreciation to my advisors Dr. Dennis Philip and Dr. Panayiotis Andreou for their continuous encouragement, guidance and support through the completion of this thesis. Without them this journey would not have been the same. I am really proud of being one of their students. I am also grateful to Prof. Paulo Maio for his valuable comments on the third empirical chapter of the thesis.

I would also like to thank Konstantinos Gavriilidis and Dima Smayra for their friendship and support during the years of my study at Durham. A huge thank you goes to my close friends, Symeon Chouvardas, Aristofanis Sakarellos and Maria Frantzela for their endless encouragement all these years. Moreover, special thanks go to Ioannis Koutrakos for his generous hospitality during my several visits to London.

Finally, I would like to deeply thank my family and especially my parents Sofia and Konstantinos and my sisters Christina and Aggeliki. Whatever I have achieved or will achieve in the future is a direct outcome of their continuous love and support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The classic work of Black and Scholes (1973), Merton (1973) and Black (1976) in option pricing assumes that the underlying asset price follows a geometric Brownian motion with constant volatility. This implies that the option value can be replicated by a dynamic hedging strategy in the underlying asset and the risk-free rate and hence options are *redundant* securities. Rubinstein (1994) shows that while index option prices have been roughly consistent with the Black-Scholes-Merton assumptions before the crash of October 1987, this is not the case in the post-1987 crash period. In particular, a typical post-crash plot of implied volatilities across strike prices exhibits a convex and explicitly downward-sloping pattern, which constitutes a strong violation of the Black-Scholes-Merton model.

To account for this anomaly, researchers typically either introduce additional risk factors such as stochastic volatility and/or jumps in the underlying asset price process (Bakshi et al., 1997, Bates, 2000, Chernov and Ghysels, 2000, Pan, 2002 and Jones, 2003 among others) or examine the role of additional factors that may exert impact on option prices (Peña et al., 1999, Amin et al., 2004, Bollen and Whaley, 2004, Buraschi and Jiltsov, 2006, Han, 2008, Gârleanu et al., 2009, among others). The common element in both approaches is the implication that options are not redundant securities. In fact, Bakshi et al. (2000), Buraschi and Jackwerth (2001), Coval and Shumway (2001), Bakshi and Kapadia (2003) and Jones (2006) strongly reject the hypothesis that options can be considered redundant securities and postulate that their prices reflect exposure to additional priced factors. Buraschi and Jiltsov (2006, p.2841) further assert that since options are not redundant securities,

"...they provide an economic value that at least exceeds the cost of maintaining option exchanges". The fact that options have been shown to be nonredundant assets implies that the options market encapsulates important information that is distinct from that found in the underlying asset market.

Another advantageous characteristic of option contracts is that they are by definition related to investors' *forward-looking* beliefs and risk preferences. For example, the risk-neutral distribution extracted from option prices is derived from the product of the conditional distribution of future returns under the physical measure and the conditional pricing kernel. Moreover, the trading volume and open interest of options of different type, moneyness or maturity are affected by investors' different subjective expectations and can also be linked to informed trading (Easley et al., 1998). Therefore, over the last ten years, a large body of literature has evolved that uses option-related measures for forecasting purposes. For example, Bollerslev et al. (2009), Bali and Hovakimian (2009), Cremers and Weinbaum (2010), Xing et al. (2010), Bakshi et al. (2011) and Feunou et al. (2014) among others make use of the information content of option prices, while Pan and Poteshman (2006), Fodor et al. (2011), Byun and Kim (2013) and Chen et al. (2013) rely on the information embedded in options' trading volume and open interest. The main conclusion drawn from the above literature is that both option prices and options' trading activity exhibit significant forecasting power for stock returns, treasury returns and macroeconomic variables.

In a recent paper, Han (2008) challenges the main assumption of the traditional option pricing models that option prices are settled as if the market consists of a rational representative investor, by showing that option prices can be seen as the weighted average of the expectations of both rational and irrational investors. More specifically, he finds that investor sentiment, which is assumed to capture investors' erroneous beliefs, has a significantly positive impact on the S&P 500 index risk-neutral skewness. Since prior empirical evidence (Aït-Sahalia et al., 2001 and Rosenberg and Engle, 2002) indicates that the conditional physical probability density is approximately symmetric, this result implies that investors' erroneous expectations are a strong determinant of the pricing kernel. Motivated by a stream of papers that relate investor sentiment not only to irrationality but also to rational updating of beliefs (Brown and Cliff, 2005, Lemmon and Portniaguina, 2006 among others), in Chapter 3 we define investor sentiment as investors' overall attitude towards future market returns and postulate that it comprises beliefs driven both by factors related to fundamentals and by factors unrelated to fundamentals. Therefore, we decompose aggregate investor sentiment into an economic fundamentals component and an error in beliefs component. To this end, we employ both a parsimonious set of variables that reflect the information embedded in eight major macroeconomic categories and a set of latent factors that summarize the variations in a large dataset of 131 macroeconomic variables. We examine the sentiment of three different groups of investors - large speculators, investment advisors and individual investors - as it is possible that different investor categories respond differently to fundamentals and also trade differently in the options market. Furthermore, unlike Han (2008) whose sample period ends at 1997:06, our sample period extends up to 2011:06, thus offering us the opportunity to examine whether the relationship between investor sentiment and the S&P 500 index risk-neutral skewness has changed over time. The results presented in Chapter 3 provide clear evidence in favor of this hypothesis. In particular, aggregate investor sentiment has a strong positive impact on risk-neutral skewness during the first period considered (1990:01 - 1997:06) - which largely coincides with Han's (2008) period - but this relationship vanishes in the second and most recent period (1997:07 - 2011:06). By performing the analysis with the two sentiment components, we find that in the first period both components contribute to the significantly positive effect on risk-neutral skewness, while in the second period only the fundamentals component is significantly related to skewness. Moreover, the documented in the second period relationship between the fundamentals component and the risk-neutral skewness is more pronounced during periods of worsened stock market conditions and implies that options traders' beliefs are in line with large speculators' expectations regarding a reversal of recent economic conditions. Finally, we investigate the relationship between the two sentiment components and skewness proxies created separately from call and put options. Our results demonstrate that while the impact of the fundamentals component on puts is in line with the effect on the overall skewness, its impact on calls has the exactly opposite direction, hence supporting Constantinides et al.'s (2011) assertion that the two markets are segmented.

As discussed above, the shape of the risk-neutral distribution reflects investors' forward-looking beliefs and risk aversion. It is important to note, however, that this forward-looking information spans a horizon equal to the maturity of the options used for the risk-neutral distribution estimation. Therefore, on a given day it is possible to construct a term structure of risk-neutral moments that will encompass investors' beliefs and attitudes towards risk for several horizons. This additional information embedded in the term structure of the risk-neutral moments has been recently used by Bakshi et al. (2011) for improving the forecasting power of risk-neutral variance. In particular, they create measures of forward variance for one up to four months ahead using options on the S&P 500 index and show that the estimated forward variances can jointly improve the predictability of future real activity, T-bill returns and stock market returns. In Chapter 4 we extend the concept of forward variances by creating forward skewness coefficients. Our aim is to explicitly capture investors' crash worries for one up to four months ahead and explore their information content. This is of particular importance, since the forward variances estimated by Bakshi et al. (2011) are not robust to the inclusion of jumps in the price process and therefore it is possible that they underestimate the true forward variance. In contrast, our forward moments are based on a newly established technique suggested by Neuberger (2012) and their main characteristic is that they are unbiased estimates of the true forward moments even in the presence of jumps in the price process as long as the asset price is a martingale. We investigate the predictive power of the estimated forward skewness coefficients for various macroeconomic variables, stock market returns as well as risk and uncertainty variables, controlling for the effect of forward variances. We focus on the joint significance of each forward moments group because our primary goal is to investigate whether taking into consideration the term structure of each risk-neutral moment as a whole, is valuable for forecasting purposes or not. The results suggest that indeed the information embedded in the term structure of risk-neutral skewness provides significant improvement in the forecastability of several variables. In particular, forward skewness coefficients are jointly significant for the majority of real activity, money and credit variables examined for horizons up to twelve months ahead, while they also exhibit significant forecasting power for treasury yields for a short 1-month horizon. Furthermore, they are also important for predicting future stock market returns, systemic risk and equity uncertainty especially for a 4-month horizon that matches the time period spanned by the estimated forward skewness coefficients.

In Chapter 5, we exploit the simple fact that trading in options of different strike prices reflects different expectations about future returns, in order to create a novel measure of dispersion in beliefs. More specifically, the proposed measure is derived from the dispersion of the volume-weighted strike prices and captures the dispersion in the beliefs of options traders. The empirical results show that the dispersion in options traders' beliefs forecasts negative excess market returns for horizons up to two years ahead, exhibiting a predictive ability comparable to that of the well-established variance risk premium and outperforming all other traditional predictors considered both in-sample and out-of-sample. Moreover, trading strategies that are based on the out-of-sample forecasting power of the dispersion in beliefs measure at both the aggregate and the portfolio level provide a mean-variance investor with significant utility gains compared to a buy-hold strategy. Furthermore, the information embedded in dispersion is not subsumed by other option-implied measures that proxy for variance and jump risk, or reflect investors' hedging demand. We provide two alternative explanations for the strong negative relationship between dispersion in options traders' beliefs and future market returns. If our measure proxies for the level of disagreement in the underlying asset market, then this result is in line with the limits-to-arbitrage model of Miller (1977) who shows that in the presence of short-sale constraints asset prices are settled according to the opinions of the most optimistic investors since pessimistic investors have no means to express their views. Therefore, higher disagreement leads to higher asset prices and lower subsequent returns. Alternatively, if we consider a framework wherein the underlying asset market participants are homogeneous and update their beliefs based - to some extent- on the trading activity in the options market, then the dispersion in options trading volume across strikes can be regarded as a proxy for the representative investor's ambiguity about the true return generating model. In such a case, the documented negative relationship can be explained within the context of the recursive smooth ambiguity model of Ju and Miao (2012) in case of an elasticity of intertemporal substitution that is lower than one. More specifically, in this setting higher ambiguity increases the pricing kernel but also leads to an increased demand for the risky asset since investors are willing to substitute current consumption with increased future consumption. Therefore, the positive covariance between the pricing kernel and the risky asset return decreases the equity premium.

In summary, this study sheds more light on the interrelationships arising between the options market and the macroeconomic and stock market environment. The first empirical chapter (Chapter 3) contributes to the literature that explores the determinants of the shape of the risk-neutral distribution by departing from the rational expectations paradigm. In particular, it elaborates on the previously documented impact of investor sentiment on index option prices (Han, 2008, Lemmon and Ni, 2011) by showing that in recent years the risk-neutral skewness and hence the pricing kernel is affected by the sentiment component driven by fundamentals and not by investors' unjustified optimism or pessimism. This result first indicates that incorporating investors' irrationality into option pricing models is not likely to improve the fit of observed option prices at least in mature markets such as the S&P 500 index options market. Second, given the central role played by the pricing kernel in determining the price of all assets according to the consumption-based model, the above result implies that the impact of investors' erroneous beliefs in other mature markets such as the US stock market is likely to be small as well. In support of this conjecture, Sibley et al. (2013) find that the component of the Baker and Wurgler (2006) sentiment index that is unrelated to fundamentals has very limited explanatory power for the cross-section of stock returns. Moreover, the finding that call and put options are oppositely affected by the fundamentals sentiment component, provides further evidence in favor of the hypothesis that call and put options markets are segmented.

The second empirical chapter (Chapter 4) contributes to the ongoing research that extracts option-related variables for forecasting purposes by creating measures of index forward skewness coefficients and exploring their predictive power for future macroeconomic conditions, asset prices as well as variables related to risk and uncertainty. The documented significant predictive power of the estimated forward skewness coefficients for the majority of the variables examined, gives further support to the idea that the time dimension of the implied volatility surface provides important information about the underlying asset dynamics and investors' risk aversion in addition to that provided by the moneyness dimension.

Finally, the third empirical chapter (Chapter 5) contributes to the literature that explores the information content of options' trading volume but also to the literature that investigates the impact of dispersion in expectations on asset returns. Therefore, we uncover a new dimension of predictability stemming from the trading activity in the options market that has not been explored before. Second, we propose a new dispersion in expectations measure which, compared to other proxies that stem from analysts' forecasts or portfolio holdings, exhibits several advantages. More specifically, it captures all the beliefs that are expressed in a highly liquid options market in the form of trading activity, refers directly to asset returns and not to alternative economic indicators such as corporate earnings, can be estimated even on a daily basis, and is designed to disentangle between different levels of both positive and negative expectations. Additionally, since the intraday data in options trading activity are publicly and freely available, our measure can be easily used by investors and regulators.

The remainder of this thesis is structured as follows. Chapter 2 reviews the literature on the various techniques for extracting nonparametrically the risk-neutral distribution and the respective moments from observed option prices, while Appendix A discusses the parametric techniques. Chapters 3-5 present the main empirical findings of the thesis. Each of these chapters is accompanied by an appendix (Appendices B-D) that provides additional results and where appropriate complementary discussions in support of the arguments presented in the main body of the thesis. Finally, Chapter 6 discusses some of the limitations of the thesis, proposes some avenues for future research and concludes.

Chapter 2

Risk-Neutral Distributions and their Moments

2.1 Introduction

Cox and Ross (1976) show that the price of any European option can be seen as its discounted expected payoff under the risk-neutral measure:

$$C(X) = e^{-rt} \int_X^{\infty} g(S_t)(S_t - X)\, dS_t \tag{2.1}$$

$$P(X) = e^{-rt} \int_{-\infty}^{X} g(S_t)(X - S_t)\, dS_t, \tag{2.2}$$

where $C(X)$ and $P(X)$ denote the price of a European call and put option respectively with strike price X and time to maturity t, r is the risk-free rate and $g(S_t)$ stands for the risk-neutral density - RND for short - function of the underlying asset price at time t.[1] Under the risk-neutral measure, the discounted expected payoff of the option is a martingale, thus the risk-neutral measure is also called a martingale measure.

The seminal models of Black and Scholes (1973), Merton (1973) and Black (1976) rely on the assumption that the risk-neutral distribution of the underlying asset price is lognormal. The empirical evidence across markets, however, shows that observed option prices cannot be reconciled with this hypothesis. In particular, option prices in equity (Shimko, 1993, Rubinstein, 1994), foreign exchange (Campa et al., 1998),

[1] We keep this notation consistent in the rest of this chapter. Moreover, all option pricing formulas discussed refer to European options unless otherwise stated.

commodities (Sherrick, Garcia and Tirupattur, 1996) and interest rate (Dutta and Babbel, 2005) markets imply a risk-neutral distribution which is more negatively skewed and more leptokurtic than the lognormal one.

Since the risk-neutral distribution is closely linked to the pricing kernel and the physical distribution, its shape encapsulates important forward-looking information about investors' expectations and risk preferences. As a result, a large number of studies have developed alternative methods for extracting RNDs and their respective moments. Such methods can be divided into two main groups: The *parametric* methods and the *nonparametric* ones. The advantage of the parametric methods is that there are only a few parameters that have to be estimated. On the other hand, since every parametric model has a specific structure, it is not always easy to account sufficiently for the observed data. In other words, there is always a probability that the model will be misspecified. In contrast to the parametric methods, the nonparametric ones are very flexible but sometimes they can be quite data-intensive and may lead to data overfitting.

Due to their high flexibility and their *model-free* nature nonparametric methods have emerged as the primary methods for extracting risk-neutral densities and moments. In the subsequent empirical analysis we will use nonparametric methods that extract directly the moments of the risk-neutral distribution. This is because it is computationally easier and faster to extract directly the moments. However, such methods share common characteristics with various methods that extract the whole risk-neutral distribution. Therefore, the aim of this chapter is to provide an overview of the main methods found in the literature for extracting nonparametrically both risk-neutral distributions and risk-neutral moments directly. For the sake of completeness, an overview of the main parametric methods for extracting RNDs can be found in Appendix A.

2.2 Extracting Risk-Neutral Densities

The techniques for extracting RNDs nonparametrically can be further divided into *kernel* methods, *maximum entropy* methods, *RND-fitting* methods and *implied volatility curve-fitting* methods. Most of the subsequently described methods rely on Breeden and Litzenberger's (1978) observation that the second derivative of the call price with respect to its exercise price gives the RND function:

$$g(S_t) = e^{rt} \frac{d^2 C\left(X\right)}{dX^2} \mid_{X=S_t} . \tag{2.3}$$

2.2.1 Kernel methods

Kernel methods are conceptually similar to a nonlinear regression, in the sense that they are used to fit a function given a set of observed data. Aït-Sahalia and Lo (1998) suggest the usage of a kernel estimator in order to obtain an option pricing function $\widehat{C}(.)$ that matches the observed option prices and use Breeden and Litzenberger's (1978) double differentiation rule to subsequently extract the RND. In particular, given a number of observed option prices $\{C_i\}$ and their characteristics $\{Z_i \equiv [S_{0_i}, X_i, t_i, r_{t;i}, d_{t;i}]'\}$ - where d_i is the dividend yield corresponding to call C_i and the rest of the letters are defined as before - they minimize the following mean squared error formula:

$$\min_{C(.)\in G} \sum_{i=1}^{n} \left[C_i - C\left(Z_i\right)\right]^2 , \tag{2.4}$$

where G is the space of twice continuously differentiable functions. The conditional expectation of C given the information set Z is estimated using a nonparametric kernel regression. For every specific value Z_{i_0}, this type of regression takes a weighted average of all the $C_i s$ by assigning higher weights to observations with characteristics Z_i that are closer to Z_{i_0}. The option prices are assumed to depend on five variables, therefore a five-dimensional kernel function $K(Z)$ that integrates to one is selected.

The Nadaraya-Watson kernel estimator has the generic form:

$$\widehat{C}(Z) = \widehat{E}\left[C \mid Z\right] = \frac{\sum_{i=1}^{n} K\left(\left(Z - Z_i\right)/h\right) C_i}{\sum_{i=1}^{n} K\left(\left(Z - Z_i\right)/h\right)}, \tag{2.5}$$

where h is the bandwidth. The higher the value of h, the smoother the function becomes, whereas the closer it is to zero the more peaked it becomes. Aït-Sahalia and Lo (1998) show that while the choice of kernel function does not play any crucial role to the final result, the choice of the correct bandwidth is of great importance. Furthermore, they find that more accurate estimates are obtained when the dimensions of the kernel function are reduced. Therefore, they also propose another semiparametric methodology where the pricing function is the Black-Scholes one but the volatility parameter depends on the futures price, the strike price and the time to maturity. This idea is depicted as following:

$$\widehat{C}(Z) = C_f(F_0, X, t, r_t; \widehat{\sigma}(F_0, X, t)). \tag{2.6}$$

where $C_f(.)$ stands for the Black-Scholes pricing formula and $F_0 = S_0 e^{(r_t - \delta_t)t}$. The conditional expectation of σ on F_0, X and t is calculated by the following three dimensional kernel estimator:

$$\widehat{\sigma}(F_0, X, t) = \frac{\sum_{i=1}^{n} k_F\left(\frac{F_0 - F_{0_i}}{h_F}\right) k_X\left(\frac{X - X_i}{h_X}\right) k_t\left(\frac{t - t_i}{h_t}\right) \sigma_i}{\sum_{i=1}^{n} k_F\left(\frac{F_0 - F_{0_i}}{h_F}\right) k_X\left(\frac{X - X_i}{h_X}\right) k_t\left(\frac{t - t_i}{h_t}\right)}. \tag{2.7}$$

Since the foregoing methodology uses both cross-sectional and time-series option prices, its main characteristic is that it is stable across time. On the other hand, there may be some dates where the estimated risk-neutral distribution is not consistent with the observed cross section of option prices.

Bondarenko (2003) incorporates the idea of a kernel function in his research but in a different way. He states that the RND function can be described by the convolution of a kernel function k and another positive function u. In general, the convolution of two integrable functions can be represented by:

$$\xi \cdot \eta = \int_{-\infty}^{\infty} \xi\left(\chi - \psi\right) \eta\left(\psi\right) d\psi. \tag{2.8}$$

Thus, in this case $k_{S_t} \cdot u\left(S_t\right) = g\left(S_t\right)$. If there are n available calls in the market each with a strike price $X_i, i = 1, ..., n$ and having in mind Breeden and Litzenberger's (1978) double differentiation rule, the RND can be calculated by solving numerically the minimization problem:

$$\min_g \sum_{i=1}^{n} \left(C_i - D^{-2}\widehat{g}\left(X_i\right)\right)^2, \tag{2.9}$$

where:

$$D^{-2}g\left(X\right) = \int_{-\infty}^{X} \left(\int_{-\infty}^{\varphi} g\left(\zeta\right) d\zeta\right) d\varphi, \tag{2.10}$$

is the second integral of $g\left(X\right)$. For computational issues, however, Bondarenko (2003) chooses to discretize the possible values that the underlying asset can take.

2.2.2 Maximum entropy methods

Given some constraints, the maximum entropy distribution is the one that maximizes the information one misses when the value of a random variable is unknown and therefore can be described as the least prejudiced. For a continuous distribution $p(x)$ the entropy formula that is maximised is:

$$\Im\left(g\right) = -\int_{0}^{\infty} g\left(S_t\right) \ln g\left(S_t\right) dS_t \tag{2.11}$$

This idea can be used in the context of the RNDs in order to obtain the least prejudiced RND given the observed option data.

More specifically, Buchen and Kelly (1996) maximize the above formula subject to the constraints:

$$\int_{0}^{\infty} g(S_t) d\left(S_t\right) = 1 \tag{2.12}$$

$$g(S_t) \geq 0 \tag{2.13}$$

$$E_0\left[\mathcal{F}_i\left(S_t\right)\right] = \mathcal{D}_i = \int_0^\infty g(S_t)\mathcal{F}_i\left(S_t\right)dS_t, \tag{2.14}$$

where $\mathcal{D}_i$ is the observed option price and $\mathcal{F}_i\left(S_t\right)$ is the discounted payoff of the option i, for $i = 1, ..., m$ where m is the number of observed option data. The final formula to be maximized takes the form:

$$\begin{aligned}\mathcal{L}\left(g\right) &= -\int_0^\infty g\left(S_t\right)\ln g\left(S_t\right)dS_t + \left(1+\lambda_0\right)\int_0^\infty g(S_t)dS_t + \\ &+ \sum_{i=1}^{m}\lambda_i\int_0^\infty g(S_t)\mathcal{F}_i\left(S_t\right)dS_t,\end{aligned} \tag{2.15}$$

where $\lambda_i, ...\lambda_m$ are the Langrange multipliers which are computed numerically. The resulting RND becomes:

$$g\left(S_t\right) = \frac{\exp\left(\sum\limits_{i=1}^{m}\lambda_i\mathcal{F}_i\left(S_t\right)\right)}{\int_0^\infty \exp\left(\sum\limits_{i=1}^{m}\lambda_i\mathcal{F}_i\left(S_t\right)\right)dS_t}. \tag{2.16}$$

It has to be mentioned that the above formulas are discretised by the authors and the results are assumed to approximate the respective continuous distributions.

Buchen and Kelly (1996) suggest also another similar methodology which is based on the "Principle of Minimum Cross-Entropy". More specifically, if there is some prior information about the probability distribution of the value of the underlying asset at time t, which can be described by a PDF $q\left(S_t\right)$ then the distribution $g\left(S_t\right)$ can be estimated by minimizing the entropy distance between the two distributions. This difference in the uncertainty implicit in each distribution is called "cross entropy". Hence given the same constraints and a prior distribution $q\left(S_t\right)$ the cross entropy formula that is minimized is:

$$\wp\left(g, q\right) = \int_0^\infty g\left(S_t\right)\ln\frac{g\left(S_t\right)}{q\left(S_t\right)}dS_t, \tag{2.17}$$

and the RND becomes:

$$g\left(S_t\right) = \frac{q\left(S_t\right)\exp\left(\sum\limits_{i=1}^{m}\lambda_i\mathcal{F}_i\left(S_t\right)\right)}{\int_0^\infty q\left(S_t\right)\exp\left(\sum\limits_{i=1}^{m}\lambda_i\mathcal{F}_i\left(S_t\right)\right)dS_t}. \tag{2.18}$$

This distribution is regarded again as the least prejudiced subject to the existing constraints. When there is no prior information about the distribution $g\left(S_t\right)$, $q\left(S_t\right)$ can be seen as a uniform distribution and in this case the resulting risk-neutral PDF will be the maximum entropy distribution. Thus, the two methodologies coincide when there is no prior information about the distribution of the random variable.

The minimum cross entropy methodology is also implemented by Stutzer (1996). Stutzer takes a large number of t-period past returns[2] and creates a uniform prior distribution with $h = 1, ...H$ possible outcomes. He uses the following constraint as an approximation of the martingale property that has to hold for the future returns of the underlying asset:

$$1 = \sum_{h=1}^{H}\frac{R\left(-h\right)}{e^{rt}}\frac{g\left(h\right)}{q\left(h\right)}q\left(h\right), \tag{2.19}$$

where $R\left(.\right)$ denotes the H past t-period returns that are used as a proxy for the future return of the asset. Since the author uses a uniform prior distribution the minimum cross entropy technique has no difference from the maximum entropy method and the risk-neutral density function becomes (in discrete form):

$$g\left(S_t\right) = \frac{\exp\left(\gamma^*\frac{R(-h)}{e^{rt}}\right)}{\sum_h\exp\left(\gamma^*\frac{R(-h)}{e^{rt}}\right)}, \tag{2.20}$$

where γ^* is a Lagrange multiplier found by solving numerically the following convex problem:

$$\gamma^* = \arg\min_\gamma\sum_h\exp\left[\gamma\left(\frac{R\left(-h\right)}{e^{rt}} - 1\right)\right]. \tag{2.21}$$

[2]Where t is equal to the maturity of the option he is interested in.

2.2.3 RND-fitting methods

This class of methods approximates the RND directly using appropriate optimization criteria. In Rubinstein (1994), for example, the risk-neutral distribution is estimated by applying the following least squares formula:

$$\min_{P_j} \sum_j \left(P_j - P_j^{'}\right)^2, \tag{2.22}$$

subject to:

$\sum_j P_j = 1$ and $P_j \geq 0$ for $j = 0, ..., n$,

$S^b \leq S \leq S^a$ where $S = \frac{\left(d^t \sum_j P_j S_j\right)}{R_f^t}$,

$C_i^b \leq C_i \leq C_i^a$ where $C_i = \frac{\left(\sum_j P_j \max[0, S_j - X_i]\right)}{R_f^t}$ for $i = 1, ..., m$,

where:

P_j denotes the nodal implied risk-neutral probabilities.

$P_j^{'}$ denotes the prior distribution which is derived from a n-step standard binomial tree and for a large enough n can be considered to be lognormal.

S_j is the final nodal asset price j.

$S^b\,(S^a)$ is the current observed bid (ask) price of the underlying asset.

$C_i^b\,(C_i^a)$ is the current observed bid (ask) call option price with strike price X_i.

d is the annualized payout return.

R_f is the annualized risk-free interest return.

t is the time to maturity of the option.

It is easy to observe that Rubinstein's (1994) methodology provides as the resulting RND the one closest to lognormal for which the present values of all the options and underlying asset fall between their bid and ask prices. Therefore, if all option prices calculated based on the lognormal distribution lie between the actual bid and ask prices then $P_j = P_j^{'}$ for every j. Moreover, as the available set of option prices increases, the resulting probability distribution will depend less on the prior distribution. In the extreme case that $m \to \infty$, P_j will become independent of $P_j^{'}$.

Based on Rubinstein's (1994) technique, Jackwerth and Rubinstein (1996) propose another approach where the objective is to find the implied distribution with the maximum smoothness. In this case there is no prior distribution. The objective function is:

$$\min_{P_j} \Omega = \sum_j \left(\frac{\partial^2 P_j}{\partial X_j^2} \right)^2 , \tag{2.23}$$

which becomes in discrete form:

$$\min_{C_j} \Omega = \sum_j \left(C_{j-2} - 4C_{j-1} + 6C_j - 4C_{j+1} + C_{j+2} \right)^2 , \tag{2.24}$$

subject to $C_j = C_i^m$ whenever $X_j = X_i$ for $j = 0, ...n$ and $i = 1, ..., m$, where C_j (C_i^m) is the option value (observed mid-point price) at exercise price X_j (observed strike price X_i). The first order condition for the cases where there is no available option with strike price X_j becomes in discrete form:

$$\begin{aligned} \frac{\partial \Omega}{\partial C_j} &= 2C_{j-4} - 16C_{j-3} + 56C_{j-2} - 112C_{j-1} + 140C_j - 112C_{j+1} \\ &\quad +56C_{j+2} - 16C_{j+3} + 2C_{j+4} \\ &= 0. \end{aligned} \tag{2.25}$$

However, whenever there is an option with strike price equal to X_j, the model option prices should coincide with the observed midpoint prices. Therefore, a penalty term is added to function Ω and the new objective function becomes:

$$\Omega' = \Omega + a \sum_i \left(C_i - C_i^m \right)^2 , \tag{2.26}$$

and taking the first order condition for the cases where $X_j = X_i$ becomes:

$$\begin{aligned} \frac{\partial \Omega'}{\partial C_j} &= 2C_{j-4} - 16C_{j-3} + 56C_{j-2} - 112C_{j-1} + (140 + 2a)\, C_j - 112C_{j+1} \\ &\quad +56C_{j+2} - 16C_{j+3} + 2C_{j+4} \\ &= 2aC_i^m. \end{aligned} \tag{2.27}$$

Jackwerth and Rubinstein (1996) solve the above system of equations by setting the prices of call options with very high strike prices equal to zero and the prices of call options with very low strike prices equal to their intrinsic values. This procedure also ensures that the resulting risk-neutral probabilities, calculated by the butterfly approximation of the Breeden and Litzenberger's (1978) formula, sum up to one. In case of negative risk-neutral probabilities, the authors increase the number of options with predetermined prices until they find an optimum RND.

2.2.4 Implied volatility curve-fitting methods

The purpose of this group of methodologies is to fit the implied volatility smile using some type of polynomial function. Then, they transform the fitted implied volatility smile into option prices and use the double differentiation rule in order to estimate the risk-neutral distribution either numerically or analytically. It has to be mentioned that the usage of the Black-Scholes model for transforming the option prices into implied volatilities and vice versa, does not imply the validity of the model's assumptions.

Shimko (1993) is the first to propose that instead of interpolating the observed option prices, it is preferable to interpolate the observed Black-Scholes implied volatilities. This way, the resulting option pricing function turns out to be smoother. Therefore, Shimko (1993) assumes that the implied volatility function for an expiration date t with respect to the strike price has a quadratic form:

$$\widehat{\sigma}(X) = a_0 + a_1 X + a_2 X^2. \tag{2.28}$$

Given the implied volatilities for all the call options available in the market, the coefficients a_0, a_1, a_2 can be estimated by linear least squares. Then, he creates a smoothed call option pricing formula which is a function of $\widehat{\sigma}(X)$ and by differentiating twice he finds the RND function. The analytic expression of the function is:

$$g(S_t = X) = n\left(d_2\right)\left[d_{2x} - \left(a_1 + 2a_2 X\right)\left(1 - d_2 d_{2x}\right) - 2a_2 X\right], \tag{2.29}$$

where $n(.)$ is the PDF for a standard normal distribution, d_1, d_2 are defined as in the case of the Black-Scholes formula, and:

$$\begin{aligned} d_{1x} &= -\frac{1}{X\sigma\sqrt{t}} + \left(\frac{1-d_1}{\sigma\sqrt{t}}\right)(a_1 + 2a_2 X) \\ d_{2x} &= d_{1x} - (a_1 + 2a_2 X). \end{aligned} \tag{2.30}$$

Since this method provides probabilities only for the range of prices corresponding to the option strike prices observed in the market, the author assumes that the tails of the distribution are lognormal.

Malz (1997) modifies the above technique by expressing the volatility smile in terms of $\Delta = \frac{\partial C}{\partial S}$ (delta of the option) and uses a quadratic polynomial to approximate the implied volatility function as well. Then, he transforms the volatility smile from a function of delta to a function of strike price. By substituting this implied volatility function to the Black-Scholes formula and differentiating numerically twice, he finds the RND function. The advantage of Malz's (1997) approach compared to that of Shimko (1993) is that Δ can take values only from 0 to 1 and these boundaries represent the whole probability distribution. Therefore, there is no need for any assumptions regarding the tails of the distribution. Another advantage of Malz's (1997) technique is that the resulting distribution is more flexible in the centre where in general the data are more reliable (Bliss and Panigirtzoglou 2002).

Campa et al. (1998) amend Shimko's (1993) method by fitting a cubic spline instead of a quadratic polynomial to the implied volatilities of the observed options. In this case, the cubic polynomial that links two knots can be different for every different couple of data points. One constraint is that at each point the first derivatives of the two polynomial functions should be equal and differentiable. Moreover, for the area before the first data point and after the last one, the first and last polynomials are used for a length that equals that of the first and last data interval. Beyond this extended range of available implied volatilities, the implied volatility smile is assumed to be flat.

Bliss and Panigirtzoglou (2002) combine the two previously described methods and use a smoothing cubic spline to fit a volatility function with respect to the Δ of the options.[3,4] The main characteristic of the smoothing cubic spline is that it penalizes excess curvature by a smoothness parameter λ. Moreover, outside the available data range the spline becomes linear. The objective function that is minimized is:

$$\min_{\Theta} \sum_{i=1}^{N} w_i \left(IV_i - \widehat{IV_i}\left(\Delta_i, \Theta\right) \right)^2 + \lambda \int_{-\infty}^{\infty} f''\left(x; \Theta\right)^2 dx, \tag{2.31}$$

where Θ is the matrix of the parameters of the cubic spline, $f\left(\Theta\right)$ is the implied volatility function, w_i is a weight that corresponds to the option's vega ($v = \frac{\partial C}{\partial \sigma}$) and $\widehat{IV_i}\left(\Delta_i, \Theta\right)$ is the fitted implied volatility at Δ_i given the parameters Θ. When the implied volatility function with respect to Δ is estimated, a large number of equally spaced points are selected and are transformed to option prices with respect to strike prices. Finally, by approximating numerically the second derivative of these option prices with respect to the respective strike prices, the risk-neutral density function is obtained.

Jackwerth (2000) suggests another methodology that fits the observed implied volatilites and minimizes the curvature of the volatility smile. The author first discretizes the possible future values of the underlying asset with equal intervals of δ and in a way that all the available strike prices in the market are covered. Then the following objective function is minimized:

$$\min_{\sigma_j} \left(i - p\right) \sum_{j=0}^{J} \left(\sigma_j'' \right)^2 + p \sum_{i=1}^{I} \left(\frac{\sigma_i - \overline{\sigma_i}}{STD_i} \right)^2, \tag{2.32}$$

where J is the number of possible outcomes of the stock price, I is the number of observed options in the market, σ_j is the implied volatility of every hypothetical option j, $\overline{\sigma_i}$ is the implied volatility of the observed option i, σ_i is the model implied

[3] According to Figlewski (2010), this methodology solves the problem of the discontinuous first derivative of the risk-neutral distribution which is present when interpolating a cubic spline to the volatility smile expressed in terms of strike prices.

[4] Similar smoothing spline techniques are also implemented by Aparicio and Hodges (1998) and Anagnou-Basioudis et al. (2005). However, in these studies the volatility smile is plotted against strike prices and not against option deltas.

volatility for the observed option i, $\sigma_j^{''}$ is the second derivative of implied volatility with respect to strike price approximated by $\sigma_j^{'} = (\sigma_{j+1} - 2\sigma_j + \sigma_{j-1})/\delta^2$, STD_i is the standard deviation of the implied volatility during the day for each option i and p is the trade-off parameter. The second part of the function minimizes the error between the observed and the model implied volatilities while the first part minimizes the curvature of the constructed volatility smile. As long as the implied volatility function is estimated, the respective Black-Scholes prices are calculated and the double differentiation rule provides the following formula for the RND:

$$g(S_j) = e^{rt}\begin{bmatrix} \frac{e^{-rt}n(d2_j)}{S_j\sigma_j\sqrt{t}}\left[1 + 2S_j\sqrt{t}d1_j\sigma_j^{'}\right] + \\ S_0 d^{-t}\sqrt{t}n(d1_j)\left[\sigma_j^{''} + \frac{d1_j d2_j}{\sigma_j}\left(\sigma_j^{'}\right)^2\right] \end{bmatrix}, \tag{2.33}$$

where:

$n(.)$ is the standard normal density function,

$d1_j = \left(\frac{\ln\left(\frac{S_0 d^{-t}}{S_j e^{-rt}}\right)}{\sigma_j\sqrt{t}}\right) + \frac{1}{2}\sigma_j\sqrt{t}$,

$d2_j = d1_j - \sigma_j\sqrt{t}$,

$d = 1+$ dividend yield,

S_j is the asset price at time t equal to the respective strike price X_j and

$\sigma_j^{'}$ = is the first derivative of implied volatility with respect to strike price, approximated by $\sigma_j^{'} = (\sigma_{j+1} - \sigma_{j-1})/2\delta$. A similar approach is also presented by Jackwerth (2004). However, in this case the optimization function is:

$$\min_{\sigma_j}\left(\frac{\delta^4}{2(J+1)}\right)\sum_{j=0}^{J}\left(\sigma_j^{''}\right)^2 + \frac{p}{2I}\sum_{i=1}^{I}(\sigma_i - \overline{\sigma_i})^2. \tag{2.34}$$

Figlewksi (2010) also interpolates the implied volatilities plotted as a function of strike prices but uses a different technique. He fits a $4th$ order spline assuming only one knot for the at-the-money option and minimizing the weighted sum of squared deviations between the curve and the market implied volatility midpoints. The constructed weighting function assigns higher weights to the deviations that fall outside the bid-ask spread. Moreover, the author chooses to discard from his dataset quotes of deep-out-of-the-money options and forms the RND tails according to the GEV distribution.

2.3 Extracting Risk-Neutral Moments

It is evident from the previous section that there exists a plethora of alternative methods for extracting RNDs, most of which differ substantially in terms of their objectives, assumptions and constraints to be satisfied. However, while all studies typically lead to RNDs that are skewed to the left and exhibit fat tails, there is no agreement in the literature about which method delivers the most accurate results. Furthermore, while it is possible to calculate the moments of an estimated RND, it is easier and faster to compute the risk-neutral moments directly from the observed option prices. Therefore, this section outlines the main techniques for estimating risk-neutral moments given a set of available option prices.

The common characteristic of all the methods discussed in this section is that the risk-neutral moments can be estimated from a portfolio of out-of-the-money European call and put options with weights that depend on the current underlying asset price and the respective strike prices. This is because they rely on the spanning result of Bakshi and Madan (2000) and Carr and Madan (2001) who show that any twice-continuously differentiable payoff function $H(S_t)$ of the terminal asset price S_t can be written as:

$$H(S_t) = H(S_0) + (S_t - S_0)H'(S_0) + \int_{S_0}^{\infty} H''(X)(S_t - X)^+ dX + \int_{\infty}^{S_0} H''(X)(X - S_t)^+ dX, \quad (2.35)$$

where S_0 is the current asset price, $(S_t - X)^+ = max(0, S_t - X)$ and $(X - S_t)^+ =$

$max(0, X - S_t)$. The above equation can be rewritten as:

$$H(S_t) = [H(S_0) - H'(S_0)S_0] + H'(S_0)S_t + \int_{S_0}^{\infty} H''(X)(S_t - X)^+ dX + \int_{\infty}^{S_0} H''(X)(X - S_t)^+ dX, \quad (2.36)$$

and implies that the payoff function $H(S_t)$ can be replicated by a $H(S_0) - H'(S_0)S_0$ positioning in zero-coupon bonds, a $H'(S_0)$ positioning in the asset and a $H''(X)dX$ positioning in out-of-the-money call and put options of all strikes. Intuitively, the positions in the bond and the asset form a tangent to the payoff curve at the initial asset price S_0, while the positions in the out-of-the-money options generate the necessary curvature to match the payoff curve for the rest of the potential terminal prices.

The main issue regarding the implementation of these methods is the requirement of a continuum of out-of-the-money calls and puts across strike prices, while in reality options are traded only for a finite range of discrete strike prices. To overcome this difficulty, researchers typically fit the implied volatility curve in a way similar to what described in the previous section. More specifically, they first fit the implied volatility curve using cubic splines inside the range of available data and the respective implied volatility boundary values outside the range of available data, then they create a large number of artificial option prices and finally they estimate the integrals that appear in the formulas using the trapezoidal approximation (see for example Jiang and Tian, 2005, Chang et al., 2013 and Neumann and Skiadopoulos, 2013).

Carr and Madan (1998), Demeterfi et al. (1999) and Britten-Jones and Neuberger (2000) show that under the assumption that the underlying asset price follows a diffusion process, it is possible to estimate exactly the integrated variance of the asset's returns under the risk-neutral measure over a period $[0, t]$ using a portfolio of out-of-the-money call and put options expiring at time t:

$$IV(0,t) = 2e^{rt}\left[\int_{F_0}^{\infty}\frac{C(X)}{X^2}dX + \int_{0}^{F_0}\frac{P(X)}{X^2}dX\right], \tag{2.37}$$

where $C(X)$ and put $P(X)$ denote call and put option prices respectively with time to maturity t and strike price X, and F_0 stands for the asset's current forward price. The above formula is used by the Chicago Board Options Exchange (CBOE) for the estimation of the well-known VIX index. Jiang and Tian (2005) and Carr and Wu (2009) further assert that equation (2.37) can estimate the quadratic variation over a period $[0,t]$ with a small approximation error in case the underlying asset price follows a discontinuous process with jumps. However, Broadie and Jain (2008), Du and Kapadia (2012), Rompolis and Tzavalis (2013) and Bondarenko (2014) show that in the presence of negative jumps, this method consistently underestimates the quadratic variation of log-returns and this underestimation is not negligible in turbulent periods.

In light of this, Neuberger (2012) and Bondarenko (2014) propose an alternative definition of variance. In particular, they suggest the function $\varrho^* = 2\left(e^x - 1 - x\right)$, to be used as variance instead of the commonly used $\varrho = (x)^2$, where x is the log-return of a martingale price. When the are no jumps in the underlying asset process both variance definitions, ϱ^* and ϱ, converge to the integrated variance at the continuous-time limit and can be replicated exactly by equation (2.37). In the presence of jumps, however, the two methods respond differently and only the quadratic variation stemming from ϱ^* can be replicated exactly by equation (2.37). Moreover, ϱ^* when sampled at a high frequency serves as an unbiased estimator of the true conditional long-horizon variance, while this is not the case for ϱ.

Similarly, Neuberger (2012) and Kozhan, Neuberger and Schneider (2013) propose the function $\psi^* = 6\left(xe^x - 2e^x + x + 2\right)$, to be used as the third moment of an asset's log-return x, instead of the commonly used $\psi = (x)^3$. The reason is that, unlike the traditional definition ψ, the alternative definition ψ^* when sampled at a high frequency serves as an unbiased estimator of the true conditional long-horizon third moment. Based on their alternative specification, Neuberger (2012) and Kozhan, Neuberger and Schneider (2013) show that it is possible to estimate the third moment of an asset's returns under the risk-neutral measure over a period $[0, t]$, and this estimate will be exact even in the presence of jumps in the underlying asset price process, as long as the price is martingale. The respective formula takes the form:

$$TM(0,t) = 6e^{rt}\left[\int_{F_0}^{\infty} \frac{X - F_0}{X^2 F_0} C(X)dX - \int_0^{F_0} \frac{F_0 - X}{X^2 F_0} P(X)dX\right], \tag{2.38}$$

and by standardizing with the implied quadratic variation from equation (2.37), the respective risk-neutral skewness coefficient becomes:

$$SC(0,t) = \frac{TM(0,t)}{IV(0,t)^{\frac{3}{2}}}. \tag{2.39}$$

In a slightly different context, Bakshi, Panayotov and Skoulakis (2011) assume that the underlying asset price follows a pure diffusion process and based on the theoretical evidence presented in Carr and Lee (2009) construct exponential claims on integrated variance of an asset's returns under the risk-neutral measure for a period $[0, t]$:

$$H(0,t) \equiv e^{-rt}\mathbb{E}^{\mathbb{Q}}\left[e^{-IV(0,t)}\right], \tag{2.40}$$

The price of such an exponential claim is given by the following formula:

$$\begin{aligned} H(0,t) = e^{-rt} &- \int_{S_0}^{\infty} \frac{\frac{8}{\sqrt{14}}\cos\left(\arctan\left(\frac{1}{\sqrt{7}}\right) + \frac{\sqrt{7}}{2}\ln\left(\frac{X}{S_0}\right)\right)}{\sqrt{S_0}X^{\frac{3}{2}}} C(X)dX \\ &- \int_0^{S_0} \frac{\frac{8}{\sqrt{14}}\cos\left(\arctan\left(\frac{1}{\sqrt{7}}\right) + \frac{\sqrt{7}}{2}\ln\left(\frac{X}{S_0}\right)\right)}{\sqrt{S_0}X^{\frac{3}{2}}} P(X)dX. \end{aligned} \tag{2.41}$$

Since their final goal is to provide forward variance estimates, Bakshi, Panayotov and Skoulakis (2011) perform a small manipulation of equation (2.40) and get:

$$-\ln H(0,t) = rt - \ln \mathbb{E}^{\mathbb{Q}}\left[e^{-IV(0,t)}\right]. \tag{2.42}$$

Therefore, their estimates are affected not only by the expectation of the integrated variance under the risk-neutral measure but also by the level of the risk-free rate. The effect of the presence of jumps in the underlying asset price process on the quadratic variation estimates from this method has not been investigated so far.

Bakshi, Kapadia and Madan (2003) construct formulas for all the risk-neutral moments of the continuously compounded log-return $\ln\left(\frac{S_t}{S_0}\right)$. In particular they show that the risk-neutral variance, skewness and kurtosis of the asset's return over a period $[0,t]$ can be calculated from the prices of out-of-the-money call and put options expiring at time t:

$$Var(0,t) = e^{rt}V(0,t) - \mu(0,t)^2, \tag{2.43}$$

$$Skew(0,t) = \frac{e^{rt}W(0,t) - 3\mu(0,t)\, e^{rt}V(0,t) + 2\mu(0,t)^3}{\left[e^{rt}V(0,t) - \mu(0,t)^2\right]^{3/2}}, \tag{2.44}$$

$$Kurt(0,t) = \frac{e^{rt}Q(0,t) - 4\mu(0,t)\, e^{rt}W(0,t) + 6e^{rt}\mu(0,t)\, V(0,t) - 3\mu(0,t)^4}{\left[e^{rt}V(0,t) - \mu(0,t)^2\right]^2}, \tag{2.45}$$

where:

$$\mu(0,t) = e^{rt} - 1 - \frac{e^{rt}}{2}V(0,t) - \frac{e^{rt}}{6}W(0,t) - \frac{e^{rt}}{24}Q(0,t), \tag{2.46}$$

and $V(0,t)$, $W(0,t)$ and $Q(0,t)$ are the prices of three contracts that represent the second, third and fourth noncentral moment respectively of the asset's log-return:

$$V(0,t) = \int_{S_0}^{\infty} \frac{2(1-\ln(\frac{X}{S_0}))}{X^2} C(X)dX + \int_0^{S_0} \frac{2(1+\ln(\frac{S_0}{X}))}{X^2} P(X)dX, \tag{2.47}$$

$$
\begin{aligned}
W(0,t) &= \int_{S_0}^{\infty} \frac{6\ln(\frac{X}{S_0}) - 3\left[\ln(\frac{X}{S_0})\right]^2}{X^2} C(X)dX - \\
&\quad - \int_0^{S_0} \frac{6\ln(\frac{S_0}{X}) + 3\left[\ln(\frac{S_0}{X})\right]^2}{X^2} P(X)dX, \qquad (2.48)
\end{aligned}
$$

$$
\begin{aligned}
Q(0,t) &= \int_{S_0}^{\infty} \frac{12\left[\ln(\frac{X}{S_0})\right]^2 - 4\left[\ln(\frac{X}{S_0})\right]^3}{X^2} C(X)dX + \\
&\quad + \int_0^{S_0} \frac{12\left[\ln(\frac{S_0}{X})\right]^2 + 4\left[\ln(\frac{S_0}{X})\right]^3}{X^2} P(X)dX. \qquad (2.49)
\end{aligned}
$$

Rompolis and Tzavalis (2013) extend the work of Bakshi, Kapadia and Madan (2003) for moments of order higher than fourth. Unlike the risk-neutral moments estimated from the previous techniques, the moments estimated from Bakshi, Kapadia and Madan's (2003) technique do not have theoretically equivalent realized measures that can be constructed from high-frequency returns. However, Du and Kapadia (2012) show that in the presence of jumps in the underlying asset price process, the Bakshi, Kapadia and Madan (2003) risk-neutral variance captures much more accurately the quadratic variation of the asset's returns than the formula (2.37) does.

2.4 Conclusion

In this chapter, we outline the main methods for extracting risk-neutral distributions and moments nonparametrically from observed option prices. Such methods will be subsequently used for the empirical analysis presented in this thesis. We first describe the different techniques for extracting the whole distribution of the future underlying asset price (or return). Such techniques can be further divided into those that use a kernel estimator to estimate the option pricing function, those that obtain the least prejudiced distribution, i.e. the one exhibiting the maximum entropy, those that fit directly the risk-neutral density and finally those that fit the implied volatility curve.

Assuming that the risk-neutral distribution is accurately estimated, it is straightforward to calculate its moments. However, given the computational difficulties that are embedded in most of the aforementioned methods and the fact that there is no consensus in the literature regarding the method that provides the most accurate results, it is easier and more reliable to estimate the risk-neutral moments directly from the observed option prices. Therefore, in the second part of this chapter we analyze the different approaches for estimating the risk-neutral variance, skewness and kurtosis using option prices of a given maturity.

Chapter 3

Investor Sentiments, Rational Beliefs and Option Prices

3.1 Introduction

The smirk pattern, which characterizes the cross-sectional plot of index options' implied volatilities, constitutes evidence of a pronounced negative skewness in the risk-neutral distribution of the index returns. This phenomenon cannot be fully captured even by sophisticated option pricing models that incorporate stochastic volatility and jumps (see Bakshi et al., 1997, Bates, 2000 and Pan, 2002 among others). Under the representative-investor paradigm, consumption-based asset pricing theory accounts for the variation in risk-neutral skewness by describing the determinants of the pricing kernel. In particular, it suggests that when investors are pessimistic (optimistic) about future consumption their marginal utility is high (low). Consequently, when investors are bearish (bullish) about the market, they drive up (down) the prices of Arrow-Debreu securities that pay off when the index level is low. This is equivalent to a more (less) negatively sloped pricing kernel, and assuming that the conditional physical probability distribution is always approximately symmetric (Aït-Sahalia et al., 2001 and Rosenberg and Engle, 2002), it implies a more (less) negative risk-neutral skewness. Shefrin (2008) postulates that the pricing kernel can be decomposed into two parts: one that is driven by investors' erroneous beliefs and a second part that is driven by investors' rational expectations about future consumption.[1] Recent literature attributes the negative risk-neutral skewness to factors such as investor sentiment (Han, 2008, Lemmon and Ni, 2011), limits to arbitrage (Bollen and Whaley, 2004), market momentum (Amin et al., 2004), heterogeneous beliefs (Buraschi and Jiltsov, 2006) and market default risk (Andreou, 2013).

In this chapter, we examine how investor sentiment related to economic fundamentals and investors' erroneous beliefs impact the risk-neutral skewness of index returns. Unlike Han (2008) who assumes that investor sentiment reflects only investors' unjustified expectations,[2] a recent stream of papers acknowledges that investor sentiment is not purely driven by irrationality but also incorporates rational updating of beliefs. For example, Brown and Cliff (2005, p. 417) note that "when people say they are bullish on the market, this can be a rational reflection of prosperous times to come, an irrational hope for the future, or some combination of the two".[3] In light of this, we define investor sentiment as investors' overall attitude towards future market returns and argue that it captures their beliefs driven both by changes in fundamentals and by factors unrelated to fundamentals. Therefore, for the subsequent analysis we decompose aggregate investor sentiment into two components: an economic fundamentals (EF) component, which corresponds to investors' rational updating of beliefs regarding future market returns due to changes in economic conditions,[4] and an error in beliefs (EB) component, which captures investors' expectations that are not associated with the economic conditions (expressed in the form of unjustified optimism or pessimism). In this respect, this study contributes to the ongoing research on the impact of behavioral biases on index option prices and the pricing kernel.

In order to extract risk-neutral skewness estimates from S&P 500 index options

[1]It is important to note that Shefrin (2008) uses the word sentiment to describe only the error in beliefs sentiment component.

[2]His assumption is reinforced by the fact that the main results of his study do not change after controlling for four popular macroeconomic indicators.

[3]For similar discussions, see also Qiu and Welch (2004), Baker and Wurgler (2006), Lemmon and Portniaguina (2006) and Sibley et al. (2013).

[4]The term *economic conditions* is used to describe the combination of both macroeconomic and financial conditions.

we use the well-established model-free approach of Bakshi, Kapadia and Madan (2003). To capture investor sentiment we utilize three proxies which can be regarded as reflecting the sentiment of three different classes of investors. More specifically, we use the net position of non-commercial traders in S&P 500 futures to capture the sentiment of large speculators, the bull-bear spread based on the surveys of Investors Intelligence to proxy for the sentiment of investment advisors and a newly established measure introduced by Ben-Rephael et al. (2012) that captures individual investor sentiment through the net exchanges of the equity funds.

We estimate the economic fundamentals sentiment component by regressing each sentiment proxy on a vector of eight major economic indicators that practically cover all aspects of the macroeconomic and financial environment that could influence investors' expectations regarding future market returns. The fitted values from the regressions are regarded as the component that is related to economic activity and corresponds to investors' rational updating of beliefs. The residuals from the regressions are regarded as the error in beliefs component that is not associated with the economic conditions and reflects investors' unjustified optimism or pessimism. We check the robustness of our sentiment decomposition by using a set of common latent factors that summarize the information embedded in a large panel of 131 macroeconomic variables. In particular, our alternative sentiment decompositions are based a) on the estimated common factors and b) on the variables that are most correlated with those factors. Both alternative sentiment decompositions provide qualitatively similar results with those from the main decomposition.

Our sample period extends from 1990:01 to 2011:06. We conduct our empirical analysis over two sample periods, before and after 1997:06. This enables us to compare our findings to those of Han (2008) whose sample period ends at 1997:06. Therefore, we examine whether the results reported in Han (2008) continue to hold in recent times. Furthermore, our subsample analysis is motivated by Gârleanu et al. (2009) who suggest the high likelihood of a structural change in the S&P 500 index options market in 1997 due to the introduction of new competing securities such as the S&P 500 E-mini futures and futures options on the Chicago Mercantile Exchange (CME) and the Dow Jones options on the Chicago Board Options Exchange (CBOE). The introduction of these alternative securities may have triggered a shift of investors between markets, thus changing the characteristics of the representative S&P 500 index options investor.

The empirical results show that in the first sample period (before 1997:06), there exists a significantly positive relationship between all three sentiment proxies and skewness. This means that a more bullish (bearish) investor sentiment leads to a less (more) negative risk-neutral skewness. Hence, our results corroborate the findings of Han (2008). However after 1997:06, we see that this pattern changes and none of the three sentiment proxies exhibits any significant relationship with skewness. Our findings indicate that the impact of aggregate sentiment on the risk-neutral skewness is insignificant in the recent years.

To shed light on the nature of this change, we repeat our analysis for the EF and EB components of sentiment. The results of the first period show that the previously reported relationship between aggregate investor sentiment and risk-neutral skewness stems mainly from the EF component, but the EB component also exhibits some explanatory power. However, the results for the second period provide a striking contrast. The EF component remains strongly significant, while the respective EB component is consistently insignificant. The results suggest that aggregate sentiment, which is previously found to be insignificant, cannot always account for the impact of investors' rational and erroneous expectations on option prices as it constitutes only a noisy aggregation of these two components. A further examination of the EF component of sentiment shows that the risk-neutral skewness is influenced by the expectations of large speculators regarding a reversal of recent economic conditions.

We further examine the market conditions under which the documented in the second period relation between the EF sentiment component and the risk-neutral skewness is more pronounced. Our analysis suggests that this significant relation mainly stems from periods of worsened stock market conditions. Thus, on average the S&P 500 index options traders react to bad news by forcing risk-neutral skewness to become less negative but do not react to good news in a way that would make risk-neutral skewness more negative. We validate this finding by examining whether the significant relation for the EF component of each sentiment proxy is mainly driven by periods of more bullish or bearish expectations depending on the way each proxy responds to recent economic conditions.

Finally, motivated by Constantinides et al.'s (2011) assertion that the index options market is segmented, in the sense that out-of-the-money puts are mainly traded by hedgers while out-of-the-money calls are mainly traded by speculators, we create different slope measures for calls and puts. The analysis (restricted in the second period) suggests that the slope of the calls' implied volatility smirk is mainly driven by the expectations of investment advisors and individual investors regarding a continuation of recent economic conditions, while the slope of the puts' implied volatility smirk is mainly driven by the expectations of large speculators regarding a reversal in the economy. In contrast, the EB component of all sentiment proxies has no explanatory power either for the slope measures from call options or for the slope measure from put options. The above evidence is consistent with that of Constantinides et al. (2011) and implies that the demand for call and the demand for put options originate from different sources.

Our findings provide useful insights regarding the role of investor sentiment on asset prices. In particular, we document that investor sentiment has an important component that represents investors' rational updating of beliefs, the impact of which cannot be ignored on asset prices. In fact, in the S&P 500 index options market, the pricing kernel is mainly driven by its economic fundamentals component and not by the errors in beliefs component. This finding is important as the index options market provides us with valuable forward-looking information about the pricing kernel. Moreover, in contrast to Han (2008), our results suggest that incorporating investors' irrationality into sophisticated option pricing models is no longer valuable.[5]

The remainder of the chapter is structured as follows. Section 3.2 presents in detail the related literature. Section 3.3 describes the data and the construction of the variables used in the study. Section 3.4 provides the empirical results. Finally, Section 3.5 concludes.

3.2 Related Literature

This chapter is mainly related to the literature that examines the determinants of the risk-neutral moments - or equivalently the shape of the implied volatility smirk - extracted from index and individual stock options. Bollen and Whaley (2004) investigate whether the net buying pressure for index and stock options affects the shape of the implied volatility smirk. The underlying assumption is that limits to arbitrage force market makers to charge a higher price for options when their short positions become large. Their results indicate that the shape of the index options' implied volatility smirk is driven by the demand for index put options consistent with the hypothesis that investors seek portfolio insurance. On the other hand, stock options' implied volatility smirk is affected by the demand for stock call options. Gârleanu et al. (2009) provide further time-series and cross-sectional evidence showing that option expensiveness and hence the level and steepness of the implied volatility smirk is positively linked to demand pressure.

Dennis and Mayhew (2002) investigate the systematic and firm-specific factors that affect the risk neutral skewness implied by individual stock options. They find that a larger (smaller) firm size, a higher (lower) stock beta, a lower (higher) stock liquidity, a period of high (low) market volatility and a period of more (less) negative market skewness are related to a more (less) negative individual stock risk-

[5] In fact, recent literature on general equilibrium and option pricing models attempts to account for the implied volatility smirk anomaly by extending the traditional rational representative-agent paradigm. For example, Liu et al. (2005) incorporate uncertainty aversion for rare events, while Benzoni et al. (2011) and Du (2011) build a long-run risk and a habit formation model respectively, with a jump component to the consumption growth process. Finally, Christoffersen et al. (2013) construct a GARCH option pricing model with a pricing kernel that depends both on stock returns and volatility.

neutral skewness. Similar results are also obtained by Taylor et al. (2009). Duan and Wei (2009) demonstrate that the shape of the individual stock options' implied volatility smirk is associated with the systematic risk proportion of the underlying asset. In particular, systematic risk proportion has a positive effect on the level and the steepness of the smirk. Furthermore, Bradshaw et al., (2010) find a positive relation between firm opacity and the steepness of the implied volatility smirk.

Han (2008) and Lemmon and Ni (2011) examine the impact of investor sentiment on S&P 500 index and individual stock option prices. Both studies find that a more bullish investor sentiment is related to a less negatively skewed risk-neutral density and a flatter implied volatility smirk, while a more bearish sentiment is associated with a more negative risk-neutral skewness and a steeper implied volatility smirk. Buraschi and Jiltsov (2006) create a heterogeneous-belief equilibrium model which leads to negative risk neutral skewness. Moreover, they provide empirical evidence that a higher heterogeneity of beliefs about expected returns increases the options' trading volume and makes the slope of the S&P 500 implied volatility smirk steeper. Friesen et al. (2012) draw the same conclusion by examining the implied volatility smirk of individual stock options.

Restricting our attention to the index options markets, Amin et al. (2004) examine the impact of stock market momentum on the S&P 100 index option prices and find that negative market returns lead to a substantially steeper and more curved smirk pattern in both call and put options' implied volatility functions. David and Veronesi (2014), show that the slope of the S&P 500 index implied volatility smirk is positively related to investors' perceived economic uncertainty, probability of a recession and probability of deflation. Andreou (2013) investigates the impact of market default risk on the S&P 500 index risk-neutral moments and shows that it is positively related to variance and skewness and negatively related to kurtosis.

With respect to the determinants of index option prices outside the US, Nordén and Xu (2012) indicate that an increase (decrease) in relative liquidity between out-of-the-money put and at-the-money call options leads to a less (more) negatively sloped implied volatility smirk of the Swedish OMXS30 index. Peña et al. (1999) show that the shape of the Spanish IBEX-35 index implied volatility smirk is mainly influenced by the market's transaction costs, proxied by the options' bid-ask spread, with higher (lower) transaction costs being associated with a more (less) curved smirk. In a similar study, Chang et al. (2009) indicate that the slope of the implied volatility smirk of the Hang Seng index is negatively related to the variance of the underlying asset and the options' bid-ask spread and also exhibits a "Monday effect", by being steeper on Mondays.[6]

This chapter is also related to a strand of the literature that examines the impact of macroeconomic announcements and macroeconomy in general on the equity market risk-neutral moments. Graham et al. (2003) investigate the impact of eleven macroeconomic announcements on the equity market risk-neutral variance as captured by VIX and state that five of them have a significantly negative effect. In a similar study, Nikkinen and Sahlström (2004) find that VIX decreases after the Federal Open Market Committee (FOMC) meetings and after the release of the employment report from the Bureau of Labor Statistics. Vähämaa and Äijö (2011) further document that the effect of the FOMC meetings on VIX is stronger during periods of expansive monetary policy. Beber and Brandt (2009) provide evidence that the reduction in the risk-neutral variance of cyclical stocks after a non-farm payroll announcement is more pronounced in periods of high macroeconomic uncertainty. At an intraday level, Nofsinger and Prucyk (2003) find that the implied volatility of S&P 100 index options increases for some hours after the release of macroeconomic news, a phenomenon which is mostly attributed to bad news. In the same vein, Bailey et al. (2014) show that during the recent financial crisis, VIX is significantly increased five minutes before and after macroeconomic announcements.

Steeley (2004) examines the impact of news about inflation, unemployment, government borrowings, interest rates and money supply in UK on the risk-neutral moments of the FTSE 100 index. In most cases the effect is significant especially for

[6] For the determinants of the risk-neutral moments in the interest rate and commodities markets, the reader can refer to Deuskar et al. (2008), Pan (2012) and Ruf (2012).

skewness and kurtosis. In a similar study, Äijö (2008) investigates the impact of both US and UK macroeconomic news on the shape of the risk-neutral distribution implied by FTSE 100 index options. His conclusion is that good (bad) news have a negative (positive) impact on implied volatility but a positive (negative) impact on skewness and kurtosis. At an intraday level, Kim and Lee (2011) examine whether six macroeconomic news announcements from South Korea and US influence the risk-neutral moments of KOSPI 200 index. They find that after an announcement risk-neutral volatility increases, kurtosis decreases while skewness increases after good news and decreases after bad news. In contrast to Äijö (2008) they state that changes in kurtosis do not depend on the quality of the news.[7]

Bekaert et al. (2013) investigate the effect of monetary policy actions on VIX. They show that a lower (higher) real interest rate is related to a lower (higher) VIX, an effect which mainly stems from the variance risk premium component. Mixon (2002) explores the effect of several stock market and macroeconomic variables on the at-the-money implied volatility of S&P 500 index options. S&P 500 index and Nikkei index past returns exhibit the highest explanatory power, having a negative relation with implied volatility. Secondarily, the short-term risk-free rate and the corporate yield spread have a significantly negative and positive effect respectively but only for a one-month horizon. Glatzer and Scheicher (2005) extract the DAX index RND and investigate whether macroeconomic and financial conditions in Germany and US influence the higher moments and the left tail of the distribution. Their results show that only the USD/Euro exchange rate has a significant impact on all the higher moments, whereas the US stock market momentum and variance mainly affect the implied volatility and to a lesser extent the implied skewness. The effect of German macroeconomy on the DAX RND shape is weak and limited only to the left tail of the distribution.

[7]For studies examining the impact of macroeconomic news on the risk-neutral variance and higher moments in the interest rate markets the reader can refer to Ederington and Lee (1996), Sun and Sutcliffe (2003), Vähämaa et al. (2005) and Beber and Brandt (2006). Kim and Kim's (2003) study focuses on the foreigh exchange markets.

3.3 Data and Variables

3.3.1 Options data and risk-neutral skewness

We obtain daily S&P 500 index call and put options data from IVolatility.com for the period 1990:01 to 2011:06.[8] Following the standard practice, option prices are calculated as the midpoint between the best bid and best ask price. Expiration time is calculated assuming 360 calendar days per year. Each trading day is matched with the respective dividend yield which is obtained from Bloomberg. Moreover, each option contract is matched with the appropriate continuous risk-free rate that is found after interpolating the one-, three-, six- and twelve-month Treasury Constant Maturity rates downloaded from the FRED database of the Federal Reserve Bank of St. Louis.

Standard filtering rules are applied to the dataset to eliminate measurement errors and outliers that are mainly caused by thinly traded options (Aït-Sahalia and Lo, 1998, Han, 2008, Neumann and Skiadopoulos, 2013). First, we discard options that violate no-arbitrage boundaries. Second, we exclude observations with zero bid prices and midpoint prices of less than 0.25 index point. Options with implied volatility of more than 100% are also removed. Finally, we take into consideration only options with a non-zero trading volume and maturity from 5 to 270 calendar days.

Risk-neutral skewness is estimated using the model-free method of Bakshi, Kapadia and Madan (2003), and specifically the formulas (2.44) and (2.46)-(2.49) presented in the previous chapter. This method has been extensively used in the literature (Dennis and Mayhew, 2002, Han, 2008, and Duan and Wei, 2009, among others) and is considered the standard approach for risk-neutral moments estimation. In order to create a monthly time-series we estimate skewness on the last trading day of each month. We consider only cross-sections that have at least two calls with $K/S > 1$ and two puts with $K/S < 1$, where K is the strike price of the

[8] Unlike the following chapters, the sample period examined in this chapter ends at 2011:06 due to the availability of the investor sentiment data.

option and S is the index level.

The main issue regarding the implementation of the method is that it requires a continuum of option prices while the available data is discrete. Therefore, following Chang et al. (2013) and Neumann and Skiadopoulos (2013), for each cross-section of options we interpolate implied volatilities into the range of available options data using a smoothing cubic spline and extrapolate outside this range using the respective boundary values. Our final goal is to obtain a set of 1000 implied volatilities that cover the moneyness range from 0.0001 to 3. The implied volatility data points for moneyness < 1 are then converted into put prices and those for moneyness > 1 are converted into call prices. Finally, the trapezoidal approximation is used to calculate the integrals in the Bakshi, Kapadia and Madan (2003) formulas. Following this procedure, we calculate the risk-neutral skewness for the two time horizons that are nearest to one month, and then linearly interpolate to find the risk-neutral skewness for exactly one-month ahead.

Figure 3.1 (top left panel) plots the monthly time series of S&P 500 index risk-neutral skewness from 1990:01 to 2011:06. We observe that the risk-neutral skewness is negative throughout the sample period and fluctuates substantially from month to month. It is notable that the level of the risk-neutral skewness increases during the period of the recent financial crisis, a phenomenon that has also been discussed by Birru and Figlewski (2011) and Coakley et al. (2013). As reported in Table 3.1, its sample mean is -1.559 and its autocorrelation coefficient 0.547.

3.3.2 Sentiment measures

The first sentiment proxy is related to the trading activity of large speculators in S&P 500 futures. The Commodity Futures Trading Commission (CFTC) requires clearing members, futures commission merchants and foreign brokers to report daily their futures and options positions if they are above a specified level. Based on those data, the CFTC releases the Commitments of Traders report, which provides a breakdown of each Tuesday's open interest for markets in which there are 20 or more reported positions. Since October 1992 the report is released on Friday and the data included refer to the previous Tuesday. Prior to that, it was released twice a month in the middle and at the end of each month. Every individual trader is classified by the CFTC as either "commercial" or "non-commercial". Commercial traders use futures for hedging purposes while non-commercial traders are large speculators. Similar to Han (2008), we derive the sentiment of large speculators as the net position (long contracts minus short contracts) of non-commercial traders scaled by the total open interest in S&P 500 futures (Spec-Sent).

The second sentiment proxy comes from Investors Intelligence's advisors' sentiment index. In particular, Investors Intelligence performs a weekly survey of more than 120 independent financial market newsletter writers. Each newsletter is categorized as bullish, bearish or correction, based on the expectations of future market movements. The survey started as monthly in 1963, became fortnightly until June 1969 and since then has been weekly. It is published every Wednesday but the historical data are matched with Friday dates since the majority of the newsletters are written after the markets close each Friday. Following Brown and Cliff (2004, 2005) we use the bull-bear spread (percentage of bullish investors minus percentage of bearish investors) in order to capture the sentiment of investment advisors (Adv-Sent).

The monthly time-series of the aforementioned sentiment proxies are created by using the data closest to the end of each month. The two sentiment proxies are shown by Han (2008) to be positively and strongly related to the S&P 500 index risk-neutral skewness for the period 1988:01-1997:06.

The third sentiment measure considered in this study is the normalized aggregate net exchanges of the equity funds introduced by Ben-Rephael et al. (2012). In particular, the Investor Company Institute (ICI) provides monthly data of aggregate mutual fund flows for bond funds, domestic equity funds, international equity funds and mixed funds. The last three categories constitute the overall equity funds category.[9] The sentiment proxy is calculated as the "exchanges in" minus the "ex-

[9]Ben-Rephael et al. (2012) include mixed funds in the equity funds category, because their

changes out" of the equity funds, normalized by the fund assets at the beginning of each month (Ind-Sent). Intuitively, higher sentiment leads investors to alter their asset allocation from bonds to equities and vice versa. Ben-Rephael et al. (2012) mention that the vast majority (more than 85%) of mutual fund assets are held by households, hence this measure is regarded as an individual investor sentiment proxy.

Summary statistics for the sentiment measures can be found in Table 3.1. We observe that the Adv-Sent is on average bullish with a mean value of 0.139, while the Spec-Sent is on average slightly bearish with a mean value of -0.045. Moreover, both the Adv-Sent and the Spec-Sent are quite persistent with autocorrelation coefficients of 0.746 and 0.817 respectively. In contrast, the Ind-Sent has an almost zero mean value and is much less persistent, with an autocorrelation coefficient of 0.219. Figure 3.1 plots the three proxies from 1990:01 to 2011:06. We see that Adv-Sent and Spec-Sent tend to move together in the first period but this pattern reverses in the second period. The Ind-Sent tends to follow the Adv-Sent across the whole period and especially during the latest years. This is expected since investment advisory services are mainly used by individual investors in order to form their beliefs. Comparing the plots of the sentiment proxies with that of the risk-neutral skewness, we observe that all sentiment proxies move similarly to the risk-neutral skewness in the first period. This is not the case in the second period.

The aforementioned relations are also confirmed by the correlation coefficients that are reported in Table 3.2. In the first period all sentiment proxies are positively related to each other with the pair of Adv-Sent and Spec-Sent having the highest correlation coefficient (0.51) and the pair of Spec-Sent and Ind-Sent having the lowest correlation coefficient (0.22). Risk-neutral skewness appears to be positively related to all sentiment proxies with the correlation coefficients ranging from 0.61 in the case of Adv-Sent to 0.37 in the case of Ind-Sent. This pattern changes in the second period. More specifically, the Adv-Sent and the Ind-Sent remain positevely

normalized net exchanges are positively correlated with those of equity funds but negatively correlated with those of bond funds. Moreover, the beta of mixed funds with equity funds is 0.93 while their beta with bond funds is 0.02.

correlated (0.57), but the Spec-Sent is negatively correlated with both the other sentiment measures. Similarly, risk-neutral skewness is negatively correlated with the Adv-Sent and the Ind-Sent (insignificantly in the case of the Ind-Sent), and positively but weakly correlated with the Spec-Sent. The correlations presented in Table 3.2 suggest that the relations among the sentiments of the three groups of investors examined (large speculators, investment advisors and individual investors) and the risk-neutral skewness have changed substantially after 1997:06. Therefore, the correlation coefficients for the overall period provide mixed results, sometimes driven by the first period (e.g. Adv-Sent and Spec-Sent) and sometimes driven by the second period (e.g. Adv-Sent and Skewness).

3.3.3 Rational updating of beliefs estimation

Brown and Cliff (2005) and Baker and Wurgler (2006) acknowledge that investor sentiment can be seen as the sum of two components: one reflecting investors' rational expectations about future returns and a second reflecting investors' irrational beliefs. Both components can affect the shape of the risk-neutral density, through their impact on the slope of the pricing kernel. Chen (1991) finds that a number of state variables such as the market dividend yield, the term spread and the default spread contain valuable information about expected excess market return, due to their correlation with current and future growth rates of economic activity. Therefore, it is reasonable to assume that such variables also have an impact on investor sentiment and reflect rational updating of investors' beliefs. In this study, we use a rich set of macroeconomic variables that practically cover all economic indicators that can possibly affect investors' beliefs about future market returns, and decompose investor sentiment into two components: one that is related to economic activity and corresponds to investors' rational expectations regarding future market returns, and a second component that is unrelated to economic activity and captures investors' unjustified optimism or pessimism.

Our initial macroeconomic dataset is comprised of monthly observations covering the period 1990:01 to 2011:06 for 131 macro variables.[10] Similar datasets have been previously used by Stock and Watson (2002, 2005), Ludvigson and Ng (2007, 2009, 2011) and Maio and Philip (2014a,b). The macroeconomic variables belong to eight main categories, namely, output and income; employment; housing; consumption, orders and inventories; money and credit; interest rates, exchange rates and spreads; prices and the stock market.[11] It can be seen that the dataset includes not only pure macroeconomic variables but also financial variables. This is important since it is reasonable to assume that investors' rational beliefs about future market returns are affected by both sources of information.

3.3.3.1 Main sentiment decomposition

For our main sentiment decomposition we consider eight major economic indicators, one from each category, thus capturing all the different aspects of the macroeconomic and financial environment that investors observe in order to form their expectations. We choose to use only one variable from each category, in order to have a parsimonious representation. In particular the variables we select are industrial production (IP: total), nonfarm payroll (Emp: total), housing starts (Starts: nonfarm), Purchasing Managers' Index (PMI), money supply M2 (M2), term spread (10 yr-FF spread), Personal Consumption Expenditure deflator (PCE defl) and aggregate stock market momentum (S&P 500). Most of the variables have been used in a similar context by Brown and Cliff (2005), Baker and Wurgler (2006) and Lemmon and Portniaguina (2006). In order to estimate the macroeconomic fundamentals driving the sentiment measures, we estimate the following regression for each of the sentiment proxies:

$$Sent_{it} = a + \boldsymbol{\beta}' \mathbf{z}_t + e_{it}, \tag{3.1}$$

[10]The original dataset consists of 132 variables. However, the variable representing the "Non-borrowed reserves of depository institutions" is eliminated as it takes negative values in 2008. This is a measurement error due to the fact that the total reserves should have been consistent with the Federal Open Market Committee's objective for the federal funds rate. Therefore, since the borrowings of the Term Auction Facility were larger than the total reserves, the non-borrowed reserves appeared to be negative.

[11]A detailed description of the dataset is provided in Table B.1 of Appendix B.

where $Sent_{it}$ is the sentiment proxy i at time t, and $\mathbf{z}_t$ is the vector of the macroeconomic variables. The fitted values from the regressions are regarded as the component that is related to economic activity and corresponds to investors' rational updating of beliefs (EF component), while the residuals from the regressions are regarded as the component that is not related to economic conditions (EB component), i.e. investors' unjustified optimism or pessimism. Note that including both sentiment components into a model that explains risk-neutral skewness is econometrically almost identical to including the aggregate sentiment together with the macroeconomic variables. However, we prefer to follow the first approach since it is more widely used in the literature and the results can be more easily interpreted.

Table 3.3 reports the results of regressing each sentiment proxy on the vector of macroeconomic variables described above. Panel A reports the results for the first period, 1990:01-1997:06, while Panel B reports the results for the second period, 1997:07-2011:06. We observe that in general the macroeconomic variables do not have the same impact on the sentiment proxies in the two periods. Panel A shows that in the first period all three sentiment proxies are mainly driven by the stock market momentum. In particular, higher (lower) momentum leads to a more optimistic (pessimistic) sentiment for all three groups of investors. Other variables such as industrial production, nonfarm payroll, housing starts and term spread appear to play some role, but their effect is not consistent across sentiment proxies. Panel B shows that in the second period stock market momentum continues to be the main determinant of Adv-Sent and Ind-Sent but has no significant effect on Spec-Sent. Moreover, the sentiments of advisors and individual investors tend to be positively related to current macroeconomic and financial conditions, while the opposite is true for the sentiment of large speculators.

The above empirical evidence regarding the differential reaction of the Spec-Sent versus the Adv-Sent and Ind-Sent to fundamentals in the most recent years constitutes a novel finding and can account for the negative correlations between Spec-Sent and the other two sentiment measures presented in Panel B of Table 3.2.

Further support comes from the correlation coefficients between the EF components of the three measures. While in the first period they are all positively correlated to each other, with the pair of Adv-Sent and Ind-Sent having the highest correlation coefficient (0.70) and the pair of Adv-Sent and Spec-Sent having the lowest correlation coefficient (0.49), in the second period the EF component of Spec-Sent is negatively correlated with the respective components of both Adv-Sent (-0.62) and Ind-Sent (-0.34). In contrast, the EF components of Adv-Sent and Ind-Sent have a remarkable positive correlation of 0.91. In essence, the decomposition results show that investment advisors and individual investors consistently expect a continuation of recent economic conditions across the whole time period examined. In contrast, large speculators tend to expect a continuation of recent economic conditions during the first period but a reversal of recent economic conditions during the second period.

3.3.3.2 Alternative Sentiment Decompositions

While our main sentiment decomposition is based on a set of eight indicators that represent eight major segments of the economy, it is possible that a broader range of variables is needed in order to capture all the macroeconomic information that drives investors' expectations. Therefore, we also consider an alternative sentiment decomposition that makes use of our entire dataset of the 131 macroeconomic variables. This is of particular importance, since the key signalling variables that drive investor sentiment are unobserved. In order to utilize our full macroeconomic dataset, we create a set of latent common factors using the asymptotic principal component analysis (APCA) method of Connor and Korajczyk (1986). Details about the method can be found in Appendix B. These factors capture the common information among the 131 macroeconomic variables of our dataset. Using the second information criterion of Bai and Ng (2002), we find that the first eight factors adequately summarize the macroeconomic variations. Therefore, the alternative sentiment decomposition relies on regressions of the aggregate sentiment proxies on the estimated common factors.

While these factors by construction are associated with all the macroeconomic variables, some factors load heavily on particular groups of variables. Hence, it is possible to characterize the factors by examining how they load on each macroeconomic variable. To this end, we regress each factor on each of the 131 variables and plot the respective R^2s (Figure B.1 in Appendix B). This way, each factor can be associated with one or more groups of macroeconomic variables. The first factor mostly loads on the variables of output, employment and orders, so it can be considered a "real activity" factor. The second factor loads heavily on price indices. The third factor is mostly associated with interest rate spreads, while the fourth factor is mostly correlated with inventories and consumption variables. The fifth factor loads mainly on output and stock market variables. The sixth factor is mainly related to interest rates and exchange rates. Finally, the seventh factor loads mainly on housing variables, while the eighth factor is mainly driven by money supply and bank reserves.

Since the estimated common factors summarize the information embedded in all 131 variables that constitute our macroeconomic dataset, it is likely that they also eclipse the idiosyncratic signal encapsulated in each variable. Another concern is the potential for look-ahead bias, given the fact that the construction of the factors requires the usage of data from the full sample period. Therefore, as a robustness check, we consider a third alternative sentiment decomposition that uses a selection of variables that are highly correlated with the common factors. In particular, for each factor we consider the three variables with the highest R^2s and choose from those the one that is most important and widely used in the literature. The variables we obtain following this approach are nonfarm payroll (Emp: total), Consumer Price Index (CPI-U: all), term spread (10 yr-FF spread), inventories to sales ratio (M&T invent/sales), aggregate stock market momentum (S&P 500), Baa corporate bond yield (Baa bond), housing starts (Starts: nonfarm) and money supply M1 (M1). It is apparent that the information embedded in the aforementioned explanatory variables is similar to the information embedded in the variables used in the main sentiment decomposition, hence validating our initial selection.

3.3.4 Control variables

Similar to Han (2008), we include a series of control variables in the regression analysis. First, we control for the autocorrelation in the risk-neutral series by adding the lagged skewness value (LagRNS). The second control variable, is the ratio of the open interest of out-of-the-money (OTM) puts to the open interest of near-the-money (NTM) calls and puts, and represents relative demand pressure (RelDem) (see Bollen and Whaley, 2004, Gârleanu et al., 2009). More specifically, a higher relative demand value implies that there is high expectation among investors about a downturn in the market and therefore the demand for OTM puts for hedging purposes increases. The next variable is the options' trading volume (TrVlm), which is considered a proxy for dispersion in investors' beliefs (see Buraschi and Jiltsov, 2006). In particular, we take the natural logarithm of the detrended trading volume. Further, we include the contemporaneous volatility of S&P 500 index (Vol) proxied by the VIX index, as it is considered the main determinant of risk-neutral skewness in stochastic volatility pricing models (e.g. Heston, 1993) and is also theoretically linked to skewness by Bakshi, Kapadia and Madan (2003).[12]

Summary statistics for the control variables can be found in Table 3.1. RelDem has a mean value of 1,837 with a close to zero autocorrelation coefficient, while TrVlm has a zero mean value due to the deterministic time trend adjustment and an autocorrelation coefficient of 0.393. Vol has a mean value of 0.203 and it is quite persistent with an autocorrelation coefficient of 0.861.

The correlation coefficients between the control variables and risk-neutral skewness can be found in Table 3.2. RelDem and TrVlm are negatively related to risk-neutral skewness in both periods (for both variables, however, the correlation is

[12]Han (2008) considers aggregate stock market momentum as an additional control variable in his explanatory model. In our case, the information of momentum is embedded in the EF sentiment component.

insignificant in the first period) with coefficients ranging from -0.01 to -0.26. The documented negative correlation for RelDem is in accordance with the limits to arbitrage hypothesis of Bollen and Whaley (2004). Intuitively, a higher demand for OTM puts in relation to NTM options drives the prices of those contracts up because of the markets makers' increased risk exposure and hedging costs. Furhermore, the documented negative correlation for TrVlm is in accordance with the heterogeneous agents model of Buraschi and Jiltsov (2006). Unlike the other two variables, Vol is negatively correlated with risk-neutral skewness during the first period (-0.25) but exhibits a positive correlation in the second period (0.24).[13] The documented negative correlation for Vol in the first period is in accordance with the theoretical prediction of Bakshi, Kapadia and Madan (2003) in the presence of excess kurtosis in the physical density. The documented positive correlation for Vol in the second period is in line with stochastic volatility models, such as Heston (1993). The correlation coefficients for the overall period show that only TrVlm exhibits a significant (negative) correlation with risk-neutral skewness.

3.4 Empirical Analysis

This section explores in detail the linkages between the S&P 500 index risk-neutral skewness and the sentiment of large speculators, investment advisors and individual investors. It is possible for sentiment to influence risk-neutral skewness due to the existence of limits to arbitrage in the options market. In particular, a bearish or bullish sentiment creates demand for a specific class of options (e.g. OTM puts). As market makers satisfy this demand they face more difficulty in hedging their positions and therefore charge higher prices. From their point of view, investors

[13] Unlike our finding, Han (2008) documents a positive relation between volatility and risk-neutral skewness for the period 1988:01-1997:06. However, the negative relation we find for the period 1990:01-1997:06 is driven only by the observations of year 1990. If we remove this year from our sample Vol exhibits a positive correlation of 0.19 with skewness. Since the correlation between the two variables appears to be very sensitive to the sample period considered, we attribute the difference in results between this study and Han (2008) to the difference in sample periods. We confirm that the same relations between Vol and risk-neutral skewness hold if instead of our skewness estimates we use the CBOE SKEW index available at `http://www.cboe.com/micro/skew/introduction.aspx`.

are willing to accept these higher prices due to their sentiment. Hence, we observe changes in the shape of the risk-neutral distribution and implicitly in the pricing kernel. In this section, we first examine the relation of the risk-neutral skewness with the aggregate sentiment and then with the two distinct sentiment components, the EF component and the EB component. Subsequently, we investigate whether the documented relations between skewness and the EF component exhibit an asymmetric pattern depending on whether stock market conditions improve or worsen and whether the EF component becomes more optimistic or pessimistic. Finally, we examine whether there is a differential impact of the two sentiment components on slope measures of the implied volatility smirk created separately by call and put options. We conduct our analysis over two periods - the first dating from 1990:01 to 1997:06 and the second from 1997:07 to 2011:06. This enables us to compare out findings to those of Han (2008), whose sample period ends at 1997:06. Further, we are also able to account for the possible structural change in the S&P 500 index options market due to the introduction of the E-mini contracts and the Dow-Jones options in 1997.

3.4.1 Risk-neutral skewness and aggregate sentiment

Table 3.4 shows the results of regressing S&P 500 index risk-neutral skewness on the three sentiment measures used in the study. Panel A reports the results for the period 1990:01-1997:06. Similar to Han (2008), all sentiment measures are both positive and statistically significant. This result implies that a more bearish (bullish) sentiment of either large speculators, investment advisors or individual investors leads to a more (less) negatively sloped pricing kernel and a more (less) negative risk-neutral skewness. In other words, when investors are pessimistic (optimistic) about future market returns, they are willing to pay more (less) to protect their portfolios from possible downturns in the stock market. In economic terms, a one standard deviation increase of Adv-Sent, Spec-Sent and Ind-Sent is followed by an increase of approximately 0.19, 0.09 and 0.08 in the risk-neutral skewness. These values represent about 43%, 19% and 18% of skewness' sample standard deviation respectively. Han (2008) argues that such changes cannot be attributed to measurement errors in option prices such as bid-ask bounce. The above results remain statistically significant once the control variables are introduced into the analysis, consistent with Han's (2008) findings.

The results in Panel B are intriguing and show that the previous pattern changes substantially in the second period. In particular, all three sentiment proxies become negative with Adv-Sent and Ind-Sent being also significant. However, both of them turn insignificant once the control variables are included into the explanatory model. Apparently, in the second period there is no considerable relationship between risk-neutral skewness and the sentiment of the three investor groups examined. A possible explanation for the results in Panel B is that in the second period the market makers are willing to provide liquidity to investors at lower prices than those that the investors are willing to accept when their sentiment is either low or high. In the same vein, Han (2008) asserts that the risk-neutral skewness and sentiment relation is much weaker in periods of low limits to arbitrage. Another possible explanation, however, which will be examined in the next section, is that the aggregate sentiment constitutes a noisy aggregation of the two sentiment components which can separately affect the risk-neutral skewness.

Regarding the control variables, they are insignificant in the first period and while mainly negative, they turn positive in a few cases. In the second period the picture is clearer with RelDem and TrVlm being consistently negative and Vol being consistently positive. Furthermore, the results for TrVlm and Vol are significant in all but one case (Vol for the Adv-Sent). These relations are in line with the correlations between the control variables and risk-neutral skewness described in the previous section.

Summarizing the above mentioned empirical evidence, we find that aggregate sentiment plays an important role in determining the level of the risk-neutral skewness only in the period 1990:01-1997:06. More specifically, aggregate sentiment of all three groups of investors is positively related to risk-neutral skewness, implying that a more bearish (bullish) investor sentiment leads to a more (less) negatively sloped pricing kernel. In the second and most recent period 1997:07-2011:06 there is no significant relation for any of the sentiment measures once we control for relative demand pressure, heterogeneity in beliefs and contemporaneous volatility.

3.4.2 Risk-neutral skewness and sentiment components

Previous literature documents that the variation in the risk-neutral skewness is mainly driven by changes in the slope of the pricing kernel due to the approximately symmetric conditional physical probability distribution (Aït-Sahalia et al., 2001 and Rosenberg and Engle, 2002). Shefrin (2008) asserts that the pricing kernel can be decomposed into two components: one component that is driven by investors' rational expectations about future consumption and a second component that stems from investors' erroneous beliefs. In accordance with Shefrin's model, in Section 3.3.3 we decomposed investor sentiment into an economic fundamentals and an error in beliefs components. In light of this, the main aim of this section is to examine which part of sentiment drives the variation in the index risk-neutral skewness. In doing so, we will gain useful insights about the way the economic fundamentals and error in beliefs components affect index option prices and the pricing kernel.

If aggregate sentiment has a significant impact on risk-neutral skewness, this can originate from the EF component, the EB component or a combination of both. In that respect, the analysis of the first period will allow us to draw inferences about the source of the positive relationship between risk-neutral skewness and investor sentiment documented by Han (2008) and also confirmed in the previous section. Han (2008) conjectures that this relation stems from investors' erroneous expectations and his assumption is reinforced by the fact that the main results of his study do not change after controlling for four popular macroeconomic indicators. Our empirical analysis seeks to scrutinize Han's assumption by examining which of the two sentiment components actually affects risk-neutral skewness and hence the pricing kernel.

If aggregate sentiment does not significantly affect risk-neutral skewness, as in the second period, this does not necessarily mean that no component affects it. In fact it is possible that one (or even both) of the two components has a significant impact on skewness, which vanishes when we sum the two components into an aggregate sentiment. In that respect, the analysis of the second period will allow us to investigate whether either of the two separate sentiment components influences risk-neutral skewness. We hypothesize that if the market matures with time, it is more likely that the EF component will be significant. If no component turns out to be significant, this finding will imply either that there are no limits to arbitrage in the index options market or that the information embedded in sentiment is subsumed by some of the control variables.

The results from regressing risk-neutral skewness on both parts of each sentiment proxy following the main decomposition are reported in Table 3.5. Panel A shows the results for the first period and Panel B the results for the second period. For each sentiment proxy we report results before and after controlling for Vol, since stock market volatility is known to depend on the macroeconomic conditions (Brandt and Kang, 2004), which also drive the variation in the EF sentiment component. Therefore, we examine how the EF component reacts to the inclusion of market volatility in the analysis. In Panel A, we observe that the positive and significant relation between sentiment and risk-neutral skewness documented in the previous section mainly comes from the EF component. This finding combined with the sentiment decomposition result provided in Section 3.3.3 implies that during the first period risk-neutral skewness is driven by the similar expectations of all three investor groups regarding a continuation of recent economic conditions. In the case of Adv-Sent the EB component also appears significant, even when the control variables are included into the model. In economic terms, however, a one standard deviation increase of the Adv-Sent, Spec-Sent and Ind-Sent EF (EB) component is associated with an increase in risk-neutral skewness corresponding to 37%, 17% and

32% (16%, 10% and 9%) of its sample standard deviation respectively. Hence, both in statistical and economic terms the impact of the EF component on skewness is stronger than that of the EB component.

Panel B shows that the pattern is different in the second period. In particular, the EF component of all three sentiment proxies remains strongly significant, while their EB component is always insignificant even in the absence of any control variables. Another intriguing result from Panel B is the negative sign of the coefficients of the Adv-Sent and the Ind-Sent EF components in contrast to the positive sign of the Spec-Sent EF component coefficient. These signs combined with the evidence of Section 3.3.3 imply that in the second period risk-neutral skewness is driven by the expectations of large speculators regarding a reversal of recent economic conditions.[14] In economic terms a one standard deviation increase of the Adv-Sent and Ind-Sent (Spec-Sent) EF component is related to a decrease (increase) in risk-neutral skewness by about 17% and 13% (15%) of its sample standard deviation respectively.

The empirical evidence regarding the second period is in line with our hypothesis that the S&P 500 index options market has become more mature with time, as option prices are only driven by investors' rational updating of beliefs due to changes in fundamentals. Of course, the fact that the EF sentiment component is significant implies that there are still limits to arbitrage that prevent market makers from having flat supply curves. In a similar vein, Constantinides et al. (2009) find that OTM calls have been systematically overpriced during 1997 - 2003, even when bid-ask spreads and trading costs are taken into consideration.

Turning to the control variables, the results are qualitatively similar to the ones presented in the previous section with the exception of Vol in the second period which is now positively but insignificantly related to risk-neutral skewness. This implies that the information embedded in Vol for skewness is subsumed by the EF component of all three sentiment proxies. In essence, the explanatory power of stock

[14]This is because conventional wisdom dictates a positive relationship between risk-neutral skewness and the sentiment of those trading in the options market.

market volatility for risk-neutral skewness stems from the fact that they are both related to macroeconomic fundamentals. In fact, the correlation of Vol with the EF component of Adv-Sent, Spec-Sent and Ind-Sent is -0.61, 0.33 and -0.53 respectively.

Table 3.6 presents the results from regressing risk-neutral skewness on the sentiment components following the decomposition based on the common factors. It can be seen that the results are qualitatively similar to those from Table 3.5, but the EF component is slightly less significant in all cases, apart from the case of Spec-Sent in the first period. Furthermore, the regression R^2s in the second period are always lower or equal to the respective R^2s presented in Table 3.5. This empirical evidence implies that at least in the second period, the idiosyncratic component of various major economic indicators is important for explaining the risk-neutral skewness with macroeconomic fundamentals. The results from regressing skewness on the sentiment components following the decomposition based on the alternative selection of macroeconomic variables are qualitatively and quantitatively almost identical to those presented in Table 3.5 and thus are presented in Table B.2 of Appendix B.

In summary, this section shows that the significant relationship between risk-neutral skewness and aggregate sentiment in the first period stems mainly from the EF sentiment component but the EB component exhibits also some explanatory power. In the second period, the explanatory power of the EB component vanishes and only the EF component has a significant impact on skewness. Moreover, the relation indicates that index options traders' beliefs are in line with the sentiment of large speculators which on average reflects an expectation of reversal of recent economic conditions. Finally, there is evidence that the EF sentiment component captures additional information to that captured by stock market volatility.

3.4.3 Risk-neutral skewness and sentiment components in different periods

The empirical evidence presented in the previous section establishes a strong link between the EF sentiment component and the S&P 500 index risk-neutral skewness.

It is possible, however, that traders in the index options market react differently to worsened than to improved economic conditions. Therefore, the aim of this section is to investigate whether the documented relation between the EF component and risk-neutral skewness exhibits an asymmetric pattern depending on the recent economic conditions and subsequently on whether the EF component becomes more bullish or bearish. To this end, we repeat the analysis of Section 3.4.2 investigating separately months of improved and deteriorating stock market momentum and months of increased and decreased EF sentiment component. In particular, we create dummy variables based on whether the stock market momentum and the EF component increase or decrease relative to the value of the previous month.[15] Similar to Han (2008) these dummy variables are used as interaction terms for all the regressors except for the lagged dependent variable. Due to the relatively low number of observations in the first period we conduct this analysis only for the second period.

Table 3.7 presents the results. Panel A shows the results when the sample is split based on past momentum and Panel B the results when the sample is split based on the EF component. From Panel A we observe that the coefficients of the EF component in cases of a decreased stock market momentum are always much higher in absolute value than the respective coefficients in cases of an increased stock maket momentum. Furthermore, for all three proxies the EF component is significant when momentum decreases relative to its previous value but insignificant when momentum increases relative to its previous value. Therefore, it is apparent that the strong relation between risk-neutral skewness and the EF component documented in the previous section is mainly driven by the periods of worsened stock market conditions which lead to a less negative risk-neutral skewness since on average index options traders anticipate a reversal.

The above empirical evidence combined with the observation that during the second period the EF component of the Adv-Sent and the Ind-Sent tend to be

[15]For example, the increased (decreased) momentum dummy takes a value of one whenever the one-month S&P 500 index return is higher (lower) than that of the previous month and zero otherwise.

positively related to recent economic conditions whilst the EF component of the Spec-Sent tends to be negatively related to recent economic conditions, implies a more pronounced relation when the EF component of the Adv-Sent and the Ind-Sent becomes more bearish and the EF component of the Spec-Sent becomes more bullish. This is exactly what we observe in Panel B of Table 3.7. The coefficients of the Adv-Sent and Ind-Sent (Spec-Sent) EF component are much higher and much more significant when there is a decrease (increase) in the EF component relative to its previous value than when there is an increase (decrease) relative to its previous value.

Turning to the control variables, it is interesting to note that RelDem and Vol mainly affect risk-neutral skewness in periods of declined stock market conditions, while the impact of TrVlm mostly comes from periods of improved stock market conditions.

Overall, the empirical evidence of this section reveals that while after 1997:06 the S&P500 index risk-neutral skewness is driven by the reversal expectations embedded in the sentiment of large speculators, this is more pronounced when recent stock market conditions deteriorate. In particular, a decreased stock market momentum leads to a more bullish (bearish) Spec-Sent (Adv-Sent and Ind-Sent) and a less negative risk-neutrals skewness. The opposite relation does not appear to be significant.

3.4.4 Slope measures from calls and puts and sentiment components

The empirical analysis of this section investigates whether the EF and EB sentiment components have different impact on implied volatility slope measures created separately from call and put options. Our motivation comes from Constantinides et al.'s (2011) assertion that their results are consistent with an equilibrium in a segmented index options market where OTM puts are mainly traded by hedgers for portfolio insurance and OTM calls are mainly traded by optimistic investors for speculative purposes. In the same vein, Lemmon and Ni (2011) find that the positive-exposure demand for index calls is more correlated with the positive-exposure demand for stock options than with the positive-exposure demand for index puts. Recall from the results of the Section 3.4.2 that in the second period S&P 500 index risk-neutral skewness is mainly affected by large speculators' reversal expectations. Since the negative risk-neutral skewness is mainly driven by the high prices of OTM puts (Rubinstein, 1994, Jackwerth and Rubinstein, 1996), we expect to find that the same relation holds for the skewness proxy extracted from put options as well. On the contrary, if call and put options markets are indeed segmented, we hypothesize that the same relation may not hold for the skewness proxy extracted from call options.

To investigate the possibility of a differential impact of the two sentiment components on call and put options prices, we construct for each type of options two measures of the slope of the respective implied volatility smirk. In particular, for each cross-section of either calls or puts we interpolate implied volatilities into the range of available options data using a smoothing cubic spline and extrapolate outside this range using the respective boundary values. Then for call options we create the following slope measures capturing the difference in implied volatilities between OTM and deep-out-of-the-money (DOTM) calls and at-the-money (ATM) and DOTM calls:

$$CO\_Slope = IV_{OTM} - IV_{DOTM} \tag{3.2}$$

$$CA\_Slope = IV_{ATM} - IV_{DOTM} \tag{3.3}$$

where IV_{DOTM}, IV_{OTM} and IV_{ATM} are the implied volatilities of call options corresponding to $K/S = 1.125$, $K/S = 1.075$ and $K/S = 1$ respectively. Similarly for puts we create the following slope measures capturing the difference in implied volatilities between DOTM and OTM puts and DOTM and ATM puts:

$$PO\_Slope = IV_{DOTM} - IV_{OTM} \tag{3.4}$$

$$PA\_Slope = IV_{DOTM} - IV_{ATM} \tag{3.5}$$

where IV_{DOTM}, IV_{OTM} and IV_{ATM} are the implied volatilities corresponding to $K/S = 0.875$, $K/S = 0.925$ and $K/S = 1$ respectively. If for a given cross-section two or more of the desired moneynesses are outside the range of available options data, then this cross-section is discarded. We estimate the slope measures on the last trading day of the month for the two time horizons that are nearest to one month and interpolate to find the exact one-month ahead slope measures for calls and puts markets. We investigate only the second part of our sample period as in the first part the liquidity of high-moneyness calls is quite low and there are many days for which we cannot estimate the slopes measures for call options. The summary statistics of the above variables can be found in Table 3.1 Panel B. The slope measures from puts have higher mean values than the respective slope measures from calls, showing that the slope of the implied volatility function is steeper on its left side than on its right side. Moreover, while the slope measures from puts are invariably positive, the slope measures from calls turn occasionally negative implying the existence of an implied volatility smile pattern. All the slope measures are moderately persistent with autocorrelation coefficients ranging from 0.327 to 0.380.

Table 3.8 reports the results of regressing the slope measures from call and put options on the two components of each sentiment proxy. Panel A reports the results for the slope measures created only by call options, while Panel B shows the results for the slope measures created only by put options. In both Panels A and B there is a clear picture showing that the EB sentiment component is never significant, while the EF component is most of the times strongly significant. This evidence further supports our previous findings regarding the absence of any significant relation between investors' erroneous beliefs and option prices in the second period. Moreover, the EF component is significant in all cases for CO_Slope and PO_Slope but in two out of six cases for CA_Slope and PA_Slope. This is an intuitive result as the prices of away-from-the-money options mainly reveal investors' expectations about future returns while the prices of ATM options have the highest vegas and hence are also related to investors' expectations about future volatility (Ni et al., 2008).

In Panel A, we observe that the EF sentiment coefficients for the Adv-Sent and the Ind-Sent are negative while those for the Spec-Sent are positive. Since a negative (positive) change in the slope measures from call options implies that the slope becomes flatter (steeper), the above evidence suggests that the slope measures from call options are mainly driven by the expectations of investment advisors and individual investors regarding a continuation of recent economic conditions. The coefficients of the TrVlm and the Vol are always positive and in the majority of the cases strongly significant across sentiment proxies and slope measures, indicating that higher trading volume and volatility are associated with a steeper calls' slope. The effect of RelDem is always insignificant due to the fact that by definition it captures hedging demand pressure for OTM puts and therefore cannot account for the variation in the implied volatility slope of call options.

In Panel B, we observe that the EF sentiment coefficients for the Adv-Sent and the Ind-Sent are positive, while those for the Spec-Sent are all negative. Since a negative (positive) change in the slope measures from put options implies that the slope becomes flatter (steeper), we conclude that the slope measures from put options are driven by the expectations of large speculators regarding a reversal in the economy. As expected this relationship is similar to the one documented for the risk-neutral skewness. The coefficients of the control variables are consistently positive across sentiment proxies and slope measures (with the exception of Vol when considering the effect of Ind-Sent on PA_Slope), with RelDem and TrVlm being significant mostly for the PA_Slope. These results imply that a higher hedging demand and higher volume are related to a steeper implied volatility slope of put options.[16]

Summarizing the above empirical evidence, our results support Constantinides et al.'s (2011) statement that the index options market is segmented. In particular, the call options traders' beliefs are in line with the expectations of advisors and indi-

[16]If we do not control for the sentiment components, the relation between Vol and the two slopes measures from puts is negative, hence resembling the relation between Vol and risk-neutral skewness. It is insignificant, however, due to the fact that the implied volatility slope measures do not contain only information about skewness but also about the interaction of skewness with volatility and kurtosis (Mixon, 2011).

vidual investors regarding a continuation of recent economic conditions. In contrast, the put options traders' beliefs are in line with the expectations of large speculators regarding a reversal of recent economic conditions. Again there is no evidence that investors' erroneous beliefs have any significant impact on options prices in the second period.

3.5 Conclusion

In this study, we decompose the sentiment of three main groups of investors, i.e. large speculators, investment advisors and individual investors, into two parts - one part that is driven by economic fundamentals (EF) and represents investors' rational updating of beliefs about future returns, and a second part that is unrelated to fundamentals and represents errors in investors' beliefs (EB). The main aim of the study is to examine how the rational and the irrational components of investor sentiment drive the variations in risk-neutral skewness and hence the pricing kernel.

We estimate the EF component using a parsimonious set of variables that reflect the information embedded in eight main macroeconomic categories. In this way we take into consideration various aspects of the macroeconomic and financial environment that can possibly drive investors' beliefs about future market returns. The predicted values from the regression of aggregate sentiment measures on these macroeconomic variables constitute the estimated EF component, while the residuals are regarded as the EB component unrelated to fundamentals. Alternative sentiment decompositions based on common latent factors estimated using asymptotic principal component analysis provide qualitatively similar results.

We conduct our analysis for two time periods: the first from 1990:01 to 1997:06 and the second from 1997:07 to 2011:06. Similar to previous literature, we find that aggregate investor sentiment affects S&P 500 index risk-neutral skewness only in the first period. Our results show that this relation mainly stems from the EF component but the EB component appears to play some role too. Contrarily, in the second period (after 1997:06) the significant effect of the EB component vanishes, while the EF component remains strongly significant and implies that skewness is driven by the sentiment of large speculators and reflects an anticipation of a reversal of the economy. Moreover, the significant impact of the EF component on skewness is more prominent in periods of worsened stock market conditions. We further document that the EF component has opposite effects on implied volatility slope measures created separately from calls and puts. Our results show that the slope measures from calls are mainly driven by the expectations of investment advisors and individual investors regarding a continuation of recent economic conditions, while the slope measures from puts are mainly driven by the expectations of large speculators regarding a reversal in the economy. This result provides strong evidence in favor of Constantinides et al.'s (2011) assertion that the call and put options markets are segmented.

The empirical evidence in this study has important implications for the asset pricing literature as options encapsulate forward-looking information about the pricing kernel. In particular, our results demonstrate that the pricing kernel is mainly driven by investors' rational updating of beliefs and most importantly that in the second period investors' erroneous beliefs play no role at all. Therefore, incorporating investors' irrationality into sophisticated option pricing models does not appear to be a necessary extension anymore, at least for mature options markets such as S&P 500 index options market.

Table 3.1: Summary statistics of variables

	Mean	**StDev**	**Min**	**Max**	**Auto**
Panel A: Skewness and Explanatory Variables					
RNS	-1.559	0.447	-3.048	-0.483	0.547
Reldem	1.837	1.476	0.594	17.629	-0.020
TrVlm	0.000	0.598	-3.089	1.825	0.393
Vol	0.203	0.078	0.104	0.599	0.861
Adv-Sent	0.139	0.152	-0.264	0.423	0.746
Spec-Sent	-0.045	0.057	-0.205	0.105	0.817
Ind-Sent	-0.006	0.175	-1.090	0.748	0.219
Panel B: Implied Volatility Slopes					
CO_Slope	0.001	0.009	-0.021	0.042	0.327
CA_Slope	0.033	0.021	-0.017	0.147	0.380
PO_Slope	0.045	0.011	0.014	0.067	0.337
PA_Slope	0.104	0.022	0.055	0.164	0.352

This table reports summary statistics of the variables used in the empirical analysis. The sample period for the variables in Panel A is 1990:01-2011:06 while the sample period for variables in Panel B is 1997:07-2011:06. RNS is the S&P 500 index risk-neutral skewness estimated using the model-free method of Bakshi, Kapadia and Madan (2003). RelDem is the relative demand pressure as captured by the ratio of the open interest of OTM puts to the open interest of NTM calls and puts. TrVlm is the heterogeneity of beliefs, proxied by the detrended logarithm of options trading volume. Vol is the index instantaneous volatility as proxied by VIX. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds. CO_Slope and CA_Slope denote the difference in implied volatility between OTM and DOTM calls and ATM and DOTM calls respectively. PO_Slope and PA_Slope denote the difference in implied volatility between DOTM and OTM puts and DOTM and ATM puts respectively.

Table 3.2: Correlation coefficients

Panel A: 1990:01 - 1997:06							
	Skewness	**RelDem**	**TrVlm**	**Vol**	**Adv-Sent**	**Spec-Sent**	**Ind-Sent**
Skewness	1.00						
	[1.00]						
RelDem	-0.01	1.00					
	[0.92]	[1.00]					
TrVlm	-0.15	-0.13	1.00				
	[0.16]	[0.23]	[1.00]				
Vol	-0.25	-0.11	-0.04	1.00			
	[0.02]	[0.28]	[0.70]	[1.00]			
Adv-Sent	0.61	-0.08	-0.25	-0.17	1.00		
	[0.00]	[0.47]	[0.02]	[0.12]	[1.00]		
Spec-Sent	0.38	-0.04	-0.14	-0.31	0.51	1.00	
	[0.00]	[0.72]	[0.20]	[0.00]	[0.00]	[1.00]	
Ind-Sent	0.37	0.08	-0.02	-0.53	0.35	0.22	1.00
	[0.00]	[0.45]	[0.84]	[0.00]	[0.00]	[0.03]	[1.00]
Panel B: 1997:07 - 2011:06							
	Skewness	**RelDem**	**TrVlm**	**Vol**	**Adv-Sent**	**Spec-Sent**	**Ind-Sent**
Skewness	1.00						
	[1.00]						
RelDem	-0.16	1.00					
	[0.04]	[1.00]					
TrVlm	-0.26	0.11	1.00				
	[0.00]	[0.14]	[1.00]				
Vol	0.24	-0.22	0.06	1.00			
	[0.00]	[0.00]	[0.45]	[1.00]			
Adv-Sent	-0.23	0.16	-0.19	-0.64	1.00		
	[0.00]	[0.04]	[0.01]	[0.00]	[1.00]		
Spec-Sent	0.08	0.11	0.04	0.02	-0.16	1.00	
	[0.30]	[0.15]	[0.58]	[0.83]	[0.03]	[1.00]	
Ind-Sent	-0.12	0.14	-0.21	-0.49	0.57	-0.15	1.00
	[0.11]	[0.07]	[0.01]	[0.00]	[0.00]	[0.06]	[1.00]
Panel C: 1990:01 - 2011:06							
	Skewness	**RelDem**	**TrVlm**	**Vol**	**Adv-Sent**	**Spec-Sent**	**Ind-Sent**
Skewness	1.00						
	[1.00]						
RelDem	0.02	1.00					
	[0.76]	[1.00]					
TrVlm	-0.15	0.02	1.00				
	[0.02]	[0.75]	[1.00]				
Vol	0.02	-0.20	-0.05	1.00			
	[0.73]	[0.00]	[0.42]	[1.00]			
Adv-Sent	-0.12	-0.11	-0.28	-0.25	1.00		
	[0.06]	[0.08]	[0.00]	[0.00]	[1.00]		
Spec-Sent	-0.04	-0.14	-0.15	0.18	0.34	1.00	
	[0.49]	[0.03]	[0.02]	[0.00]	[0.00]	[1.00]	
Ind-Sent	0.15	0.13	-0.05	-0.44	0.27	-0.06	1.00
	[0.02]	[0.03]	[0.38]	[0.00]	[0.00]	[0.30]	[1.00]

This table reports the correlation coefficients of the variables used in the empirical analysis. The respective p-values are shown in brackets. Panel A reports the correlations for the period 1990:01 - 1997:06, Panel B reports the correlations for the period 1997:07 - 2011:06, while Panel C reports the correlations for the period 1990:01 - 2011:06. RNS is the S&P 500 index risk-neutral skewness estimated using the model-free method of Bakshi, Kapadia and Madan (2003). RelDem is the relative demand pressure as captured by the ratio of the open interest of OTM puts to the open interest of NTM calls and puts. TrVlm is the heterogeneity of beliefs, proxied by the detrended logarithm of options trading volume. Vol is the index instantaneous volatility as proxied by VIX. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds.

Table 3.3: Sentiment decomposition

	Panel A: 1990:01 - 1997:06			Panel B: 1997:07 - 2011:06		
	Adv-Sent	**Spec-Sent**	**Ind-Sent**	**Adv-Sent**	**Spec-Sent**	**Ind-Sent**
IP: total	8.512***	0.803	1.619	1.022	-0.811***	-0.096
	(4.042)	(0.682)	(0.295)	(0.664)	(-2.610)	(-0.091)
Emp: total	-24.329***	9.116*	1.383	19.926**	-4.919	2.839
	(-3.346)	(1.846)	(0.081)	(2.421)	(-1.590)	(0.431)
Starts: nonfarm	0.232**	0.059	-0.525	0.073	-0.055	0.127*
	(2.458)	(1.140)	(-1.527)	(1.034)	(-1.330)	(1.975)
PMI	-0.006	-0.002	-0.003	0.002	-0.001	0.002
	(-1.249)	(-0.873)	(-0.591)	(0.620)	(-0.509)	(0.916)
M2	6.284	-0.577	-8.436	0.426	0.338	-0.621
	(1.500)	(-0.317)	(-0.735)	(0.348)	(0.558)	(-0.599)
10 yr-FF spread	0.034***	0.003	0.077***	0.013	-0.010**	0.002
	(2.657)	(0.460)	(2.887)	(1.572)	(-2.245)	(0.246)
PCE defl	7.202	1.484	-18.376	0.886	-1.165	-2.920
	(1.304)	(0.552)	(-0.980)	(0.390)	(-1.028)	(-1.421)
S&P 500	3.179***	0.561***	4.368***	1.497***	0.038	1.805***
	(8.752)	(3.210)	(3.327)	(8.386)	(0.398)	(6.451)
$\widetilde{\mathbf{R}}^2$	0.486	0.096	0.289	0.470	0.196	0.512

This table reports the results of monthly regressions of each investor sentiment proxy on a series of macroeconomic variables. A constant term is included in all the regressions but omitted for brevity. Panel A reports the results for the period 1990:01 - 1997:06, while Panel B reports the results for the period 1997:07 - 2011:06. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds. A description of the macroeconomic variables can be found in Table B.1 of Appendix B. Newey-West t-statistics are reported in parentheses below the coefficients. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Table 3.4: Risk-neutral skewness and aggregate sentiment measures

	Adv-Sent		**Spec-Sent**		**Ind-Sent**	
			Panel A: 1990:01 - 1997:06			
LagRNS	0.227***	0.213**	0.369***	0.357***	0.446***	0.438***
	(2.733)	(2.448)	(3.856)	(3.535)	(5.925)	(5.538)
RelDem		0.004		-0.001		-0.005
		(0.607)		(-0.115)		(-0.577)
TrVlm		0.019		-0.017		-0.020
		(0.272)		(-0.278)		(-0.362)
Vol		-0.942		-0.813		0.065
		(-1.402)		(-0.991)		(0.075)
Sent	1.274***	1.278***	1.526**	1.324*	0.456***	0.468***
	(4.213)	(4.208)	(2.033)	(1.939)	(3.202)	(2.650)
$\widetilde{\mathbf{R}}^2$	0.383	0.377	0.235	0.217	0.289	0.265
			Panel B: 1997:07 - 2011:06			
LagRNS	0.486***	0.423***	0.513***	0.433***	0.506***	0.437***
	(6.825)	(5.556)	(6.807)	(5.264)	(6.707)	(5.266)
RelDem		-0.041		-0.045		-0.040
		(-0.999)		(-1.090)		(-0.965)
TrVlm		-0.129***		-0.115**		-0.124**
		(-2.617)		(-2.293)		(-2.469)
Vol		0.463		0.774***		0.607**
		(1.125)		(2.910)		(2.041)
Sent	-0.385**	-0.329	-0.025	0.225	-0.365*	-0.267
	(-2.075)	(-1.126)	(-0.045)	(0.417)	(-1.720)	(-1.087)
$\widetilde{\mathbf{R}}^2$	0.269	0.292	0.257	0.288	0.265	0.291

This table reports the results of monthly regressions of S&P 500 index risk-neutral skewness on the sentiment proxies used in the study and a set of control variables. A constant term is included in all the regressions but omitted for brevity. Panel A reports the results for the period 1990:01 - 1997:06, while Panel B reports the results for the period 1997:07 - 2011:06. Risk-neutral skewness is estimated using the model-free method of Bakshi, Kapadia and Madan (2003). LagRNS is the lagged skewness value. RelDem is the relative demand pressure as captured by the ratio of the open interest of OTM puts to the open interest of NTM calls and puts. TrVlm is the heterogeneity of beliefs, proxied by the detrended logarithm of options trading volume. Vol is the index instantaneous volatility as proxied by VIX. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds. Newey-West t-statistics are reported in parentheses below the coefficients. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Table 3.5: Risk-neutral skewness and EF, EB sentiment components

	Adv-Sent			**Spec-Sent**			**Ind-Sent**		
				Panel A: 1990:01 - 1997:06					
LagRNS	0.258***	0.267***	0.246***	0.382***	0.376***	0.370***	0.422***	0.430***	0.443***
	(3.278)	(3.430)	(2.993)	(4.021)	(3.901)	(3.725)	(5.675)	(5.781)	(5.863)
RelDem		0.006	0.003		-0.002	-0.003		0.004	0.005
		(0.827)	(0.487)		(-0.287)	(-0.409)		(0.572)	(0.792)
TrVlm		0.032	0.022		-0.016	-0.020		0.023	0.030
		(0.462)	(0.331)		(-0.243)	(-0.306)		(0.325)	(0.425)
Vol			-0.835			-0.422			0.554
			(-1.288)			(-0.504)			(0.697)
EF	1.651***	1.678***	1.640***	3.739*	3.792*	3.448	0.980***	0.999***	1.069***
	(5.603)	(5.680)	(5.738)	(1.800)	(1.807)	(1.493)	(4.206)	(4.102)	(3.789)
EB	0.774***	0.802**	0.803***	0.996	0.973	0.926	0.214	0.197	0.233
	(2.642)	(2.537)	(2.673)	(1.267)	(1.237)	(1.213)	(1.418)	(1.238)	(1.412)
$\widetilde{\mathbf{R}}^{\mathbf{2}}$	0.405	0.395	0.397	0.248	0.231	0.224	0.347	0.333	0.328
				Panel B: 1997:07 - 2011:06					
LagRNS	0.462***	0.401***	0.401***	0.472***	0.409***	0.397***	0.483***	0.433***	0.421***
	(6.262)	(4.930)	(4.959)	(6.959)	(5.284)	(5.074)	(6.237)	(5.100)	(4.975)
RelDem		-0.039	-0.036		-0.063	-0.051		-0.043	-0.036
		(-0.948)	(-0.878)		(-1.616)	(-1.294)		(-1.028)	(-0.870)
TrVlm		-0.136***	-0.135***		-0.127***	-0.133***		-0.126**	-0.127**
		(-2.845)	(-2.824)		(-2.714)	(-2.810)		(-2.542)	(-2.556)
Vol			0.262			0.463			0.515
			(0.597)			(1.482)			(1.636)
EF	-0.796***	-0.936***	-0.797**	3.228**	4.060***	3.530**	-0.693**	-0.807***	-0.559*
	(-3.100)	(-3.453)	(-2.167)	(2.517)	(3.172)	(2.470)	(-2.500)	(-2.884)	(-1.684)
EB	-0.023	-0.153	-0.083	-0.935	-0.815	-0.669	-0.001	-0.119	-0.026
	(-0.085)	(-0.561)	(-0.252)	(-1.359)	(-1.297)	(-1.007)	(-0.004)	(-0.402)	(-0.087)
$\widetilde{\mathbf{R}}^{\mathbf{2}}$	0.277	0.301	0.298	0.280	0.306	0.308	0.268	0.288	0.290

This table reports the results of monthly regressions of S&P 500 index risk-neutral skewness on the EF and EB components of the sentiment proxies used in the study and a set of control variables. A constant term is included in all the regressions but omitted for brevity. Panel A reports the results for the period 1990:01 - 1997:06, while Panel B reports the results for the period 1997:07 - 2011:06. Risk-neutral skewness is estimated using the model-free method of Bakshi, Kapadia and Madan (2003). LagRNS is the lagged skewness value. RelDem is the relative demand pressure as captured by the ratio of the open interest of OTM puts to the open interest of NTM calls and puts. TrVlm is the heterogeneity of beliefs, proxied by the detrended logarithm of options trading volume. Vol is the index instantaneous volatility as proxied by VIX. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds. EF and EB are the two components of each sentiment proxy estimated as described in Section 3.3.3. Newey-West t-statistics are reported in parentheses below the coefficients. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Table 3.6: Risk-neutral skewness and EF, EB sentiment components using APCA

	Adv-Sent			**Spec-Sent**			**Ind-Sent**		
				Panel A: 1990:01 - 1997:06					
LagRNS	0.240***	0.261***	0.240***	0.359***	0.360***	0.355***	0.439***	0.446***	0.454***
	(2.968)	(3.331)	(2.892)	(3.863)	(3.822)	(3.690)	(5.586)	(5.670)	(5.839)
RelDem		0.006	0.003		-0.002	-0.003		0.000	0.000
		(0.898)	(0.534)		(-0.229)	(-0.401)		(-0.055)	(0.034)
TrVlm		0.068	0.058		0.002	-0.003		0.016	0.020
		(1.008)	(0.871)		(0.027)	(-0.047)		(0.229)	(0.290)
Vol			-0.899			-0.406			0.352
			(-1.338)			(-0.602)			(0.411)
EF	1.686***	1.805***	1.776***	5.269***	5.284***	5.062***	1.033***	1.048***	1.092***
	(4.999)	(5.242)	(5.383)	(3.298)	(3.294)	(3.227)	(3.306)	(3.187)	(2.964)
EB	0.770**	0.738**	0.724**	0.616	0.610	0.538	0.300**	0.295**	0.321*
	(2.223)	(2.220)	(2.269)	(0.839)	(0.819)	(0.735)	(2.229)	(2.080)	(1.985)
$\widetilde{\mathbf{R}}^2$	0.409	0.406	0.411	0.295	0.278	0.271	0.326	0.311	0.304
				Panel B: 1997:07 - 2011:06					
LagRNS	0.468***	0.408***	0.409***	0.472***	0.412***	0.402***	0.502***	0.451***	0.439***
	(6.126)	(4.980)	(5.052)	(6.738)	(5.210)	(5.052)	(6.174)	(5.132)	(5.098)
RelDem		-0.041	-0.038		-0.065	-0.054		-0.049	-0.041
		(-1.010)	(-0.943)		(-1.646)	(-1.340)		(-1.121)	(-0.975)
TrVlm		-0.134***	-0.132***		-0.121**	-0.126***		-0.123**	-0.123**
		(-2.811)	(-2.760)		(-2.582)	(-2.682)		(-2.473)	(-2.459)
Vol			0.298			0.426			0.628**
			(0.669)			(1.237)			(2.033)
EF	-0.711***	-0.833***	-0.652	3.638***	4.407***	3.774**	-0.449	-0.564*	-0.202
	(-2.611)	(-2.901)	(-1.571)	(2.840)	(3.260)	(2.382)	(-1.557)	(-1.795)	(-0.505)
EB	-0.155	-0.291	-0.219	-0.827	-0.664	-0.528	-0.307	-0.410	-0.296
	(-0.557)	(-1.066)	(-0.691)	(-1.313)	(-1.166)	(-0.850)	(-0.889)	(-1.211)	(-0.888)
$\widetilde{\mathbf{R}}^2$	0.271	0.294	0.291	0.280	0.305	0.305	0.261	0.282	0.286

This table reports the results of monthly regressions of S&P 500 index risk-neutral skewness on the EF and EB components of the sentiment proxies used in the study and a set of control variables. A constant term is included in all the regressions but omitted for brevity. Panel A reports the results for the period 1990:01 - 1997:06, while Panel B reports the results for the period 1997:07 - 2011:06. Risk-neutral skewness is estimated using the model-free method of Bakshi, Kapadia and Madan (2003). LagRNS is the lagged skewness value. RelDem is the relative demand pressure as captured by the ratio of the open interest of OTM puts to the open interest of NTM calls and puts. TrVlm is the heterogeneity of beliefs, proxied by the detrended logarithm of options trading volume. Vol is the index instantaneous volatility as proxied by VIX. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds. EF and EB are the two components of each sentiment proxy estimated using APCA as described in Section 3.3.3. Newey-West t-statistics are reported in parentheses below the coefficients. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Table 3.7: Risk-neutral skewness and EF, EB sentiment components in different periods

	Adv-Sent		**Spec-Sent**		**Ind-Sent**	
	Panel A: Momentum					
	Decrease	**Increase**	**Decrease**	**Increase**	**Decrease**	**Increase**
RelDem	-0.067	0.012	-0.097**	0.027	-0.096*	0.066
	(-1.358)	(0.210)	(-2.103)	(0.502)	(-1.948)	(1.140)
TrVlm	-0.094	-0.187***	-0.099	-0.198***	-0.071	-0.195***
	(-1.303)	(-3.513)	(-1.513)	(-3.698)	(-1.045)	(-3.196)
Vol	0.569	-0.314	1.004**	-0.168	0.717*	0.022
	(1.073)	(-0.725)	(2.534)	(-0.471)	(1.850)	(0.054)
EF	-1.286***	-0.383	5.890***	0.889	-0.934*	-0.474
	(-3.005)	(-0.833)	(3.803)	(0.454)	(-1.968)	(-0.948)
EB	-0.210	0.267	-1.105	-0.243	-0.154	0.530
	(-0.466)	(0.493)	(-1.446)	(-0.193)	(-0.429)	(0.845)
$\widetilde{\mathbf{R}}^{\mathbf{2}}$	0.332		0.343		0.316	
	Panel B: EF Sentiment					
	Decrease	**Increase**	**Decrease**	**Increase**	**Decrease**	**Increase**
RelDem	-0.005	-0.103	-0.078	-0.050	-0.057	0.000
	(-0.113)	(-1.118)	(-1.078)	(-1.175)	(-1.239)	(-0.003)
TrVlm	-0.088	-0.185***	-0.115**	-0.172**	-0.065	-0.193***
	(-1.031)	(-3.479)	(-2.398)	(-2.083)	(-0.924)	(-3.103)
Vol	0.145	0.321	0.619	0.127	0.319	0.366
	(0.291)	(0.654)	(1.577)	(0.332)	(0.709)	(0.963)
EF	-1.227***	-0.420	3.664*	5.238***	-1.119**	-0.492
	(-2.796)	(-0.838)	(1.843)	(2.720)	(-2.421)	(-0.757)
EB	-0.054	0.151	-0.818	-1.027	-0.131	0.437
	(-0.113)	(0.360)	(-0.707)	(-1.269)	(-0.348)	(0.610)
$\widetilde{\mathbf{R}}^{\mathbf{2}}$	0.291		0.297		0.289	

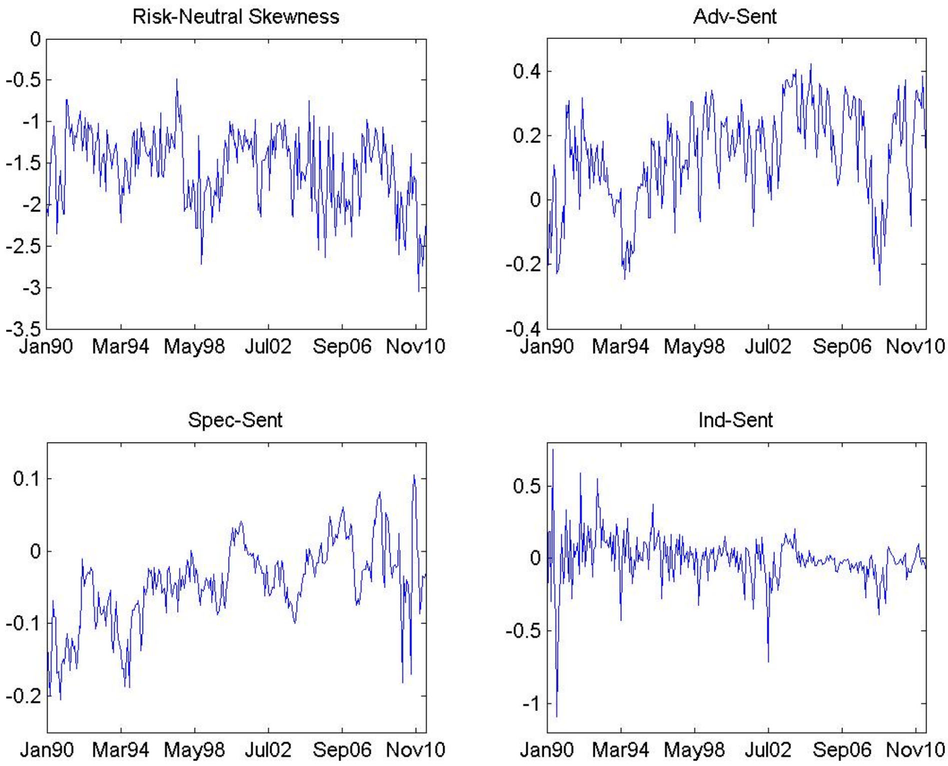
This table reports the results of monthly regressions of S&P 500 index risk-neutral skewness on the EF and EB components of the sentiment proxies used in the study and a set of control variables. A constant term and a lagged dependent variable are included in all the regressions but omitted for brevity. The sample period is 1997:07 - 2011:06. Panel A reports the results when the sample is split into periods of increased and decreased momentum, while Panel B reports the when the sample is split into periods of increased and decreased EF sentiment components. Risk-neutral skewness is estimated using the model-free method of Bakshi, Kapadia and Madan (2003). RelDem is the relative demand pressure as captured by the ratio of the open interest of OTM puts to the open interest of NTM calls and puts. TrVlm is the heterogeneity of beliefs, proxied by the detrended logarithm of options trading volume. Vol is the index instantaneous volatility as proxied by VIX. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds. EF and EB are the two components of each sentiment proxy estimated as described in Section 3.3.3. Newey-West t-statistics are reported in parentheses below the coefficients. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Table 3.8: Implied volatility slope measures from calls and puts and EF, EB sentiment components

	Adv-Sent		**Spec-Sent**		**Ind-Sent**	
	Panel A: Calls					
	CO_Slope	**CA_Slope**	**CO_Slope**	**CA_Slope**	**CO_Slope**	**CA_Slope**
LagSlope	0.026	0.135	0.019	0.143	0.053	0.160
	(0.280)	(0.983)	(0.236)	(1.089)	(0.566)	(1.198)
RelDem	0.000	-0.001	-0.001	-0.001	0.000	-0.001
	(-0.522)	(-0.741)	(-1.121)	(-1.068)	(-0.457)	(-0.665)
TrVlm	0.001	0.009***	0.001	0.009***	0.002	0.010***
	(1.335)	(3.134)	(1.513)	(3.573)	(1.529)	(3.336)
Vol	0.042***	0.076***	0.057***	0.099***	0.055***	0.102***
	(3.163)	(2.629)	(4.717)	(3.694)	(3.852)	(3.265)
EF	-0.028***	-0.039**	0.076**	0.095	-0.016**	-0.019
	(-3.063)	(-2.105)	(2.334)	(1.253)	(-2.019)	(-1.132)
EB	-0.007	-0.023	0.006	-0.027	0.010	0.029
	(-0.998)	(-1.628)	(0.348)	(-0.720)	(1.403)	(1.466)
$\widetilde{\mathbf{R}}^2$	0.384	0.364	0.365	0.352	0.363	0.359
	Panel B: Puts					
	PO_Slope	**PA_Slope**	**PO_Slope**	**PA_Slope**	**PO_Slope**	**PA_Slope**
LagSlope	0.291***	0.298***	0.256***	0.267***	0.320***	0.326***
	(3.350)	(3.609)	(2.963)	(3.319)	(3.441)	(3.808)
RelDem	0.001	0.003	0.002	0.004**	0.001	0.003
	(0.870)	(1.591)	(1.414)	(2.214)	(0.902)	(1.559)
TrVlm	0.001	0.007***	0.002	0.008***	0.002	0.007***
	(1.261)	(3.075)	(1.636)	(3.565)	(1.410)	(3.116)
Vol	0.014	0.006	0.010	0.006	0.007	-0.003
	(1.174)	(0.192)	(1.148)	(0.300)	(0.796)	(-0.124)
EF	0.023***	0.033	-0.134***	-0.244***	0.016**	0.020
	(2.866)	(1.569)	(-3.419)	(-3.306)	(2.274)	(1.049)
EB	0.004	0.000	0.013	0.028	0.011	0.024
	(0.478)	(0.026)	(0.729)	(0.818)	(1.426)	(1.325)
$\widetilde{\mathbf{R}}^2$	0.118	0.153	0.156	0.192	0.111	0.151

This table reports the results of monthly regressions of the slope of the implied volatility smirk created solely by calls (Panel A) or puts (Panel B) on the EF and EB components of the sentiment proxies used in the study and a set of control variables. A constant term is included in all the regressions but omitted for brevity. The sample period is 1997:07 - 2011:06. CO_Slope and CA_Slope denote the difference in implied volatility between OTM and DOTM calls and ATM and DOTM calls respectively. PO_Slope and PA_Slope denote the difference in implied volatility between DOTM and OTM puts and DOTM and ATM puts respectively. LagSlope is the lagged value of the respective slope variable. RelDem is the relative demand pressure as captured by the ratio of the open interest of OTM puts to the open interest of NTM calls and puts. TrVlm is the heterogeneity of beliefs, proxied by the detrended logarithm of options trading volume. Vol is the index instantaneous volatility as proxied by VIX. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds. EF and EB are the two components of each sentiment proxy estimated as described in Section 3.3.3. Newey-West t-statistics are reported in parentheses below the coefficients. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Figure 3.1: Time series of risk-neutral skewness and sentiment measures



This figure plots the monthly time series of the risk-neutral skewness and the three sentiment proxies. The top left panel plots the S&P 500 index risk-neutral skewness, as estimated using the model-free method of Bakshi, Kapadia and Madan (2003). The top right panel plots the bull-bear spread based on the Investors Intelligence's advisors sentiment index (Adv-Sent). The bottom left panel plots the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest (Spec-Sent). The bottom right panel plots the normalized aggregate net exchanges of the equity funds (Ind-Sent). The sample period is 1990:01-2011:06.

Chapter 4

Forward Skewness and its Information Content

4.1 Introduction

The forward-looking nature of the risk-neutral probability distribution has made the usage of option-implied moments and surrogate measures extremely popular for forecasting purposes among researchers. The majority of studies extract information from the shape of the implied volatility curve either using short maturity options since they tend to be the most liquid, or by weighting options from all available maturities based on their trading volume or open interest. The information embedded in the time dimension of the implied volatility surface, however, is usually ignored. To this end, a new strand of the literature explores the additional predictive ability that can be offered from the term structure of option-implied moments. In particular, Bakshi, Panayotov and Skoulakis (2011) (BPS henceforth) create measures of forward 1-month stock market variance and find that they are particularly successful in predicting future real activity as well as stock market and treasury bill returns. Luo and Zhang (2012) extend these results for stock market returns by investigating the forecasting ability of forward 3-month variances. Moreover, Mueller et al. (2013) show that both the level and the slope of the implied volatility term structure in the treasury yield market exhibit significant forecasting power for future economic activity. Finally, Feunou et al. (2014) find that two factors can summarize the information embedded in the term structure of second and higher order risk-neutral cumulants and can forecast future stock market and treasury bond returns.

In this study, we create measures of stock market forward skewness and explore its predictive ability over and above forward variance. We depart from the existing literature in two major aspects. First, unlike BPS whose method relies on the assumption that the underlying asset price follows a pure diffusion process, our method is robust to the presence of jumps. In particular, our alternative measures of variance and skewness suggested by Neuberger (2012) can be replicated exactly by a positioning in a series of out-of-the-money (OTM) options, so long as the asset price follows a martingale process. This is of particular importance, since recent literature in the context of variance swap markets (see for example Du and Kapadia, 2012), suggests that the standard variance of log-returns can be severely underestimated by the widely used implied variance formula of Britten-Jones and Neuberger (2000) in the presence of jumps in the price process. Second, unlike Feunou et al. (2014) we use option prices to extract skewness coefficients and not third central moments. This is important, since the skewness coefficient represents the third central moment standardized by the second central moment and therefore isolates the tail component of the distribution. In contrast, the simple third central moment is usually highly correlated with the second central moment (variance) and hence cannot offer additional predictive power.[1] Using this standardized skewness measure, we explicitly capture investors' fears about large negative jumps and explore their information content.

At the heart of the forward variance and skewness coefficient estimation lies the *aggregation property* suggested by Neuberger (2012). The aggregation property specifies that a quantity measured over a time interval $[0, t]$ has a high-frequency, realized counterpart that serves as its unbiased estimate. Moreover, it has an implied counterpart that serves as its unbiased estimate under the risk-neutral measure and in the absence of any risk premia. Those relations only require that prices are martingales and hold even with discrete sampling. Neuberger (2012) proposes

[1] The correlation between 1-month implied variance and 1-month implied third moment is close to -0.95 in our sample.

alternative definitions of variance and skewness that satisfy the aggregation property and can also be replicated by a portfolio of OTM options. Intuitively, this means that for any time interval $0 \leq u \leq t$ Neuberger's (2012) variance and skewness over $[0, t]$ are equal to the sum of the respective variance and skewness measures over the periods $[0, u]$ and $[u, t]$. Hence, it is possible to extract at time 0 forward moments under the risk-neutral measure for the period $[u, t]$ using the respective risk-neutral moments spanning the periods $[0, t]$ and $[0, u]$. In contrast, the standard variance and skewness definitions used by Bakshi, Kapadia and Madan (2003) do not satisfy the aggregation property and therefore the respective implied moments cannot be used for accurate forward moments estimation. Given the above, in this chapter we use Neuberger's (2012) method and S&P 500 index option prices to create a term structure of implied variance and third moment for horizons of one to four months ahead. Then, we make use of the aggregation property and extract the respective forward 1-month variances and skewness coefficients. This is the first study that creates forward stock market skewness coefficients and examines their information content over and above that of forward variances. More specifically, we investigate their predictive power for a wide range of macroeconomic variables, for stock market returns and for measures of systemic risk, crash risk and uncertainty.

The main aim of the study is to examine whether taking into consideration the information embedded in the term structure of each risk-neutral moment[2] as a whole, is important for forecasting purposes or not. Therefore, our inferences are mainly based on Wald tests of joint significance of the parameters showing whether the predictive ability of a model rises when an additional set of variables is considered. Due to overlapping observations, statistical inference mainly relies on the Newey-West (1987) covariance matrix estimator but results based on the Hodrick (1992) covariance matrix estimator are also provided for robustness purposes. In the majority of the cases the results are qualitatively very similar with the two approaches. We examine the predictive power of the estimated forward moments for the macroecon-

[2]The terms "risk-neutral moment" and "implied moment" are used interchangeably in this chapter.

omy by considering two main categories of variables. The first category is related to real activity, while the second category is related to money, credit and treasury yields. For a short-horizon predictability of one month ahead, we find that forward skewness coefficients have additional predictive power over forward variances for six (four) out of ten real activity variables examined when Newey-West (Hodrick) standard errors are used. This effect is persistent across horizons up to twelve months ahead. Moreover, the increase in adjusted R^2 across horizons, when the group of forward skewness coefficients is included into the predictive model, follows a clearly upward-trending pattern for the majority of the variables. Moving to the second category of variables, the results show that for the 1-month forecasting horizon forward skewness coefficients exhibit statistically significant forecasting ability for five (four) out of nine variables when Newey-West (Hodrick) standard errors are used, three of which are treasury yields. When we increase the forecasting horizon the empirical evidence is somehow mixed, since the predictability of forward skewness coefficients for treasury yields vanishes in the case of Newey-West standard errors, but remains intact for the longest horizon in the case of Hodrick standard errors. Moreover, in the case of Newey-West (Hodrick) standard errors forward skewness coefficients are jointly significant for only three (seven) out of nine variables. Nonetheless, the change in adjusted R^2 across horizons when the augmented predictive model is considered again provides an upward-sloping pattern for all but the treasury yield variables.

The empirical evidence regarding future stock market excess returns is also important. Our results based on Newey-West standard errors indicate that forward skewness coefficients encompass important information about future market returns over and above that provided by forward variances. In particular, forward skewness coefficients are jointly significant for 3-, 6- and 9-month ahead forecasting horizons. Furthermore, a plot of the change in adjusted R^2 across horizons when augmenting the predictive model with forward skewness coefficients has a hump-shaped pattern taking its maximum value at the 4-month horizon. It needs to be mentioned, however, that the joint significance of forward skewness coefficients is to a large extent lost when Hodrick standard errors are considered.

Next, we consider the forecasting ability of the estimated forward moments for systemic risk, tail (crash) risk, equity uncertainty and economic policy uncertainty. Systemic risk reflects the aggregate risk exposure of all financial institutions, while tail risk refers to the risk of extremely negative aggregate stock returns. Equity uncertainty and economic policy uncertainty refer to uncertainty about the stock market and uncertainty about the fiscal, monetary and regulatory policy respectively. In this context, uncertainty refers to a mixture of investors' perceived risk and ambiguity about the future stock market returns and government policies. Since we do not deal with overlapping observations in these cases, statistical inference is only based on the Newey-West (1987) covariance matrix estimator. Forward skewness coefficients are found to significantly increase the predictive power of a model forecasting systemic risk for one up to six months ahead, with the effect being stronger for the 4- and 5-month horizons. The results for tail risk show that the group of forward skewness coefficients is significant mainly for the 2- and 3-month horizons but its explanatory power is less prominent than in the case of the systemic risk. Regarding the two uncertainty measures, forward skewness coefficients can significantly improve the predictive power of a model forecasting equity uncertainty three to six months ahead, exhibiting the highest significance at the 3-month horizon. In contrast, they have limited forecasting power for economic policy uncertainty where they are relevant only for a short 1-month ahead horizon. Furthermore, we observe that the plots of the change in adjusted R^2 across horizons for systemic risk and equity uncertainty have a hump-shaped pattern with its peak at four months, very similar to the one observed in the case of stock market returns.

The contributions of this study regarding the information embedded in option prices are twofold. First, we create forward variances using a technique that accurately accounts for the presence of jumps in the price process and evaluate their information content. Using our alternative method and considering an extended sample period, we corroborate to a large extent the results presented by BPS. In particular, we also find that forward variances can predict real activity and stock market returns, while our evidence regarding their predictability for treasury yields is in line with the BPS evidence about a significant relationship with future treasury bill returns.[3] Second, we explicitly model investors' fears about negative realizations in the stock market by creating forward skewness coefficients. We find that their predictive power for a large set of macroeconomic variables is highly significant and becomes stronger for most of the predicted variables once longer forecasting horizons are considered. Moreover, forward skewness coefficients exhibit significant forecasting abilty for stock market returns, systemic risk and equity market uncertainty especially for a horizon of four months ahead that matches the time period spanned by the estimated forward skewness.

The remainder of the chapter is structured as follows. Section 4.2 provides an overview of the related literature. Section 4.3 describes the theory behind the forward moments estimation, while Section 4.4 analyzes the data and the variables used in the study. Section 4.5 provides the empirical results and finally Section 4.6 concludes.

4.2 Related Literature

This chapter is similar in spirit to the studies of BPS, Luo and Zhang (2012), Mueller et al. (2013) and Feunou et al. (2014) who make use of the term structure of the second and third risk-neutral central moments in the equity and treasury markets for forecasting purposes. This is the first study, however, that investigates the information content of the term structure of the equity market risk-neutral skewness coefficients. Therefore, unlike Feunou et al. (2014) who consider the term structure of the risk-neutral third central moments,[4] we show that the term structure

[3] Some discrepancies between our results and the BPS results can be attributed to the different method employed and the different sample period examined, since by following the method of BPS and restricting our analysis to their sample period we can replicate their results almost exactly.

[4] As described in the previous section the risk-neutral third central moment is highly related to the risk neutral second central moment (variance) and therefore cannot adequately account for the

of the risk-neutral skewness coefficients encapsulates important information that is not embedded in the term structure of the risk-neutral variances.

Since forward skewness coefficients are found to exhibit significant forecasting power for future market returns, this chapter further contributes to a broader strand of the literature which investigates the information content of option prices for future equity returns and crashes at an aggregate level. More specifically, a recent stream of papers focuses on the variance risk premium, i.e. the difference between expected variance under the risk-neutral measure and expected variance under the physical measure. Bollerslev et al. (2009) are the first to use this measure for predictive purposes and show that the S&P 500 index variance risk premium is strongly positively related to future market returns especially at a quarterly horizon. Similar empirical evidence is also reported by Drechsler and Yaron (2011) and Bollerslev et al. (2012). Mueller et al. (2011) and Zhou (2012) elaborate to the predictability of the S&P 500 index variance risk premium by showing that it has a strong positive impact on future bond returns and credit spreads.

Moreover, several recent studies investigate market return predictability using alternative measures extracted from the prices of S&P 500 index options. In particular, Du and Kapadia (2012) illustrate theoretically that the difference between the Bakshi, Kapadia and Madan (2003) implied variance and the squared VIX captures the jump component of the quadratic variation and find that their implied jump index is positively related to future market returns. Karoui (2012) suggests a novel approach for estimating an option-implied equity premium and provides evidence that his measure significantly predicts stock market returns. Vilkov and Xiao (2013) create a tail loss measure from put option prices and show that it is associated with a positive market risk premium. Driessen et al. (2013) construct an implied correlation index and show that it is a strong predictor of future market returns, even when controlling for the variance risk premium. Atilgan et al. (2014) find that there is a positive relation between the slope of the implied volatility smirk and subsequent market returns. Golez (2014) estimates an option-implied dividend growth

tail of the distribution.

rate that is a strong predictor of future dividend growth and creates an amended dividend-price ratio for predicting future market returns.

As far as the predictability of jumps is concerned, Doran et al. (2007) provide evidence about the forecasting power of the slope of the S&P 100 index implied volatility smirk for future underlying asset jumps. In particular, they show that the slope of the puts' volatility smirk is significantly related to the probability of a negative jump, while the slope of the calls' volatility smirk is significantly related to the probability of a positive jump. Vilkov and Xiao (2013) show that their tail loss measure exhibits some forecasting power for future market crashes.

Finally, since forward skewness coefficients are shown to be important for predicting a number of macroeconomic variables, this study also complements a strand of the literature which - in the majority of the cases - shows that option-implied measures can be successfully used for forecasting future macroeconomic conditions. Lynch and Panigirtzoglou (2008) use futures options on the S&P 500 index, FTSE 100 index, eurodollar and short sterling to extract the respective risk-neutral moments but find limited evidence in favor of the hypothesis that there is significant predictability for macroeconomic variables such as industrial production and investment growth in the US and the UK. In contrast, Bekaert and Hoerova (2013) show that VIX forecasts negative growth in industrial production, an effect that stems from the conditional variance component and not the variance risk premium component. Moreover, Bekaert et al. (2013) find out that VIX is negatively related to future real interest rate for horizons longer than a year ahead, an effect which is attributed to both the variance risk premium and the conditional variance component. Similarly, David and Veronesi (2014) report that the S&P 500 index at-the-money implied volatility has a negative effect on the future short-term interest rate, while the opposite is true for the steepness of the S&P 500 index implied volatility smirk. In a slightly different context, Sarantopoulou-Chiourea and Skiadopoulos (2014) show that a relative risk aversion coefficient extracted from S&P 500 index risk-neutral moments is positively related to future real activity.

4.3 Forward Moments and the Aggregation Property

BPS rely on the theoretical foundations of Carr and Lee (2009) in order to price exponential claims on quadratic variation under the assumption of no jumps in the underlying asset price process. Under this assumption, the quadratic variation over an interval $[0, t]$ is equal to the integrated variance over $[0, t]$. BPS demonstrate that the exponential claims on integrated variance of log-returns can be replicated by a positioning in a series of OTM options.[5] Therefore, they construct measures of forward integrated variance the same way discount bond prices are used to provide forward treasury yields.

The main disadvantage of this method lies in the assumption that the underlying asset price follows a pure diffusion process. In particular, recent studies that investigate the impact of jumps in the context of the variance swaps (Broadie and Jain, 2008, Du and Kapadia, 2012, Rompolis and Tzavalis, 2013 and Bondarenko, 2014), find that the risk-neutral variance of Carr and Madan (1998), Demeterfi et al. (1999) and Britten-Jones and Neuberger (2000) can lead to substantial underestimation of the quadratic variation of log-returns in the presence of large negative jumps. For example, Du and Kapadia (2012) show that when jumps constitute 70% of the quadratic variation, then the approximation error in terms of annualized volatility is an economically significant 1%. Therefore, since the Carr and Lee (2009) theory of claims on exponential quadratic variation only holds for pure diffusion processes, it is reasonable to assume that a similar bias can arise in the case of the BPS quadratic variation estimates as well.

In this study, we employ the newly established concept of the *aggregation property* (Neuberger, 2012) in order to create our measures of forward moments. This has two main advantages. First, our alternative variance measure can always be replicated *exactly* by equation (2.37) as long as the underlying asset price follows a martingale

[5] See equation (2.41). In fact, the prices of those exponential claims are also affected by the level of the risk-free rate.

process. Therefore, in contrast to the variance measure estimated by BPS, it is robust to the presence of jumps in the price process. Second and most important, it allows us to create forward measures of the third (standardized) moment of the asset returns as well. This way, we explicitly model investors' future crash worries and explore their information content.

Neuberger (2012) postulates that any real-valued function g of an adapted process X has the aggregation property if for any $0 \leq u \leq t$,

$$E_0\left[g\left(X_t - X_0\right)\right] = E_0\left[g\left(X_t - X_u\right)\right] + E_0\left[g\left(X_u - X_0\right)\right]. \tag{4.1}$$

Assuming that the forward asset price F is a martingale, Neuberger (2012) and Kozhan, Neuberger and Schneider (2013) define log and entropy variance respectively as:

$$G^V_{0,t} = E_0\left[\frac{F_t}{F_0} - 1 - \ln\left(\frac{F_t}{F_0}\right)\right], \tag{4.2}$$

$$G^E_{0,t} = E_0\left[2\left(\frac{F_t}{F_0}\ln\left(\frac{F_t}{F_0}\right) - \frac{F_t}{F_0} + 1\right)\right]. \tag{4.3}$$

The functions inside the brackets have the aggregation property and converge to the second moment of returns. Intuitively, under the Black and Scholes (1973) assumptions, log variance is the implied variance of a log contract, i.e. a contract that pays $\ln\left(F_t\right)$, while entropy variance is the implied variance of an entropy contract, i.e. a contract that pays $F_t \ln\left(F_t\right)$.[6] Moreover, under the risk-neutral measure $G^V_{0,t}$ can be replicated exactly by equation (2.37). Therefore, following Neuberger (2012) and Bondarenko (2014), we adopt the proposed alternative definition of variance shown in equation (4.2) and estimate expected quadratic variation under the risk-neutral measure by employing equation (2.37). In this case, we regard $G^V_{0,t}$ as the implied variance of stock returns, i.e. $G^V_{0,t} = IV_{0,t}$.

Similarly, skewness is alternatively defined by Neuberger (2012) and Kozhan,

[6] In this case the term *entropy* is used due the similarity of the payoff of the contract with entropy as used in thermodynamics and information theory.

Neuberger and Schneider (2013) as:

$$G_{0,t}^{S} = E_0\left[6\left(\frac{F_t}{F_0}\ln\left(\frac{F_t}{F_0}\right) - 2\frac{F_t}{F_0} + \ln\left(\frac{F_t}{F_0}\right) + 2\right)\right], \tag{4.4}$$

where the function inside the brackets has the aggregation property and converges to the third moment of returns. $G_{0,t}^{S}$ can be written as the difference between the two previously described variance measures and under the risk neutral measure can be replicated by equation (2.38). Thus, $G_{0,t}^{S}$ is regarded as the implied third moment of stock returns, i.e. $G_{0,t}^{S} = TM_{0,t}$. Both implied variance and skewness are unbiased estimates of the true variance and skewness in the absence of any risk premia. From equation (4.1) we can write for any $0 \leq u \leq t$:

$$IV_{0,t} = IV_{0,u} + E_0\left[G_{u,t}^{V}\right], \tag{4.5}$$

$$TM_{0,t} = TM_{0,u} + E_0\left[G_{u,t}^{S}\right]. \tag{4.6}$$

Rearranging equations (4.5)-(4.6) we get:

$$FV_{0;u,t} \equiv E_0\left[G_{u,t}^{V}\right] = IV_{0,t} - IV_{0,u}, \tag{4.7}$$

$$FS_{0;u,t} \equiv E_0\left[G_{u,t}^{S}\right] = TM_{0,t} - TM_{0,u}. \tag{4.8}$$

where $FV_{0;u,t}$ and $FS_{0;u,t}$ are the time 0 forward variance and third moment respectively for the period u to t implied by the prices of OTM options at time 0.

In the subsequent analysis we are interested in the forward skewness coefficient. Therefore, we estimate:[7]

$$FSC_{0;u,t} = \frac{FS_{0;u,t}}{\left(FV_{0;u,t}\right)^{\frac{3}{2}}}. \tag{4.10}$$

[7]The forward skewness coefficient is subject to a small convexity bias due to Jensen's inequality. In particular,

$$E_0\left[\left(G_{u,t}^{V}\right)^{\frac{3}{2}}\right] \geq \left(E_0\left[G_{u,t}^{V}\right]\right)^{\frac{3}{2}} = \left(FV_{0;u,t}\right)^{\frac{3}{2}}. \tag{4.9}$$

4.4 Data and Variables

4.4.1 Options data and forward moments estimation

We obtain daily S&P 500 index call and put options data from IVolatility.com for the period 1996:01 to 2012:12. Following the standard practice, option prices are calculated as the midpoint between the best bid and best ask price. Expiration time is calculated assuming 360 calendar days per year. Each trading day is matched with the respective dividend yield which is obtained from Bloomberg. Moreover, each option contract is matched with the appropriate continuous risk-free rate that is found after interpolating the 1-, 3-, 6- and 12-month Treasury Constant Maturity rates downloaded from the FRED database of the Federal Reserve Bank of St. Louis.

A series of filtering rules are applied to the dataset to eliminate measurement errors and outliers mainly caused by thinly traded options (see for example Aït-Sahalia and Lo, 1998, Han, 2008 and Chang et al., 2013). First, we discard options that do not satisfy standard no-arbitrage conditions. Second, we exclude observations with zero bid prices and midpoint prices that are less than \$3/8. Third we filter out options with zero or higher than 1 implied volatility. Finally, we take into consideration only options with non-zero trading volume and maturity between 7 and 270 calendar days.

We use equations (2.37) and (2.38) to estimate implied variance and skewness for constant maturities of 30-, 60-, 90- and 120-days ahead at the end of each month.[8] Since interpolation across the time dimension is needed for this exercise, we make sure that we consider only days with a sufficient number of available maturities. A maturity is regarded as available if it has a cross-section with at least two OTM puts and two OTM call options. Therefore, we require that there are at least four available maturities that cover the next two months and either the third or fourth month (or both) following the current month. Moreover, we do not take into consideration

[8]We do not estimate implied moments for maturities longer than 120 days since the availability of long maturity options is not high enough to provide accurate estimates of long-maturity implied moments.

days that do not have available at least one maturity shorter than or equal to 30 days and at least one maturity longer than or equal to 120 days.[9]

In order to create costant maturity implied moments, we follow the interpolation technique of Kostakis et al. (2011) and Neumann and Skiadopoulos (2013). In particular for each cross-section of options, we interpolate across implied volatilites in the delta space to obtain a grid of 1000 data points with deltas ranging from 0.01 to 0.99. Inside the available delta range we interpolate using a cubic smoothing spline with smoothing parameter 0.99 while outside the available delta range, we extrapolate using the respective boundary values. The interpolation across the time dimension for a given day proceeds as follows: First, from all the available interpolated implied volatility curves of a given day we keep the data points with delta values of 0.1, 0.2,...,0.9. Using a cubic smoothing spline, we then interpolate across the time dimension for the given constant maturities. Second, we create constant maturity implied volatility curves by fitting a cubic spline to the available nine implied volatilities. Third, the delta grid of the constant maturity implied volatility curve is converted to strike prices and the respective implied volatilities are transformed to option prices. Finally, equations (2.37) and (2.38) are discretized and estimated using the trapezoidal approximation.

Once we have the estimates of constant maturity implied moments for 30-, 60-, 90- and 120-days ahead, we use equations (4.7) and (4.8) to create vectors of forward 1-month moments. In particular we create:

$$\begin{aligned}\mathbf{fv}_0 &\equiv \left[FV_0^{(1)}\ FV_0^{(2)}\ FV_0^{(3)}\ FV_0^{(4)}\right]' \\ &\equiv \left[FV_{0;0,30}\ FV_{0;30,60}\ FV_{0;60,90}\ FV_{0;90,120}\right]', \end{aligned} \tag{4.11}$$

$$\begin{aligned}\mathbf{fs}_0 &\equiv \left[FS_0^{(1)}\ FS_0^{(2)}\ FS_0^{(3)}\ FS_0^{(4)}\right]' \\ &\equiv \left[FS_{0;0,30}\ FS_{0;30,60}\ FS_{0;60,90}\ FS_{0;90,120}\right]'. \end{aligned} \tag{4.12}$$

[9]In fact, our sample is restricted to the 1996:01-2012:12 period because of the relatively limited availability of option maturities in the pro-1996 period.

Then using equation (4.10) we create a vector of forward 1-month skewness coefficients:

$$\begin{aligned}\mathbf{fsc}_0 &\equiv \left[FSC_0^{(1)}\ FSC_0^{(2)}\ FSC_0^{(3)}\ FSC_0^{(4)}\right]' \\ &\equiv \left[FSC_{0;0,30}\ FSC_{0;30,60}\ FSC_{0;60,90}\ FSC_{0;90,120}\right]'. \end{aligned} \tag{4.13}$$

Table 4.1 reports the descriptive statistics for the estimated forward variances and skewness coefficients. All forward variances exhibit very similar statistics and their autocorrelations range from 0.775 to 0.834. In contrast, forward skewness coefficients become more negative and volatile as the horizon increases. Moreover, forward skewness coefficients are much less persistent with autocorrelation coefficients ranging from 0.300 to 0.548. Table 4.2 provides the correlation coefficients for the forward moments. Forward variances are all positively and highly correlated with correlation coefficients ranging from 85% to 94%. The respective correlations between forward skewness coefficients range from 37% to 63%. It is apparent that while each forward variance has an idiosyncratic component depending on the month it refers to, all of them share a strong common component. On the contrary, the idiosyncratic information embedded in each forward skewness coefficient is more pronounced. This can be also confirmed by looking at Figure 4.1 which plots the forward moments across time. Forward variances tend to move in lockstep, taking their highest values during the recent financial crisis. Forward skewness coefficients exhibit similar patterns but the idiosyncratic variation of each variable is evident. The correlations between forward variances and skewness coefficients are low and consistently negative apart from the case of $FV^{(1)}$ and $FSC^{(1)}$ whose correlation is positive but very close to zero (2%).

4.4.2 Forecasted and control variables

The predictive power of the estimated forward moments will be investigated in respect to three main aspects of the economy, a) macroeconomic environment, b) stock market and c) risk and uncertainty. The macroeconomic variables can be further categorized into two main groups, real activity variables and money, credit and treasury yield variables. The real activity variables consist of personal income (Pers income), industrial production (Ind prod), capacity utilization (Cap util), unemployment level (Unempl), nonfarm payroll (Payroll), housing starts (House starts), housing authorized (Build perm), manufacturing and trade inventories (M&T invent), real personal consumption expenditures (Consumption), and retail sales (Retail sales). These variables reflect the main aspects of real activity in an economy, as they capture the total productivity, the labour market, the housing sector and the total sales and consumption. The credit and treasury yield variables consist of money supply M1 (M1), real money supply M2 (M2 (real)), total reserves of depository institutions (Reserves tot), commercial and industrial loans (C&I loans), consumer price index (CPI), 3-month T-bill rate (3-m t-bill), 6-month T-bill rate (6-m t-bill), 1 year T-bond rate (1-yr t-bond) and 5 years T-bond rate (5-yr t-bond). These variables illustrate the key aspects of an economy's credit capacity as well as the closely related money stock and inflation levels. For the purposes of the predictive analysis, we construct monthly logarithmic growth rates for all the variables apart from Cap util, Unempl and the four interest rates for which we estimate monthly changes, since they are expressed in percentage terms. A detailed description of the macroeconomic dataset can be found in Table C.1 of Appendix C.

The stock market is represented by the excess return of the value-weighted index from the Chicago Center for Research in Security Prices (CRSP). We define excess market return as the difference between the monthly log-return of the CRSP value-weighted index and the 1-month Treasury bill rate obtained from Kenneth French's website.[10]

The risk variables we consider, reflect financial systemic risk and tail (or crash) risk. Systemic risk refers to the aggregate risk taken by financial institutions. Such type of risk is of particular importance since a failure of a financial institution during periods of high systemic risk can cause severe instability to the overall economy

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

(see for example Allen et al., 2012 and Brownlees and Engle, 2012). Systemic risk is proxied by the Catfin measure suggested by Allen et al. (2012) which aggregates the estimates of three different VaR methodologies for the monthly returns of all the available financial firms. The data on systemic risk can be found on Turan Bali's website.[11] Tail risk refers to the risk of extremely negative realizations in the aggregate stock market. This type of risk is important not only due to investors' natural aversion to extreme negative returns but also because increased crash risk is typically related to adverse macroeconomic shocks (Kelly and Jiang, 2014). Monthly market tail risk is constructed by applying Hill's (1975) power law estimator to the daily returns of all the available stocks in a given month, as proposed by Kelly and Jiang (2014). Finally, we also examine two forms of uncertainty namely equity uncertainty and economic policy uncertainty. Equity uncertainty refers to the uncertainty that is present in the stock market, while economic policy uncertainty refers to uncertainty about fiscal, monetary and regulatory policy. In this context uncertainty stands for a mixture of both risk and ambiguity, i.e. it encapsulates events with unknown outcomes whose probability measures are assumed to be both known (risk) and unknown (ambiguity). Baker et al. (2013) and Bloom (2014) show empirically that increased uncertainty is related to decreased concurrent and subsequent real activity. The respective data are obtained from Baker et al.'s (2013) webpage.[12] The economic policy uncertainty index consists of three components related to news coverage, federal tax code provisions set to expire in future years and economic forecaster disagreement. The equity uncertainty index is only based on news coverage. Due to their tight link with the macroeconomic conditions, the above risk and uncertainty variables have become extremely popular especially after the recent financial crisis.

As control variables, we use the yield term spread (TERM), the dividend-to-price ratio (d-p) and the earnings-to-price ratio (e-p). TERM is the difference between the 10-year bond yield and the 1-year bond yield, d-p is the difference between the

[11] `http://faculty.msb.edu/tgb27/workingpapers.html`
[12] `http://www.policyuncertainty.com`

log aggregate annual dividends and the log level of the S&P 500 index and e-p is the difference between the log aggregate annual earnings and the log level of the S&P 500 index. Data on monthly prices, dividends, and earnings are obtained from Robert Shiller's website.[13] All interest rate data are obtained from the FRED database of the Federal Reserve Bank of St. Louis.

4.5 The Information Content of Forward Skewness

The main purpose of this section is to investigate the predictive ability of forward skewness coefficients for the previously described macroeconomic variables, stock returns, risk and uncertainty variables. Throughout the empirical analysis, we proceed by first examining the forecasting power of our alternative forward variance measures when combined only with relevant control variables. Then, we augment the predictive model with the forward skewness coefficients and evaluate the increase in its explanatory power. As in Cochrane and Piazzesi (2005), Ang and Bekaert (2007) and BPS, we mainly rely on Wald tests of joint significance for the four forward variances and the four forward skewness coefficients. A Wald test of that type is identical to a J-test of overidentifying restrictions and shows whether the increase in R^2 due to the inclusion of the additional group of variables is significant or not (see the discussions in Cochrane and Piazzesi, 2005 and Cochrane, 2005). We present the p-values for those Wald tests together with the adjusted R^2 for the simple and the augmented model.

4.5.1 Forecasting macroeconomy

We examine the forecasting power of the forward moments for the macroeconomic variables for horizons of one up to twelve months ahead. In particular, for each

[13] `http://www.econ.yale.edu/~shiller/data.htm`

macroeconomic variable we run regressions of the form:

$$y_{i;t+h} = \alpha_{i;h} + \boldsymbol{\beta}'_{i;h}\mathbf{z}_t + \varepsilon_{i;t+h,h}, \tag{4.14}$$

where $y_{i;t+h} = \left(\frac{12}{h}\right)[y_{i;t+1} + y_{i;t+2} + ... + y_{i;t+h}]$ is the annualized h-month growth or change in variable i and $\mathbf{z}_t$ is the vector of predictive variables for each of the two models considered. The regression analysis covers the period 1996:01-2012:12 and for each forecasting horizon we lose h observations. Under the null of no predictability the overlapping nature of the data imposes an $MA\,(h-1)$ structure to the error term $\varepsilon_{t+h,h}$ process. To overcome this problem we base our statistical inference on both Newey and West (1987) and Hodrick (1992) standard errors with lag length equal to the forecasting horizon. In general, the Hodrick (1992) standard errors tend to be more conservative, especially in long horizons when the null of no predictability is true (Ang and Bekaert, 2007) but have lower statistical power when the null is false (Bollerslev, Marrone, Xu and Zhou, 2012). Following BPS, we include TERM as a control variable in all the regressions. The beta coefficients reported in the subsequent tables have been scaled and can be interpreted as the annualized percentage growth (or change in percentage terms for Cap util, Unempl and the interest rates) in the forecasted variables from a one standard deviation change in each regressor.

4.5.1.1 Real activity

Table 4.3 presents the results for the 1-month growth (or change) in the real activity variables when Newey-West standard errors are employed. First, it is apparent that the results for the forward variances hardly change when the predictive model is augmented with the forward skewness coefficients. This is reasonable as we have already seen that the correlations between the two groups of forward moments are rather low. An increased (decreased) $FV^{(1)}$ is related to a subsequent decline (improvement) in real activity with the relationship being statistically significant mainly for House starts, Build perm and Retail sales. From the rest of the forward variances, $FV^{(3)}$

is negatively and significantly related to Ind prod and $FV^{(4)}$ exhibits a positive and significant relationship with Ind prod and Cap util. While individual forward variances are not particularly significant across real activity variables, we have to keep in mind that due to the high cross-correlations among forward variances, it is difficult to estimate the respective regression coefficients with a small confidence interval.[14] From the forward skewness coefficients, a less (more) negative value of $FSC^{(1)}$ or $FSC^{(4)}$ is associated with a decline (improvement) in real activity, while a less (more) negative value of $FSC^{(3)}$ is associated with improved (declined) real activity. These relationships are stronger for Ind prod, Unempl, Payroll and M&T invent. Additionally, $FSC^{(3)}$ is also significant for Build perm and $FSC^{(4)}$ has a significant effect on Cap util, Consumption and Retail sales. The results for TERM are mixed, since while it is negatively and significantly related to Pers income, Payroll, M&T invent and Consumption, it is positively and significantly related to Cap util. With respect to economic sensitivity, forward variances have in general a higher impact on real activity variables than forward skewness coefficients. However, the impact of skewness is far from negligible in several cases. For example, a one standard deviation increase in $FSC^{(4)}$ causes an annualized monthly drop of 1.831% in Retail sales.

Turning to the Wald tests of joint significance, forward variances are strongly jointly significant for all the real activity variables apart from Consumption and Retail sales. Most importantly, the results for the augmented model show that the forward skewness coefficients are also jointly significant at the 1% level for Ind prod, Unempl, Payroll and M&T invent and at the 10% level for Cap util and Retail sales. These results imply that for six out of ten real activity variables the increase in explanatory power when considering the augmented model is statistically significant. Finally, the adjusted R^2 increases in the case of the augmented model for eight out

[14]Orthogonalizing $FV^{(2)}$, $FV^{(3)}$ and $FV^{(4)}$ with respect to $FV^{(1)}$, provides a strong individual significance for $FV^{(1)}$ across almost all the variables. These results are not presented here, since our main focus is on joint significance of forward variances and not on individual significance of each variance. The results of joint significance remain unaltered, whether we use the original forward variances or the orthogonalized ones.

of ten variables.

The results for the 1-month growth (or change) in the real activity variables when Hodrick standard errors are used, are reported in Table 4.4. Individual coefficient results are similar to those reported in Table 4.3, although relatively weaker in principle. Moreover, the joint significance of forward variances is much weaker across variables. The joint significance of forward skewness coefficients, however, is very similar to that presented in Table 4.3 even though it is lost in two marginal cases (Cap util and Retail sales).

Table 4.5 reports the results for the 6-month horizon (Panel A) and 12-month horizon (Panel B) predictive regressions when Newey-West standard errors are considered. To save space, only the augmented model has been reported and the individual coefficients for the forward variances and TERM have been omitted. However, the full set of results can be found in Tables C.2-C.3 of Appendix C. In principle, $FV^{(1)}$ and $FV^{(4)}$ remain significant at 6- and 12-month horizons with an effect similar to the one discussed above. Moreover, TERM clearly indicates improved real activity for the 12-month horizon as in Chen (1991) and Estrella and Hardouvelis (1991). Regarding the forward skewness coefficients, similarly to the case of the 1-month predictability $FSC^{(1)}$ and $FSC^{(4)}$ are consistently related to a decreased real activity across horizons and $FSC^{(3)}$ is consistently related to an improved real activity across horizons. Furthermore, the effect of the forward skewness coefficients appears to be stronger in almost all the cases as horizon increases from one to six months ahead and remains at similar levels when the horizon increases to twelve months.

The joint significance of the forward variances is improved as the forecasting horizon becomes longer. At the 12-month horizon, forward variances are jointly significant for all the real activity variables and for eight of them the significance is at the 1% level. Forward skewness coefficients continue to be jointly significant for Ind prod, Payroll and M&T invent at both 6- and 12-month horizons, while they remain jointly significant for Unempl at the 6-month horizon. Moreover, forward skewness coefficients become jointly significant for Pers income and Consumption at both long horizons examined. Finally, it is remarkable that for both horizons the adjusted R^2 of the augmented model is always higher than the respective adjusted R^2 of the simple model.

Table 4.6 provides the results for the 6-month horizon (Panel A) and 12-month horizon (Panel B) predictive regressions in the case of Hodrick standard errors. As in Table 4.5 only the augmented model has been reported and the individual coefficients for the forward variances and TERM have been omitted. However, the full set of results can be found in Tables C.6-C.7 of Appendix C. The individual coefficient results are qualitative very similar to those reported in Table 4.5. With respect to the Wald tests for forward variances, the main difference occurs for the two housing variables and Retail sales since the joint significance is lost. With respect to the Wald tests for forward skewness coefficients, the results are qualitatively similar and overall slightly stronger when Hodrick standard errors are utilized.

Figure 4.2 provides a clear picture of the importance of the forward skewness coefficients for predicting real activity especially at long horizons. In particular, for each variable it plots the change in adjusted R^2 when considering the augmented instead of the simple model across different forecasting horizons. Apart from the case of Cap util and Unempl all the other graphs show a clear upward sloping pattern, which implies that taking into consideration forward skewness coefficients becomes even more important as forecasting horizon increases up to twelve months ahead.

Overall, forward skewness coefficients appear to have significant predictive power over and above forward variances for the majority of the real activity variables considered and especially for Ind prod, Unempl, Payroll, M&T invent and Consumption. Moreover, in most of the cases the effect is stronger when long-horizon predictability is examined.

4.5.1.2 Money, credit and treasury yields

Table 4.7 presents the results for the 1-month growth (or change) in the money, credit and treasury yield variables when Newey-West standard errors are employed. The results for forward variances are very similar for both the simple and the augmented model. $FV^{(1)}$ exhibits the strongest statistical significance with a higher (lower) value being related to increased (decreased) M2 (real), lower (higher) CPI and lower (higher) interest rates. $FV^{(2)}$ is positively and significantly related to 1-yr t-bond and positively but weakly related to 5-yr t-bond. With regard to forward skewness coefficients, $FSC^{(3)}$ exhibits a consistent pattern, being positively and significantly associated with all treasury yields and negatively and significantly associated with the money supply variables. Moreover, $FSC^{(2)}$ has a significant negative relationship with C&I loans, while $FSC^{(4)}$ has a significant positive impact on 5-yr t-bond. TERM is significantly related, positively and negatively respectively, only to M1 and C&I loans. With respect to economic sensitivity, as in the previous section, forward variances have, on average, a higher effect on the forecasted variables than forward skewness coefficients. However, the economic impact of forward skewness coefficients is not negligible either. For example, a one standard deviation increase in $FSC^{(3)}$ is followed by an annualized monthly increase of about 40 basis points in the treasury yields.

The results of the Wald tests show that forward variances are jointly strongly significant for M2 (real), Reserves tot, CPI and all the treasury yield variables. Forward skewness coefficients are jointly significant for M2 (real), C&I loans and three out of four treasury yields (6-m t-bill, 1-yr t-bond and 5-yr t-bond). These results show that the increase in explanatory power stemming from the addition of forward skewness coefficients into the predictive model is significant for five out of nine variables. Furthermore, the augmented model is accompanied by an increase in the adjusted R^2 for six out of nine variables.

Table 4.8 shows the the results for the 1-month growth (or change) in the money, credit and treasury yield variables when Hodrick standard errors are employed.

Although slightly less significant in general, individual coefficient results are similar to those reported in Table 4.7. Forward variances lose their joint significance in four out of seven cases. Forward skewness coefficients, however, remain jointly significant for all variables but M2 (real).

Table 4.9 reports the results for 6-month (Panel A) and 12-month horizon (Panel B) predictive regressions in the case of Newey-West standard errors. As in the previous section, only the augmented model has been reported and the individual coefficients for the forward variances and TERM have been omitted but can be found in Tables C.4-C.5 of Appendix C. For those variables the pattern is similar to the 1-month horizon, apart from the fact that $FV^{(3)}$ becomes also significant for M2 (real), CPI and 5-yr t-bond with an effect opposite to that of $FV^{(1)}$. Moreover, at the longest 12-month horizon, TERM exhibits some explanatory power for future treasury yield changes, a finding which is in the spirit of Fama (1990). However, this positive impact is significant only at the 10% level. With regard to forward skewness coefficients, $FSC^{(1)}$ has a strong negative effect on C&I loans at both the 6- and 12-month horizons, while $FSC^{(2)}$ is significantly related, positively and negatively respectively, to the money supply variables and CPI at the 12-month horizon. $FSC^{(3)}$ continues to have a significantly negative impact on M1 but its impact on the treasury yields diminishes as the horizon increases. Finally, $FSC^{(4)}$ is negatively related to C&I loans and CPI, with the effect becoming stronger across horizons.

The Wald tests indicate that an increase in the forecasting horizon eliminates the significant effect of forward variances on the treasury yields but strengthens their joint effect on all the other variables. Similarly, forward skewness coefficients are not jointly significant for treasury yields when 6- and 12-month horizons are considered. At the 12-month horizon, however, they are strongly jointly significant for M2 (real), C&I loans and CPI. Furthermore, the adjusted R^2 of the augmented model is higher than the adjusted R^2 of the simple model for six out of nine variables when considering the 6-month horizon and for eight out of nine variables when considering the 12-month horizon. However, in the case of the treasury yields the existing increases are only marginal.

The results for the 6-month horizon (Panel A) and 12-month horizon (Panel B) predictive regressions in the case of Hodrick standard errors are shown in Table 4.10. Similarly to Table 4.9 only the augmented model has been reported and the individual coefficients for the forward variances and TERM have been omitted. However, the full set of results can be found in Tables C.8-C.9 of Appendix C. The results of joint significance for forward variances are generally weaker for the money and creadit variables but stronger for the treasury yield variables than those reported in Table 4.9. The results of joint significance for forward skewness coefficients are similar to those presented in Table 4.9, with the main difference being that at the 12-month horizon forward skewness coefficients appear to be significant for 3-m t-bill, 6-m t-bill and 1-yr t-bond.

The above results are also depicted in Figure 4.3 which plots the change in adjusted R^2 across different forecasting horizons when adding the forward skewness coefficients into the predictive models. C&I loans and CPI exhibit an explicit upward sloping pattern similar to that found for most of the real activity variables. M1, M2 (real) and Reserves tot provide less steep but still upward trending patterns. The treasury yields, on the other hand, provide patterns that are flat and close to zero, with only a slight increase for horizons up to two months ahead.

In summary, forward skewness coefficients appear to have significant predictive power over and above forward variances mainly for M2 (real), C&I loans and CPI, with the effect being in general stronger for long horizons. They are also important for explaining treasury yield movements but mainly for a short 1-month horizon.

4.5.2 Forecasting stock market

Having established that taking into consideration forward skewness coefficients is important for forecasting macroeconomic variables, we now turn our attention to the stock market. In particular, for horizons of one up to twelve months ahead we run regressions of the following form:

$$re_{t+h} = \alpha_h + \boldsymbol{\beta}_h^{'} \mathbf{z}_t + \varepsilon_{t+h,h}, \tag{4.15}$$

where $re_{t+h} = \left(\frac{12}{h}\right)[re_{t+1} + re_{t+2} + ... + re_{t+h}]$ is the annualized h-month excess return of the CRSP value-weighted index and $\mathbf{z}_t$ is the vector of predictive variables for each of the two models considered. The regression analysis covers the period 1996:01-2012:12 and for each forecasting horizon we lose h observations. Under the null of no predictability the overlapping nature of the data imposes an $MA\,(h-1)$ structure to the error term $\varepsilon_{t+h,h}$ process. To tackle this problem we base our statistical inference on both Newey and West (1987) and Hodrick (1992) standard errors with lag length equal to the forecasting horizon. In general, the Hodrick (1992) standard errors tend to be more conservative, especially in long horizons when the null of no predictability is true (Ang and Bekaert, 2007) but have lower statistical power when the null is false (Bollerslev, Marrone, Xu and Zhou, 2012). Motivated by prior literature (see for example, Fama and French, 1988, Campbell and Shiller, 1988a,b, Lamont, 1998 and Goyal and Welch, 2008, among others) we include d-p and e-p as control variables.[15] The beta coefficients reported in the subsequent tables have been scaled and can be interpreted as the percentage annualized excess market returns caused by a one standard deviation change in each regressor.

Table 4.11 reports the results for 1-, 3-, 6-, 9- and 12-month forecasting horizons when Newey-West standard errors are used. From the forward variances group, $FV^{(1)}$ is negatively related to future stock market returns but the effect is significant only when we consider the augmented model for horizons between six and twelve months ahead. $FV^{(4)}$ exhibits also some forecasting power for future market returns but only at a short 1-month horizon. Recall, however, that due to the high cross-correlations among forward variances, it is difficult to find strong individual significance for these variables. Therefore, our conclusions are mainly based on the Wald tests of joint significance. From the forward skewness coefficients group,

[15]Note that for our sample period the correlation between d-p and e-p is -0.05.

$FSC^{(3)}$ is consistently positively and significantly related to future market returns, with the effect being stronger at the 6- and 9-month horizons. Moreover, $FSC^{(4)}$ exhibits a negative and significant relationship with future market returns at the 3-month horizon. Recall from Section 4.5.1.1 that $FSC^{(3)}$ is positively related to real activity while $FSC^{(4)}$ is negatively related to real activity. Therefore, there is a consistent pattern for these two forward skewness coefficients with $FSC^{(3)}$ being related to increased economic activity and higher stock market returns and $FSC^{(4)}$ being related to reduced economic activity and lower stock market returns. Regarding the control variables, d-p is positively related to future market returns and in line with the literature its effect becomes stronger as the forecasting horizon increases. In contrast, e-p does not exhibit any significant relationship with future market returns during our sample period. In economic terms, a one standard deviation increase in $FSC^{(3)}$ results in an annualized excess market return ranging from 3.373% to 8.244% depending on the forecasting horizon considered. With the exception of the 12-month horizon, similar figures are also observed for $FSC^{(4)}$.

Moving to the Wald tests of joint significance, forward variances are jointly significant at the 5% level when the forecasting horizon is six months ahead and at the 10% level when the forecasting horizon is nine months ahead. In contrast, forward skewness coefficients are jointly significant at the 10% level for the 3-month horizon and at the 5% level for both the 6- and the 9-month horizons. Therefore, we find that forward skewness coefficients significantly forecast future market returns over and above forward variances and their effect is stronger than that of forward variances. Furthermore, the adjusted R^2 of the augmented model is higher than the adjusted R^2 of the simple model for all but the 1-month horizon. Looking at the change in adjusted R^2 across horizons depicted in Figure 4.4, we observe a hump-shaped pattern. In particular, the increase in adjusted R^2 is upward trending for short horizons, taking its maximum value at the 4-month horizon and then gradually declining for longer horizons.

The respective results for stock market return predictability, when Hodrick standard errors are used, are presented in Table 4.12. In this case, none of the forward variance individual coefficients appears to be significant, while the individual results for forward skewness coefficients are qualitatively similar - and in some cases stronger - to those presented in Table 4.11. Turning to the Wald tests, forward variances are jointly insignificant at all horizons, while forward skewness coefficients remain significant at 10% level only at the 6-month horizon.

Collectively, the empirical results presented in this section indicate that forward skewness coefficients encapsulate important information about future stock market returns that is not embedded in forward variances. Moreover, their effect is stronger for horizons between three and nine months ahead. It should be noted, however, that the joint impact of forward skewness coefficients appears to be limited when the alternative Hodrick standard errors are employed.

4.5.3 Forecasting risk and uncertainty

As a final step in this analysis, we examine the predictive power of the forward moments for systemic risk, tail risk, equity and economic policy uncertainty. In particular, we run the following regressions for horizons of one up to six months ahead:

$$g_{k;t+h} = \alpha_{k;h} + \boldsymbol{\beta}'_{k;h}\mathbf{z}_t + \varepsilon_{k;t+h}, \tag{4.16}$$

where $g_{k;t+h}$ denotes the value of variable k, h months ahead and $\mathbf{z}_t$ is the vector of predictive variables for each of the two models considered. The vector of explanatory variables includes always the time t value of the dependent variable k as an additional control variable. The regression analysis covers the period 1996:01 - 2012:12 and for each forecasting horizon we lose h observations. In order to control for possible autocorrelation in the error term we use the Newey and West (1987) covariance matrix estimator with lag length equal to the forecasting horizon for the individual and joint significance tests. However, since we do not have overlapping observations in this case, we do not perform an additional analysis using the Hodrick (1992) covariance matrix estimator. We further choose to include TERM as a control variable as it is considered to be a business cycle indicator (Chen, 1991, Estrella and Hardouvelis, 1991). All the variables have been standardized prior to the regression analysis, so regression coefficients represent the change in the dependent variable in terms of its standard deviation caused by a change of a one standard deviation in each regressor.

4.5.3.1 Systemic and tail risk

Table 4.13 reports the results from predicting systemic risk for 1- up to 6-month horizons. The results for forward variances are similar for both the simple and the augmented model. In particular, $FV^{(1)}$ is negatively and significantly related to future systemic risk for 5- and 6-month horizons, while $FV^{(4)}$ is positively and significantly related to future systemic risk for 1- up to 3-month horizons. Turning to the forward skewness coefficients, $FSC^{(1)}$ and $FSC^{(3)}$ are negatively associated with systemic risk, while $FSC^{(2)}$ and $FSC^{(4)}$ are positively associated with systemic risk. $FSC^{(2)}$ has the strongest effect especially for 4- and 5- month forecasting horizons, but all forward skewness coefficients exhibit some significant predictive power for at least two horizons. In economic terms, a one standard deviation increase in $FSC^{(2)}$ forecasts a 0.346 standard deviation increase in systemic risk four months ahead. TERM does not appear to have any significant predictive ability for future systemic risk. The Wald tests show that forward variances are jointly significant mainly for the 5- and 6-month horizons and marginally for the 1-month horizon in the case of the augmented model. In contrast, forward skewness coefficients are jointly significant across all horizons and have the strongest effect for the 4- and 5-month horizons. It is also worth noting that the adjusted R^2 of the augmented model at the 4-month horizon is 5.7% higher than that of the simple model. The top left panel of Figure 4.5 shows clearly that the increase in adjusted R^2 when considering the augmented model rises for horizons of one to four months ahead and then gradually declines.

The results for tail risk, presented in Table 4.14, are somewhat different. For both the simple and the augmented model, $FV^{(2)}$ is positively related to future tail risk, while $FV^{(4)}$ is negatively related to future tail risk. Moreover, the effects are significant only for the 2- and 5-month horizons. With respect to the forward skewness coefficients, $FSC^{(3)}$ forecasts increased tail risk especially for horizons between three and five months ahead, while $FSC^{(4)}$ forecasts decreased tail risk especially for the 3-month horizon. These relationships are opposite to the ones found for systemic risk, which is expected as in our sample the two measures exhibit a correlation of -0.56.[16] In economic terms, the effect of forward skewness coefficients for tail risk is lower than that reported for systemic risk. For example, a one standard deviation increase in $FSC^{(3)}$ forecasts only a 0.189 standard deviation increase in tail risk four months ahead. As in the case of systemic risk, TERM does not exhibit any significant relation with tail risk. Turning to joint significance, forward variances are jointly significant at the 5% level for the 2-month horizon and at the 10% level for the 5-month horizon. Forward skewness coefficients exhibit a similar significance for the 2- and 5-month horizons but are also significant at the 5% level for the 3-month horizon. At this horizon, there is also an increase in adjusted R^2 of 1.9 percentage points when considering the augmented model. Therefore, in the case of tail risk, forward skewness coefficients can increase the explanatory power of the simple model mainly for horizons of two and three months ahead but the effect is not particularly strong. This can also be seen in the top right panel of Figure 4.5, which shows an initial small increase in the adjusted R^2 that disappears after the third month.

In summary, the results show that forward skewness coefficients are particularly important for predicting future systemic risk especially for horizons between three and six months ahead, but provide only limited forecasting ability for future tail risk.

[16] This is due to the different methods for estimating the two types of risk. Specifically, systemic risk is based on the shape of the left tail of the distribution of returns, while tail risk is based on the shape of the left tail beyond a specified threshold level which depends on the variance of the distribution. Intuitively, the difference between the two methods is similar to the difference between VaR and CVaR estimation.

4.5.3.2 Equity and economic policy uncertainty

Table 4.15 shows the results from predicting equity uncertainty for 1- up to 6-month horizons. The pattern regarding the predictive power of forward variances is similar for both the simple and the augmented model. In particular, $FV^{(1)}$ is positively and significantly related to 1-month ahead equity uncertainty, while $FV^{(3)}$ is negatively and significantly related to equity uncertainty for one, three, four and five months ahead. With respect to forward skewness coefficients, $FSC^{(2)}$ forecasts increased future equity uncertainty in 4- and 5- month horizon predictive regressions, while $FSC^{(3)}$ forecasts increased future equity uncertainty in 3-month horizon regressions. In economic terms, a one standard deviation increase in $FSC^{(2)}$ forecasts a 0.188 standard deviation increase in equity uncertainty four months ahead. TERM does not appear to have any significant predictive ability for future equity uncertainty. The results for the Wald tests of joint significance show that forward variances are jointly significant only at the 5-month horizon. This is remarkable as equity market uncertainty has been closely linked to VIX in the literature (see for example Baker et al., 2013).[17] In contrast, forward skewness coefficients are significant at the 5% level for the 3-month horizon and at the 10% level for the 4-, 5- and 6-month horizon. Therefore, while equity uncertainty is contemporaneously related to implied variance, its future values are explained better by investors' perceptions about future skewness. As in the case of systemic risk, the increase in adjusted R^2 is highest for the 4-month horizon (5.9 percentage points). This is also depicted in the bottom left panel of Figure 4.5, which shows that similarly to systemic risk, the increase in adjusted R^2 when considering the augmented model for forecasting equity uncertainty rises for horizons of one to four months ahead and then progressively declines.

The empirical evidence regarding economic policy uncertainty predictability is presented in Table 4.16 and provides a completely different picture. Similarly to the equity uncertainty, the results of forward variances hardly change once we include

[17] In our sample period, equity uncertainty has a contemporaneous correlation of 0.59 with $FV^{(1)}$.

forward skewness coefficients into the predictive model. However, $FV^{(2)}$ is now the strongest predictor of future economic policy uncertainty, exhibiting a significantly negative effect for 4- to 6-month ahead horizons. Moreover, $FV^{(3)}$ and $FV^{(4)}$ are now positively related to future economic policy uncertainty, with the effect being significant for the 1- and 6-month horizon in the case of $FV^{(3)}$ and the 5-month horizon in the case of $FV^{(4)}$. Turning to the forward skewness coefficients, the only significant relationships come from the 1-month horizon predictive regressions. In particular, $FSC^{(1)}$ predicts increased economic policy uncertainty, with the effect being significant only at the 10% level, while $FSC^{(2)}$ predicts decreased economic policy uncertainty, with the effect being significant at the 1% level. Moreover, in economic terms the predictability of forward skewness coefficients for 1-month ahead economic policy uncertainty is relatively weak, as a one standard deviation increase in $FSC^{(2)}$ forecasts a 0.121 standard deviation decrease in economic policy uncertainty. In contrast to forward skewness coefficients, TERM has a consistently strong positive impact on future economic policy uncertainty for all horizons considered. The results for the Wald tests of joint significance provide little evidence to support the hypothesis that forward moments can predict economic policy uncertainty. In particular forward variances are only jointly significant at the 6-month horizon, while forward skewness coefficients are only jointly significant at the 1-month horizon. Moreover, the increase in adjusted R^2 when considering the augmented model to forecast 1-month ahead economic policy uncertainty is relatively low (0.4 percentage points). The limited forecasting power of forward skewness coefficients for economic policy uncertainty is also depicted in the bottom right panel of Figure 4.5, where the change in adjusted R^2 associated with the augmented model is flat across horizons and always close to zero.

In summary, the results show that forward skewness coefficients are quite important for predicting future equity uncertainty especially for horizons between three and six months ahead, but exhibit weak forecasting power for economic policy uncertainty and only for the 1-month horizon.

Finally, recall from Section 4.5.2 that the graph of the increase in adjusted R^2 when considering the augmented model for explaining future stock market returns has a hump-shaped pattern with a peak at the 4-month horizon. In this section, we observe a very similar pattern for systemic risk and equity market uncertainty. Moreover, both graphs have their peak at the 4-month horizon as well. Therefore, we conclude that forward skewness coefficients are important for explaining future stock market returns, systemic risk and equity uncertainty for a horizon that matches the time period spanned by forward skewness. After the fourth month that corresponds to the fourth forward skewness coefficient, the predictive power of the augmented model gradually drops.

4.6 Conclusion

This study investigates the information content of forward skewness inferred from portfolios of options on the S&P 500 index. In particular, we construct forward 1-month skewness coefficients for one to four months ahead and examine their predictive power over and above the respective forward variances. In contrast to previous studies, our method is robust to the presence of jumps in the underlying asset process and therefore our variance estimates are valid under very general specifications for the price process. Moreover, this approach allows us to create forward standardized skewness measures instead of relying on the term structure of the third central moment of returns.

The predictive power of the estimated forward moments is tested on a wide range of macroeconomic variables, future stock market returns as well as risk and uncertainty measures. The results show that forward skewness coefficients offer additional predictive power when included into a model containing forward variances for the majority of real activity, money, credit and treasury yield variables. Furthermore, the increase in explanatory power as measured by the change in adjusted R^2 follows an upward-sloping pattern for almost all the variables considered, apart from the treasury yields. In respect to stock market returns and the risk and uncertainty variables, forward skewness coefficients significantly improve the predictability of market returns, systemic risk and equity market uncertainty mostly for horizons between 3 and 6 months. The corresponding graphs of increase in adjusted R^2 after the inclusion of forward skewness into the predictive model all have a consistent hump-shaped pattern with its peak at the 4 month horizon.

Collectively, the analysis in this chapter shows that forward skewness coefficients encapsulate important information over and above the information contained inforward variances about future macroeconomic conditions at both short and long horizons and about future financial market conditions mainly at short horizons.

Table 4.1: Summary statistics of forward moments

	Mean	**StDev**	**Min**	**Max**	**Auto**
$\mathbf{FV^{(1)}}$	0.004	0.004	0.001	0.028	0.775
$\mathbf{FV^{(2)}}$	0.004	0.003	0.001	0.024	0.794
$\mathbf{FV^{(3)}}$	0.005	0.003	0.001	0.024	0.834
$\mathbf{FV^{(4)}}$	0.005	0.003	0.001	0.018	0.799
$\mathbf{FSC^{(1)}}$	-1.059	0.227	-1.631	-0.367	0.533
$\mathbf{FSC^{(2)}}$	-2.048	0.397	-3.611	-0.758	0.548
$\mathbf{FSC^{(3)}}$	-2.593	0.667	-4.562	2.248	0.388
$\mathbf{FSC^{(4)}}$	-3.188	0.697	-6.226	-0.574	0.300

This table reports the summary statistics of the forward variances and forward skewness coefficients constructed using the method of Neuberger (2012) and Kozhan, Neuberger and Schneider (2013). The sample period is 1996:01-2012:12. $FV^{(1)}$, $FV^{(2)}$, $FV^{(3)}$ and $FV^{(4)}$ denote the forward 1-month variances for one, two, three and four months ahead, while $FSC^{(1)}$, $FSC^{(2)}$, $FSC^{(3)}$ and $FSC^{(4)}$ denote the respective skewness coefficients.

Table 4.2: Correlation coefficients

	$\mathbf{FV^{(1)}}$	$\mathbf{FV^{(2)}}$	$\mathbf{FV^{(3)}}$	$\mathbf{FV^{(4)}}$	$\mathbf{FSC^{(1)}}$	$\mathbf{FSC^{(2)}}$	$\mathbf{FSC^{(3)}}$	$\mathbf{FSC^{(4)}}$
$\mathbf{FV^{(1)}}$	1.00							
$\mathbf{FV^{(2)}}$	0.94	1.00						
$\mathbf{FV^{(3)}}$	0.90	0.93	1.00					
$\mathbf{FV^{(4)}}$	0.85	0.92	0.91	1.00				
$\mathbf{FSC^{(1)}}$	0.02	-0.02	-0.03	-0.04	1.00			
$\mathbf{FSC^{(2)}}$	-0.10	-0.08	-0.10	-0.11	0.63	1.00		
$\mathbf{FSC^{(3)}}$	-0.03	-0.08	-0.12	-0.12	0.40	0.45	1.00	
$\mathbf{FSC^{(4)}}$	-0.02	-0.07	-0.07	-0.10	0.37	0.50	0.40	1.00

This table reports the correlation coefficients of the forward variances and forward skewness coefficients constructed using the method of Neuberger (2012) and Kozhan, Neuberger and Schneider (2013). The sample period is 1996:01-2012:12. $FV^{(1)}$, $FV^{(2)}$, $FV^{(3)}$ and $FV^{(4)}$ denote the forward 1-month variances for one, two, three and four months ahead, while $FSC^{(1)}$, $FSC^{(2)}$, $FSC^{(3)}$ and $FSC^{(4)}$ denote the respective skewness coefficients.

Table 4.3: Predicting real activity for 1-month horizon - Newey-West covariance matrix

	TERM	**FV$^{(1)}$**	**FV$^{(2)}$**	**FV$^{(3)}$**	**FV$^{(4)}$**	**FSC$^{(1)}$**	**FSC$^{(2)}$**	**FSC$^{(3)}$**	**FSC$^{(4)}$**	**Adj. R^2**	**Joint FV**	**Joint FSC**
Pers income	-0.771**	-2.261*	0.906	-1.659	0.606					0.152		
	(-2.231)	(-1.900)	(0.601)	(-1.564)	(0.592)						0.000	
	-0.810**	-2.154	0.762	-1.686	0.733	-0.099	0.407	-0.105	0.204	0.140		
	(-2.356)	(-1.494)	(0.450)	(-1.535)	(0.673)	(-0.185)	(0.773)	(-0.197)	(0.220)		0.000	0.664
Ind prod	0.324	-2.232	0.563	-4.051**	2.786***					0.123		
	(0.665)	(-1.082)	(0.272)	(-2.017)	(2.839)						0.000	
	0.707	-1.703	0.064	-3.885**	2.517***	-1.642***	0.085	1.084**	-0.874*	0.150		
	(1.508)	(-0.778)	(0.032)	(-1.981)	(2.611)	(-2.606)	(0.103)	(2.133)	(-1.741)		0.000	0.009
Cap util	1.581***	-1.530	-0.320	-2.323	1.897***					0.146		
	(4.251)	(-0.989)	(-0.195)	(-1.483)	(2.712)						0.000	
	1.776***	-1.165	-0.553	-2.314	1.665**	-0.765	-0.163	0.390	-0.685*	0.160		
	(4.833)	(-0.725)	(-0.347)	(-1.532)	(2.226)	(-1.547)	(-0.262)	(0.994)	(-1.956)		0.000	0.085
Unempl	0.002	0.494	0.175	0.510	-0.432					0.143		
	(0.017)	(1.153)	(0.286)	(1.247)	(-0.792)						0.000	
	-0.079	0.150	0.437	0.552	-0.373	0.493***	-0.190	-0.089	0.339***	0.205		
	(-0.686)	(0.379)	(0.780)	(1.571)	(-0.801)	(3.386)	(-1.244)	(-0.716)	(2.682)		0.000	0.000
Payroll	-0.577***	-1.056*	0.222	-0.655	0.357					0.393		
	(-4.003)	(-1.842)	(0.333)	(-1.314)	(0.648)						0.000	
	-0.473***	-0.790	0.002	-0.648	0.279	-0.489***	0.108	0.224*	-0.357**	0.448		
	(-3.660)	(-1.518)	(0.003)	(-1.454)	(0.630)	(-3.775)	(0.675)	(1.689)	(-2.403)		0.000	0.000
House starts	7.281	-30.823*	13.119	-2.385	5.406					0.021		
	(1.424)	(-1.957)	(0.430)	(-0.113)	(0.365)						0.070	
	7.419	-43.838**	22.363	1.501	5.540	4.375	-9.998	7.137	12.396	0.030		
	(1.428)	(-2.419)	(0.731)	(0.067)	(0.364)	(0.653)	(-1.334)	(1.084)	(1.579)		0.025	0.254
Build perm	4.855	-35.035***	15.524	2.630	8.681					0.046		
	(1.267)	(-3.507)	(0.872)	(0.178)	(0.757)						0.009	
	6.305*	-38.627***	16.550	5.982	7.610	2.494	-3.423	9.333**	-5.722	0.050		
	(1.702)	(-3.383)	(0.905)	(0.394)	(0.680)	(0.548)	(-0.553)	(2.180)	(-0.876)		0.009	0.251
M&T invent	-1.006***	-2.162*	1.996	-0.549	-0.693					0.136		
	(-2.937)	(-1.948)	(1.348)	(-0.438)	(-0.597)						0.001	
	-0.796**	-1.740	1.642	-0.476	-0.926	-0.506	-0.051	0.562*	-1.148***	0.190		
	(-2.462)	(-1.652)	(1.184)	(-0.476)	(-0.996)	(-1.347)	(-0.114)	(1.961)	(-3.153)		0.000	0.001
Consumption	-0.526*	-0.908	0.484	0.126	-0.392					0.017		
	(-1.932)	(-0.732)	(0.299)	(0.113)	(-0.214)						0.299	
	-0.454*	-0.603	0.083	0.216	-0.355	-0.099	0.502	0.364	-0.675**	0.018		
	(-1.720)	(-0.489)	(0.050)	(0.201)	(-0.195)	(-0.269)	(1.033)	(1.435)	(-2.316)		0.247	0.114
Retail sales	-0.161	-9.711**	5.030	2.849	-1.324					0.065		
	(-0.198)	(-2.063)	(1.071)	(0.770)	(-0.263)						0.163	
	0.024	-9.977**	5.059	3.247	-1.615	0.990	-0.658	1.137	-1.831**	0.065		
	(0.030)	(-2.132)	(1.065)	(0.912)	(-0.335)	(0.859)	(-0.498)	(1.291)	(-2.290)		0.112	0.072

This table reports the results of predictive regressions of 1-month growth (or change) in real activity variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.4: Predicting real activity for 1-month horizon - Hodrick covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
Pers income	-0.771**	-2.261	0.906	-1.659	0.606					0.152		
	(-2.334)	(-1.042)	(0.284)	(-0.747)	(0.617)						0.120	
	-0.810**	-2.154	0.762	-1.686	0.733	-0.099	0.407	-0.105	0.204	0.140		
	(-2.526)	(-0.936)	(0.237)	(-0.742)	(0.755)	(-0.155)	(0.734)	(-0.197)	(0.208)		0.140	0.771
Ind prod	0.324	-2.232	0.563	-4.051	2.786*					0.123		
	(0.627)	(-0.646)	(0.118)	(-1.114)	(1.936)						0.064	
	0.707	-1.703	0.064	-3.885	2.517*	-1.642**	0.085	1.084	-0.874	0.150		
	(1.388)	(-0.475)	(0.014)	(-1.048)	(1.739)	(-2.360)	(0.115)	(1.505)	(-1.449)		0.064	0.065
Cap util	1.581***	-1.530	-0.320	-2.323	1.897					0.146		
	(3.658)	(-0.613)	(-0.093)	(-0.886)	(1.474)						0.025	
	1.776***	-1.165	-0.553	-2.314	1.665	-0.765	-0.163	0.390	-0.685	0.160		
	(4.187)	(-0.450)	(-0.162)	(-0.860)	(1.281)	(-1.515)	(-0.290)	(0.823)	(-1.610)		0.022	0.135
Unempl	0.002	0.494	0.175	0.510	-0.432					0.143		
	(0.016)	(0.651)	(0.164)	(0.738)	(-0.975)						0.069	
	-0.079	0.150	0.437	0.552	-0.373	0.493***	-0.190	-0.089	0.339**	0.205		
	(-0.662)	(0.191)	(0.405)	(0.775)	(-0.832)	(3.171)	(-1.138)	(-0.566)	(2.191)		0.064	0.003
Payroll	-0.577***	-1.056	0.222	-0.655	0.357					0.393		
	(-4.548)	(-1.121)	(0.154)	(-0.659)	(0.646)						0.076	
	-0.473***	-0.790	0.002	-0.648	0.279	-0.489***	0.108	0.224	-0.357**	0.448		
	(-3.797)	(-0.792)	(0.001)	(-0.633)	(0.502)	(-2.717)	(0.521)	(1.142)	(-1.974)		0.075	0.005
House starts	7.281	-30.823	13.119	-2.385	5.406					0.021		
	(1.157)	(-1.357)	(0.341)	(-0.088)	(0.395)						0.515	
	7.419	-43.838*	22.363	1.501	5.540	4.375	-9.998	7.137	12.396	0.030		
	(1.139)	(-1.814)	(0.581)	(0.053)	(0.402)	(0.583)	(-1.293)	(1.029)	(1.430)		0.331	0.306
Build perm	4.855	-35.035**	15.524	2.630	8.681					0.046		
	(1.206)	(-2.092)	(0.562)	(0.127)	(0.962)						0.214	
	6.305	-38.627**	16.550	5.982	7.610	2.494	-3.423	9.333**	-5.722	0.050		
	(1.538)	(-2.116)	(0.592)	(0.278)	(0.849)	(0.471)	(-0.462)	(2.069)	(-0.892)		0.207	0.309
M&T invent	-1.006***	-2.162	1.996	-0.549	-0.693					0.136		
	(-3.410)	(-1.401)	(0.838)	(-0.347)	(-0.480)						0.128	
	-0.796***	-1.740	1.642	-0.476	-0.926	-0.506	-0.051	0.562*	-1.148***	0.190		
	(-2.743)	(-1.059)	(0.670)	(-0.299)	(-0.631)	(-1.234)	(-0.112)	(1.720)	(-2.713)		0.109	0.006
Consumption	-0.526*	-0.908	0.484	0.126	-0.392					0.017		
	(-1.745)	(-0.620)	(0.224)	(0.087)	(-0.187)						0.622	
	-0.454	-0.603	0.083	0.216	-0.355	-0.099	0.502	0.364	-0.675**	0.018		
	(-1.533)	(-0.402)	(0.037)	(0.149)	(-0.166)	(-0.261)	(0.967)	(1.286)	(-2.052)		0.629	0.213
Retail sales	-0.161	-9.711	5.030	2.849	-1.324					0.065		
	(-0.179)	(-1.584)	(0.617)	(0.504)	(-0.252)						0.511	
	0.024	-9.977*	5.059	3.247	-1.615	0.990	-0.658	1.137	-1.831**	0.065		
	(0.027)	(-1.666)	(0.630)	(0.558)	(-0.301)	(0.815)	(-0.467)	(1.221)	(-2.100)		0.474	0.124

This table reports the results of predictive regressions of 1-month growth (or change) in real activity variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.5: Predicting real activity for 6- and 12-month horizon - Newey-West covariance matrix

	$\mathbf{FSC^{(1)}}$	$\mathbf{FSC^{(2)}}$	$\mathbf{FSC^{(3)}}$	$\mathbf{FSC^{(4)}}$	**Adj. $\mathbf{R^2}$**	**Adj. $\mathbf{\bar{R}^2}$**	**Joint FV**	**Joint FSC**
				Panel A: h=6				
Pers income	-0.398	0.074	0.639***	-0.753**	0.314	0.276		
	(-1.046)	(0.214)	(2.649)	(-2.357)			0.000	0.054
Ind prod	-1.676***	0.567	0.978**	-0.833	0.219	0.155		
	(-3.802)	(1.090)	(2.434)	(-1.350)			0.000	0.000
Cap util	-0.714*	0.114	0.210	-0.599	0.305	0.270		
	(-1.910)	(0.301)	(0.846)	(-1.371)			0.000	0.404
Unempl	0.240**	-0.060	-0.174*	0.344***	0.358	0.268		
	(2.232)	(-0.602)	(-1.744)	(3.305)			0.000	0.019
Payroll	-0.537***	0.147	0.344***	-0.482***	0.431	0.332		
	(-3.713)	(0.913)	(2.736)	(-3.043)			0.000	0.001
House starts	1.587	-0.612	5.930**	-4.680	0.089	0.061		
	(0.634)	(-0.278)	(2.215)	(-1.322)			0.073	0.166
Build perm	0.652	-0.054	6.825***	-6.809*	0.131	0.076		
	(0.267)	(-0.020)	(2.895)	(-1.779)			0.026	0.022
M&T invent	-1.013***	0.171	0.517*	-0.899***	0.338	0.236		
	(-3.068)	(0.528)	(1.800)	(-3.088)			0.000	0.000
Consumption	-0.426**	0.265	0.555***	-0.373	0.124	0.069		
	(-1.985)	(1.281)	(3.015)	(-1.446)			0.167	0.002
Retail sales	-1.027*	0.213	1.043**	-1.505*	0.121	0.048		
	(-1.878)	(0.514)	(2.211)	(-1.781)			0.158	0.206
				Panel B: h=12				
Pers income	-0.261	-0.235	0.766**	-0.554	0.244	0.193		
	(-0.763)	(-0.660)	(2.573)	(-1.589)			0.000	0.084
Ind prod	-1.158***	0.071	1.147**	-0.574	0.117	0.063		
	(-3.464)	(0.117)	(2.013)	(-1.142)			0.000	0.001
Cap util	-0.224	-0.443	0.380	-0.337	0.279	0.257		
	(-0.855)	(-1.283)	(1.045)	(-0.985)			0.000	0.482
Unempl	0.203*	0.012	-0.193*	0.254**	0.243	0.169		
	(1.961)	(0.103)	(-1.703)	(2.233)			0.000	0.149
Payroll	-0.487***	0.060	0.427**	-0.428**	0.279	0.181		
	(-2.905)	(0.294)	(2.505)	(-2.394)			0.000	0.009
House starts	0.598	0.379	6.532**	-3.490*	0.192	0.128		
	(0.361)	(0.175)	(2.378)	(-1.892)			0.071	0.065
Build perm	-0.658	0.585	6.375**	-3.127	0.204	0.151		
	(-0.363)	(0.271)	(2.377)	(-1.631)			0.097	0.153
M&T invent	-0.964***	0.095	0.649**	-0.730**	0.304	0.203		
	(-2.821)	(0.238)	(2.326)	(-2.118)			0.000	0.020
Consumption	-0.424**	0.270	0.641***	-0.228	0.118	0.036		
	(-2.230)	(1.193)	(2.739)	(-1.058)			0.007	0.004
Retail sales	-0.562	-0.315	1.144*	-0.808*	0.086	0.031		
	(-1.224)	(-0.641)	(1.784)	(-1.827)			0.004	0.291

This table reports the results of predictive regressions of 6- and 12-month growth (or change) in real activity variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. The predictive model considered is augmented with forward skewness coefficients. Individual coefficient results regarding the constant term, term spread and forward variances have been omitted for brevity. Adj. R^2 and Adj. $\bar{R}^2$ denote the adjusted R^2 with and without forward skewness coefficients respectively. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.6: Predicting real activity for 6- and 12-month horizon - Hodrick covariance matrix

	FSC$^{(1)}$	**FSC$^{(2)}$**	**FSC$^{(3)}$**	**FSC$^{(4)}$**	**Adj. R^2**	**Adj. $\bar{\mathrm{R}}^2$**	**Joint FV**	**Joint FSC**
				Panel A: h=6				
Pers income	-0.398	0.074	0.639***	-0.753**	0.314	0.276		
	(-1.553)	(0.375)	(2.754)	(-2.588)			0.026	0.017
Ind prod	-1.676***	0.567*	0.978***	-0.833*	0.219	0.155		
	(-4.597)	(1.798)	(3.169)	(-1.739)			0.023	0.000
Cap util	-0.714**	0.114	0.210	-0.599*	0.305	0.270		
	(-2.319)	(0.463)	(0.900)	(-1.701)			0.006	0.115
Unempl	0.240**	-0.060	-0.174**	0.344***	0.358	0.268		
	(2.547)	(-0.759)	(-1.985)	(3.335)			0.018	0.002
Payroll	-0.537***	0.147*	0.344***	-0.482***	0.431	0.332		
	(-5.503)	(1.819)	(3.767)	(-4.605)			0.006	0.000
House starts	1.587	-0.612	5.930*	-4.680	0.089	0.061		
	(0.393)	(-0.167)	(1.866)	(-1.192)			0.633	0.402
Build perm	0.652	-0.054	6.825**	-6.809**	0.131	0.076		
	(0.225)	(-0.018)	(2.456)	(-2.208)			0.277	0.059
M&T invent	-1.013***	0.171	0.517**	-0.899***	0.338	0.236		
	(-4.164)	(0.920)	(2.369)	(-3.917)			0.019	0.000
Consumption	-0.426*	0.265	0.555***	-0.373**	0.124	0.069		
	(-1.719)	(1.640)	(3.562)	(-2.025)			0.540	0.001
Retail sales	-1.027	0.213	1.043**	-1.505**	0.121	0.048		
	(-1.448)	(0.430)	(2.411)	(-2.357)			0.337	0.073
				12-month				
Pers income	-0.261	-0.235	0.766***	-0.554**	0.244	0.193		
	(-1.118)	(-1.107)	(3.725)	(-2.260)			0.005	0.003
Ind prod	-1.158***	0.071	1.147***	-0.574*	0.117	0.063		
	(-3.546)	(0.227)	(3.100)	(-1.953)			0.054	0.001
Cap util	-0.224	-0.443*	0.380	-0.337	0.279	0.257		
	(-0.835)	(-1.894)	(1.363)	(-1.584)			0.048	0.104
Unempl	0.203***	0.012	-0.193***	0.254***	0.243	0.169		
	(2.658)	(0.165)	(-2.635)	(3.644)			0.002	0.001
Payroll	-0.487***	0.060	0.427***	-0.428***	0.279	0.181		
	(-6.117)	(0.784)	(4.929)	(-4.679)			0.000	0.000
House starts	0.598	0.379	6.532**	-3.490	0.192	0.128		
	(0.186)	(0.122)	(2.377)	(-1.202)			0.612	0.151
Build perm	-0.658	0.585	6.375***	-3.127*	0.204	0.151		
	(-0.300)	(0.301)	(2.844)	(-1.816)			0.359	0.026
M&T invent	-0.964***	0.095	0.649***	-0.730***	0.304	0.203		
	(-4.676)	(0.625)	(4.004)	(-3.622)			0.001	0.000
Consumption	-0.424*	0.270*	0.641***	-0.228*	0.118	0.036		
	(-1.842)	(1.870)	(4.646)	(-1.713)			0.075	0.000
Retail sales	-0.562	-0.315	1.144**	-0.808**	0.086	0.031		
	(-0.862)	(-0.742)	(2.435)	(-2.052)			0.488	0.141

This table reports the results of predictive regressions of 6- and 12-month growth (or change) in real activity variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. The predictive model considered is augmented with forward skewness coefficients. Individual coefficient results regarding the constant term, term spread and forward variances have been omitted for brevity. Adj. R^2 and Adj. $\bar{R}^2$ denote the adjusted R^2 with and without forward skewness coefficients respectively. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.7: Predicting money, credit and yield variables for 1-month horizon - Newey-West covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
M1	3.094***	4.525	-4.933	3.321	-0.229					0.130		
	(4.029)	(1.031)	(-1.025)	(1.231)	(-0.149)						0.398	
	2.872***	5.985	-5.875	2.655	-0.031	-0.323	1.434	-1.649*	-0.460	0.130		
	(3.934)	(1.290)	(-1.153)	(0.926)	(-0.019)	(-0.394)	(1.470)	(-1.765)	(-0.535)		0.320	0.347
M2 (real)	-0.392	5.402***	-2.637	0.383	-1.032					0.199		
	(-0.912)	(3.933)	(-1.627)	(0.325)	(-1.456)						0.000	
	-0.548	5.996***	-2.948*	0.006	-0.907	-0.154	0.596	-1.013***	0.193	0.207		
	(-1.331)	(4.396)	(-1.792)	(0.005)	(-1.332)	(-0.354)	(1.340)	(-2.845)	(0.507)		0.000	0.065
Reserves tot	14.591*	155.972	-92.843	-35.936	11.903					0.166		
	(1.839)	(1.650)	(-1.322)	(-0.757)	(0.471)						0.025	
	10.395	157.158*	-91.424	-40.659	14.673	10.758	3.510	-17.349	8.232	0.165		
	(1.527)	(1.755)	(-1.382)	(-0.836)	(0.609)	(0.795)	(0.352)	(-1.647)	(0.944)		0.018	0.263
C&I loans	-5.383***	1.809	-3.945	-3.054	3.083					0.090		
	(-3.307)	(0.496)	(-0.694)	(-0.626)	(0.656)						0.487	
	-5.016***	2.281	-3.347	-3.544	1.935	-1.618	-2.774*	-0.451	-0.975	0.134		
	(-3.059)	(0.607)	(-0.632)	(-0.755)	(0.494)	(-0.966)	(-1.787)	(-0.361)	(-0.725)		0.238	0.004
CPI	-0.267	-3.116***	0.201	1.130	0.495					0.209		
	(-1.110)	(-3.058)	(0.167)	(1.143)	(0.934)						0.000	
	-0.219	-3.338***	0.304	1.270	0.473	0.075	-0.161	0.359	-0.029	0.201		
	(-0.923)	(-3.069)	(0.255)	(1.279)	(0.914)	(0.292)	(-0.489)	(1.520)	(-0.131)		0.000	0.675
3-m t-bill	0.080	-1.921***	0.459	0.440	0.388					0.122		
	(0.546)	(-3.686)	(0.646)	(0.830)	(0.958)						0.000	
	0.157	-2.097***	0.528	0.575	0.366	-0.217	-0.112	0.406**	0.097	0.132		
	(1.005)	(-4.058)	(0.738)	(1.271)	(0.965)	(-1.136)	(-0.433)	(2.299)	(0.649)		0.000	0.101
6-m t-bill	0.127	-2.130***	0.814	0.205	0.362					0.176		
	(0.868)	(-5.383)	(1.512)	(0.483)	(0.852)						0.000	
	0.208	-2.320***	0.902	0.343	0.326	-0.210	-0.160	0.418***	0.081	0.190		
	(1.359)	(-5.618)	(1.590)	(0.963)	(0.833)	(-1.120)	(-0.624)	(2.651)	(0.616)		0.000	0.019
1-yr t-bond	0.169	-2.412***	1.159**	0.208	0.225					0.179		
	(1.029)	(-6.076)	(2.180)	(0.513)	(0.532)						0.000	
	0.246	-2.658***	1.301**	0.349	0.184	-0.199	-0.227	0.412***	0.165	0.194		
	(1.441)	(-6.809)	(2.334)	(0.955)	(0.483)	(-1.064)	(-0.949)	(2.706)	(1.142)		0.000	0.024
5-yr t-bond	-0.018	-2.360***	1.311	0.531	-0.112					0.073		
	(-0.081)	(-3.625)	(1.555)	(0.915)	(-0.192)						0.000	
	0.029	-2.745***	1.560*	0.683	-0.116	-0.114	-0.273	0.368**	0.427**	0.087		
	(0.133)	(-4.506)	(1.817)	(1.186)	(-0.214)	(-0.441)	(-1.146)	(2.011)	(2.033)		0.000	0.039

This table reports the results of predictive regressions of 1-month growth (or change) in money, credit and yield variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.8: Predicting money, credit and yield variables for 1-month horizon - Hodrick covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
M1	3.094***	4.525	-4.933	3.321	-0.229					0.130		
	(4.000)	(0.789)	(-0.594)	(0.569)	(-0.138)						0.699	
	2.872***	5.985	-5.875	2.655	-0.031	-0.323	1.434	-1.649*	-0.460	0.130		
	(3.861)	(1.040)	(-0.712)	(0.448)	(-0.019)	(-0.386)	(1.340)	(-1.803)	(-0.500)		0.631	0.360
M2 (real)	-0.392	5.402*	-2.637	0.383	-1.032					0.199		
	(-0.976)	(1.733)	(-0.640)	(0.136)	(-1.104)						0.174	
	-0.548	5.996*	-2.948	0.006	-0.907	-0.154	0.596	-1.013**	0.193	0.207		
	(-1.401)	(1.932)	(-0.725)	(0.002)	(-0.978)	(-0.330)	(1.234)	(-2.487)	(0.472)		0.145	0.133
Reserves tot	14.591**	155.972	-92.843	-35.936	11.903					0.166		
	(2.164)	(1.324)	(-0.786)	(-0.577)	(0.642)						0.300	
	10.395**	157.158	-91.424	-40.659	14.673	10.758	3.510	-17.349	8.232	0.165		
	(2.149)	(1.407)	(-0.830)	(-0.622)	(0.782)	(0.687)	(0.234)	(-1.473)	(1.169)		0.224	0.245
C&I loans	-5.383***	1.809	-3.945	-3.054	3.083					0.090		
	(-3.629)	(0.450)	(-0.562)	(-0.633)	(0.641)						0.582	
	-5.016***	2.281	-3.347	-3.544	1.935	-1.618	-2.774	-0.451	-0.975	0.134		
	(-3.212)	(0.503)	(-0.462)	(-0.708)	(0.399)	(-0.962)	(-1.428)	(-0.334)	(-0.644)		0.416	0.015
CPI	-0.267	-3.116*	0.201	1.130	0.495					0.209		
	(-1.249)	(-1.944)	(0.078)	(0.580)	(1.204)						0.360	
	-0.219	-3.338**	0.304	1.270	0.473	0.075	-0.161	0.359	-0.029	0.201		
	(-1.053)	(-2.054)	(0.122)	(0.624)	(1.160)	(0.276)	(-0.505)	(1.208)	(-0.137)		0.342	0.827
3-m t-bill	0.080	-1.921***	0.459	0.440	0.388					0.122		
	(0.500)	(-2.859)	(0.569)	(0.809)	(1.162)						0.002	
	0.157	-2.097***	0.528	0.575	0.366	-0.217	-0.112	0.406*	0.097	0.132		
	(0.944)	(-2.907)	(0.632)	(1.063)	(1.125)	(-0.855)	(-0.396)	(1.960)	(0.600)		0.003	0.236
6-m t-bill	0.127	-2.130***	0.814	0.205	0.362					0.176		
	(0.810)	(-2.813)	(0.831)	(0.306)	(1.007)						0.012	
	0.208	-2.320***	0.902	0.343	0.326	-0.210	-0.160	0.418**	0.081	0.190		
	(1.286)	(-2.862)	(0.891)	(0.515)	(0.930)	(-0.823)	(-0.541)	(2.219)	(0.528)		0.013	0.071
1-yr t-bond	0.169	-2.412***	1.159	0.208	0.225					0.179		
	(0.964)	(-2.619)	(1.055)	(0.284)	(0.507)						0.007	
	0.246	-2.658***	1.301	0.349	0.184	-0.199	-0.227	0.412**	0.165	0.194		
	(1.361)	(-2.750)	(1.154)	(0.479)	(0.424)	(-0.778)	(-0.780)	(2.357)	(0.904)		0.006	0.048
5-yr t-bond	-0.018	-2.360**	1.311	0.531	-0.112					0.073		
	(-0.080)	(-2.258)	(0.947)	(0.531)	(-0.172)						0.113	
	0.029	-2.745**	1.560	0.683	-0.116	-0.114	-0.273	0.368*	0.427*	0.087		
	(0.129)	(-2.498)	(1.113)	(0.669)	(-0.177)	(-0.445)	(-1.071)	(1.906)	(1.723)		0.069	0.098

This table reports the results of predictive regressions of 1-month growth (or change) in money, credit and yield variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.9: Predicting money, credit and yield variables for 6- and 12-month horizon - Newey-West covariance matrix

	FSC$^{(1)}$	**FSC$^{(2)}$**	**FSC$^{(3)}$**	**FSC$^{(4)}$**	**Adj. R^2**	**Adj. $\bar{\mathbf{R}}^2$**	**Joint FV**	**Joint FSC**
				Panel A: h=6				
M1	0.266	-0.283	-0.739*	-0.326	0.331	0.323		
	(0.313)	(-0.437)	(-1.759)	(-0.473)			0.028	0.171
M2 (real)	-0.319	0.164	-0.119	0.205	0.067	0.077		
	(-0.640)	(0.450)	(-0.616)	(0.580)			0.000	0.859
Reserves tot	10.938	-0.121	-9.856	16.644	0.045	0.018		
	(1.307)	(-0.020)	(-1.284)	(1.054)			0.077	0.683
C&I loans	-3.490***	-0.789	-0.556	-2.352***	0.413	0.243		
	(-2.672)	(-0.897)	(-0.560)	(-2.629)			0.000	0.000
CPI	-0.021	-0.041	0.064	-0.310	0.113	0.101		
	(-0.111)	(-0.194)	(0.407)	(-1.602)			0.026	0.515
3-m t-bill	-0.175	0.019	0.164**	-0.132	0.060	0.061		
	(-1.099)	(0.128)	(2.066)	(-1.176)			0.081	0.280
6-m t-bill	-0.181	-0.002	0.164**	-0.138	0.087	0.083		
	(-1.173)	(-0.016)	(2.347)	(-1.258)			0.018	0.184
1-yr t-bond	-0.156	-0.023	0.150**	-0.145	0.081	0.079		
	(-1.021)	(-0.142)	(2.071)	(-1.300)			0.011	0.293
5-yr t-bond	0.106	-0.114	0.129	-0.122	0.051	0.056		
	(0.678)	(-0.645)	(1.409)	(-1.221)			0.002	0.420
				Panel B: h=12				
M1	-0.156	0.765*	-1.055**	-0.588	0.369	0.339		
	(-0.257)	(1.720)	(-1.996)	(-1.026)			0.020	0.139
M2 (real)	-0.471	0.783***	-0.198	-0.048	0.066	0.034		
	(-1.441)	(3.410)	(-0.934)	(-0.171)			0.000	0.012
Reserves tot	3.284	8.651	-15.066	7.187	0.035	0.010		
	(0.546)	(1.041)	(-1.409)	(0.932)			0.007	0.724
C&I loans	-3.485***	-1.068	0.370	-2.977***	0.482	0.260		
	(-2.757)	(-0.895)	(0.331)	(-4.101)			0.000	0.000
CPI	0.077	-0.405***	0.135	-0.237**	0.222	0.095		
	(0.645)	(-3.154)	(1.142)	(-2.264)			0.000	0.004
3-m t-bill	-0.095	-0.076	0.193*	-0.119	0.142	0.138		
	(-0.536)	(-0.418)	(1.875)	(-1.141)			0.355	0.371
6-m t-bill	-0.092	-0.080	0.190**	-0.134	0.161	0.154		
	(-0.525)	(-0.456)	(1.976)	(-1.256)			0.206	0.313
1-yr t-bond	-0.058	-0.081	0.169*	-0.140	0.152	0.147		
	(-0.330)	(-0.486)	(1.728)	(-1.274)			0.155	0.392
5-yr t-bond	0.097	-0.075	0.049	-0.041	0.033	0.045		
	(0.670)	(-0.705)	(0.558)	(-0.431)			0.134	0.824

This table reports the results of predictive regressions of 6- and 12-month growth (or change) in money, credit and yield variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. The predictive model considered is augmented with forward skewness coefficients. Individual coefficient results regarding the constant term, term spread and forward variances have been omitted for brevity. Adj. R^2 and Adj. $\bar{R}^2$ denote the adjusted R^2 with and without forward skewness coefficients respectively. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.10: Predicting money, credit and yield variables for 6- and 12-month horizon - Hodrick covariance matrix

	FSC$^{(1)}$	FSC$^{(2)}$	FSC$^{(3)}$	FSC$^{(4)}$	Adj. R^2	Adj. $\bar{R}^2$	Joint FV	Joint FSC
				Panel A: h=6				
M1	0.266	-0.283	-0.739*	-0.326	0.331	0.323		
	(0.346)	(-0.589)	(-1.811)	(-0.549)			0.734	0.118
M2 (real)	-0.319	0.164	-0.119	0.205	0.067	0.077		
	(-0.834)	(0.740)	(-0.557)	(0.800)			0.027	0.646
Reserves tot	10.938*	-0.121	-9.856*	16.644	0.045	0.018		
	(1.807)	(-0.044)	(-1.831)	(1.530)			0.175	0.165
C&I loans	-3.490***	-0.789	-0.556	-2.352***	0.413	0.243		
	(-3.441)	(-0.978)	(-0.740)	(-3.048)			0.003	0.000
CPI	-0.021	-0.041	0.064	-0.310*	0.113	0.101		
	(-0.135)	(-0.313)	(0.420)	(-1.843)			0.161	0.238
3-m t-bill	-0.175	0.019	0.164*	-0.132	0.060	0.061		
	(-1.535)	(0.171)	(1.850)	(-1.587)			0.001	0.225
6-m t-bill	-0.181	-0.002	0.164**	-0.138*	0.087	0.083		
	(-1.609)	(-0.025)	(2.081)	(-1.806)			0.002	0.148
1-yr t-bond	-0.156	-0.023	0.150*	-0.145*	0.081	0.079		
	(-1.280)	(-0.206)	(1.747)	(-1.712)			0.009	0.232
5-yr t-bond	0.106	-0.114	0.129	-0.122	0.051	0.056		
	(0.634)	(-0.831)	(1.241)	(-0.938)			0.092	0.452
				12-month				
M1	-0.156	0.765*	-1.055**	-0.588	0.369	0.339		
	(-0.288)	(1.907)	(-2.489)	(-1.252)			0.182	0.002
M2 (real)	-0.471*	0.783***	-0.198	-0.048	0.066	0.034		
	(-1.760)	(4.025)	(-0.902)	(-0.242)			0.179	0.002
Reserves tot	3.284	8.651**	-15.066**	7.187	0.035	0.010		
	(0.967)	(1.997)	(-2.032)	(1.522)			0.111	0.271
C&I loans	-3.485***	-1.068*	0.370	-2.977***	0.482	0.260		
	(-4.172)	(-1.750)	(0.590)	(-4.697)			0.000	0.000
CPI	0.077	-0.405***	0.135	-0.237**	0.222	0.095		
	(0.629)	(-3.312)	(0.851)	(-2.221)			0.006	0.005
3-m t-bill	-0.095	-0.076	0.193**	-0.119***	0.142	0.138		
	(-0.874)	(-0.776)	(2.287)	(-2.734)			0.000	0.036
6-m t-bill	-0.092	-0.080	0.190**	-0.134***	0.161	0.154		
	(-0.860)	(-0.843)	(2.434)	(-3.125)			0.000	0.017
1-yr t-bond	-0.058	-0.081	0.169**	-0.140***	0.152	0.147		
	(-0.498)	(-0.791)	(1.983)	(-2.998)			0.000	0.029
5-yr t-bond	0.097	-0.075	0.049	-0.041	0.033	0.045		
	(0.675)	(-0.657)	(0.502)	(-0.457)			0.013	0.852

This table reports the results of predictive regressions of 6- and 12-month growth (or change) in money, credit and yield variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. The predictive model considered is augmented with forward skewness coefficients. Individual coefficient results regarding the constant term, term spread and forward variances have been omitted for brevity. Adj. R^2 and Adj. $\bar{R}^2$ denote the adjusted R^2 with and without forward skewness coefficients respectively. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.11: Predicting excess stock market returns - Newey-West covariance matrix

	d-p	e-p	$\mathbf{FV^{(1)}}$	$\mathbf{FV^{(2)}}$	$\mathbf{FV^{(3)}}$	$\mathbf{FV^{(4)}}$	$\mathbf{FSC^{(1)}}$	$\mathbf{FSC^{(2)}}$	$\mathbf{FSC^{(3)}}$	$\mathbf{FSC^{(4)}}$	Adj. $\mathbf{R^2}$	Joint FV	Joint FSC
h=1	5.000	5.184	-12.946	-10.429	6.622	21.071*					0.013		
	(0.923)	(0.946)	(-0.530)	(-0.419)	(0.439)	(1.823)						0.254	
	6.309	6.472	-15.278	-11.061	9.411	21.972*	1.296	1.022	8.244*	-5.574	0.013		
	(1.191)	(1.144)	(-0.630)	(-0.456)	(0.641)	(1.932)	(0.234)	(0.169)	(1.734)	(-1.308)		0.232	0.328
h=3	5.518	3.542	-17.225	10.609	2.508	8.436					0.047		
	(1.242)	(0.814)	(-1.352)	(1.227)	(0.289)	(1.170)						0.140	
	6.492	0.893	-16.398	9.128	3.522	6.242	-1.067	-1.846	4.769**	-6.391**	0.062		
	(1.594)	(0.196)	(-1.364)	(1.041)	(0.454)	(0.926)	(-0.310)	(-0.477)	(2.034)	(-2.040)		0.307	0.075
h=6	5.604	3.599	-9.588	4.816	4.754	6.781					0.112		
	(1.316)	(1.063)	(-1.332)	(1.032)	(0.785)	(1.400)						0.020	
	6.594*	2.242	-10.458*	4.336	6.224	5.593	-1.143	-2.116	5.697***	-4.386	0.140		
	(1.784)	(0.592)	(-1.678)	(0.860)	(1.372)	(1.160)	(-0.380)	(-0.888)	(3.474)	(-1.415)		0.046	0.017
h=9	6.452*	3.055	-6.276	3.428	4.045	3.988					0.144		
	(1.697)	(1.151)	(-1.450)	(0.921)	(0.950)	(0.866)						0.078	
	7.357**	2.683	-7.001*	2.405	5.394*	3.961	-0.019	-0.628	4.705***	-3.917	0.168		
	(2.194)	(0.936)	(-1.734)	(0.526)	(1.810)	(0.942)	(-0.007)	(-0.287)	(2.660)	(-1.572)		0.058	0.039
h=12	7.301**	2.441	-5.339	5.267	0.624	3.026					0.178		
	(2.187)	(1.229)	(-1.628)	(1.244)	(0.160)	(0.703)						0.296	
	7.771**	2.629	-6.662**	5.430	1.773	3.113	-0.843	-0.812	3.373*	-0.883	0.181		
	(2.584)	(1.235)	(-2.074)	(1.143)	(0.543)	(0.757)	(-0.381)	(-0.398)	(1.801)	(-0.481)		0.193	0.364

This table reports the results of predictive regressions of excess S&P 500 index returns for horizons of 1, 3, 6, 9 and 12 months. The sample period is 1996:01-2012:12. For each horizon the first two rows present the results of a simple model containing only dividend-to-price ratio, earnings-to-price ratio and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.12: Predicting excess stock market returns - Hodrick covariance matrix

	d-p	**e-p**	**$FV^{(1)}$**	**$FV^{(2)}$**	**$FV^{(3)}$**	**$FV^{(4)}$**	**$FSC^{(1)}$**	**$FSC^{(2)}$**	**$FSC^{(3)}$**	**$FSC^{(4)}$**	**Adj. R^2**	**Joint FV**	**Joint FSC**
h=1	5.000	5.184	-12.946	-10.429	6.622	21.071					0.013		
	(0.951)	(0.849)	(-0.521)	(-0.362)	(0.426)	(1.611)						0.385	
	6.309	6.472	-15.278	-11.061	9.411	21.972	1.296	1.022	8.244*	-5.574	0.013		
	(1.225)	(0.984)	(-0.613)	(-0.393)	(0.580)	(1.644)	(0.230)	(0.162)	(1.713)	(-1.224)		0.347	0.356
h=3	5.518	3.542	-17.225	10.609	2.508	8.436					0.047		
	(1.071)	(0.714)	(-1.642)	(1.014)	(0.271)	(1.032)						0.228	
	6.492	0.893	-16.398	9.128	3.522	6.242	-1.067	-1.846	4.769*	-6.391**	0.062		
	(1.276)	(0.173)	(-1.531)	(0.837)	(0.371)	(0.781)	(-0.293)	(-0.558)	(1.776)	(-2.163)		0.383	0.116
h=6	5.604	3.599	-9.588	4.816	4.754	6.781					0.112		
	(1.125)	(0.831)	(-1.301)	(0.728)	(0.845)	(1.082)						0.252	
	6.594	2.242	-10.458	4.336	6.224	5.593	-1.143	-2.116	5.697**	-4.386*	0.140		
	(1.343)	(0.488)	(-1.541)	(0.649)	(1.124)	(0.903)	(-0.352)	(-0.812)	(2.523)	(-1.741)		0.302	0.088
h=9	6.452	3.055	-6.276	3.428	4.045	3.988					0.144		
	(1.325)	(0.792)	(-1.029)	(0.535)	(0.940)	(0.646)						0.485	
	7.357	2.683	-7.001	2.405	5.394	3.961	-0.019	-0.628	4.705**	-3.917**	0.168		
	(1.535)	(0.660)	(-1.225)	(0.362)	(1.266)	(0.648)	(-0.007)	(-0.256)	(2.233)	(-2.014)		0.508	0.141
h=12	7.301	2.441	-5.339	5.267	0.624	3.026					0.178		
	(1.512)	(0.664)	(-1.014)	(0.925)	(0.170)	(0.512)						0.451	
	7.771	2.629	-6.662	5.430	1.773	3.113	-0.843	-0.812	3.373	-0.883	0.181		
	(1.634)	(0.690)	(-1.375)	(0.907)	(0.502)	(0.536)	(-0.311)	(-0.398)	(1.415)	(-0.580)		0.420	0.645

This table reports the results of predictive regressions of excess S&P 500 index returns for horizons of 1, 3, 6, 9 and 12 months. The sample period is 1996:01-2012:12. For each horizon the first two rows present the results of a simple model containing only dividend-to-price ratio, earnings-to-price ratio and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.13: Predicting systemic risk

	Dep_t	**TERM**	**$FV^{(1)}$**	**$FV^{(2)}$**	**$FV^{(3)}$**	**$FV^{(4)}$**	**$FSC^{(1)}$**	**$FSC^{(2)}$**	**$FSC^{(3)}$**	**$FSC^{(4)}$**	**Adj. R^2**	**Joint FV**	**Joint FSC**
h=1	0.667***	0.023	-0.066	0.101	-0.139	0.214*					0.554		
	(8.181)	(0.510)	(-0.363)	(0.500)	(-0.855)	(1.848)						0.162	
	0.654***	0.008	-0.005	0.069	-0.174	0.236**	-0.070	0.068	-0.102**	0.085	0.560		
	(7.736)	(0.194)	(-0.027)	(0.348)	(-1.113)	(2.023)	(-1.290)	(1.303)	(-2.417)	(1.488)		0.089	0.040
h=2	0.540***	0.024	-0.042	-0.221	0.006	0.378**					0.404		
	(4.760)	(0.371)	(-0.157)	(-0.747)	(0.024)	(2.002)						0.266	
	0.541***	0.015	0.066	-0.313	-0.030	0.414**	-0.139**	0.181**	-0.082	0.068	0.419		
	(4.649)	(0.253)	(0.251)	(-1.085)	(-0.119)	(2.211)	(-2.176)	(2.368)	(-1.094)	(0.823)		0.207	0.075
h=3	0.540***	0.022	0.098	-0.386	0.078	0.277*					0.349		
	(3.427)	(0.287)	(0.336)	(-1.625)	(0.397)	(1.809)						0.356	
	0.549***	-0.002	0.154	-0.433*	0.040	0.320**	-0.086	0.158*	-0.117**	0.151**	0.379		
	(3.539)	(-0.025)	(0.527)	(-1.797)	(0.222)	(2.600)	(-1.175)	(1.896)	(-2.119)	(2.102)		0.131	0.025
h=4	0.582***	0.046	-0.144	0.048	-0.095	0.076					0.241		
	(3.291)	(0.555)	(-0.541)	(0.214)	(-0.332)	(0.499)						0.807	
	0.598***	0.017	0.064	-0.125	-0.168	0.137	-0.136	0.346***	-0.177***	0.023	0.298		
	(3.291)	(0.224)	(0.223)	(-0.540)	(-0.612)	(0.898)	(-1.622)	(4.497)	(-2.950)	(0.288)		0.713	0.000
h=5	0.705***	0.011	-0.544***	0.049	0.294	-0.101					0.282		
	(3.529)	(0.142)	(-2.709)	(0.315)	(1.165)	(-0.855)						0.008	
	0.732***	-0.009	-0.391*	-0.094	0.243	-0.043	-0.150*	0.326***	-0.111	0.062	0.337		
	(3.567)	(-0.121)	(-1.936)	(-0.627)	(1.024)	(-0.360)	(-1.759)	(4.029)	(-1.524)	(0.828)		0.008	0.002
h=6	0.607***	-0.029	-0.733***	0.123	0.286	0.085					0.240		
	(3.037)	(-0.364)	(-3.538)	(0.651)	(1.522)	(0.564)						0.000	
	0.626***	-0.043	-0.683***	0.075	0.264	0.117	-0.125*	0.180**	-0.062	0.138*	0.268		
	(3.106)	(-0.529)	(-2.943)	(0.363)	(1.488)	(0.768)	(-1.750)	(2.052)	(-0.954)	(1.760)		0.000	0.025

This table reports the results of predictive regressions of systemic risk for horizons of 1 to 6 months ahead. The sample period is 1996:01-2012:12. For each horizon the first two rows present the results of a simple model containing only the dependent variable of the current month, term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.14: Predicting tail risk

	$\mathbf{Dep_t}$	**TERM**	$\mathbf{FV^{(1)}}$	$\mathbf{FV^{(2)}}$	$\mathbf{FV^{(3)}}$	$\mathbf{FV^{(4)}}$	$\mathbf{FSC^{(1)}}$	$\mathbf{FSC^{(2)}}$	$\mathbf{FSC^{(3)}}$	$\mathbf{FSC^{(4)}}$	**Adj.** $\mathbf{R^2}$	**Joint FV**	**Joint FSC**
h=1	0.595***	-0.002	0.032	0.244	-0.152	-0.154					0.362		
	(9.672)	(-0.029)	(0.134)	(1.014)	(-0.893)	(-1.090)						0.465	
	0.552***	0.024	0.002	0.229	-0.116	-0.154	-0.058	0.056	0.138*	-0.026	0.367		
	(9.107)	(0.436)	(0.007)	(0.971)	(-0.656)	(-1.052)	(-0.759)	(0.803)	(1.681)	(-0.363)		0.546	0.377
h=2	0.378***	-0.006	-0.065	0.512**	-0.003	-0.494***					0.174		
	(4.645)	(-0.067)	(-0.287)	(2.052)	(-0.017)	(-2.980)						0.029	
	0.352***	0.025	-0.034	0.449*	0.026	-0.494***	-0.070	0.086	0.140*	-0.119*	0.179		
	(4.244)	(0.316)	(-0.180)	(1.832)	(0.155)	(-2.989)	(-0.827)	(0.827)	(1.812)	(-1.836)		0.042	0.025
h=3	0.390***	-0.030	0.083	0.156	-0.073	-0.119					0.121		
	(5.213)	(-0.317)	(0.402)	(0.565)	(-0.318)	(-0.458)						0.922	
	0.404***	-0.001	0.087	0.137	-0.037	-0.145	0.017	-0.056	0.136**	-0.186**	0.140		
	(5.637)	(-0.010)	(0.422)	(0.446)	(-0.194)	(-0.554)	(0.167)	(-0.489)	(1.995)	(-2.469)		0.895	0.038
h=4	0.326***	-0.031	0.217	0.225	-0.384*	0.019					0.090		
	(3.833)	(-0.292)	(1.001)	(0.905)	(-1.658)	(0.083)						0.520	
	0.314***	0.004	0.138	0.261	-0.329	-0.008	0.018	-0.096	0.189**	-0.100	0.102		
	(3.738)	(0.038)	(0.622)	(0.972)	(-1.592)	(-0.031)	(0.145)	(-0.672)	(2.387)	(-1.129)		0.484	0.174
h=5	0.203**	0.002	0.345	0.303	-0.274	-0.360**					0.067		
	(2.068)	(0.022)	(1.439)	(1.590)	(-1.206)	(-2.015)						0.089	
	0.191**	0.025	0.226	0.378**	-0.219	-0.389**	0.112	-0.147	0.159**	-0.089	0.080		
	(2.075)	(0.234)	(1.044)	(2.007)	(-1.181)	(-2.256)	(1.045)	(-1.411)	(2.159)	(-1.227)		0.078	0.052
h=6	0.176*	0.021	0.374	0.221	-0.244	-0.373					0.060		
	(1.776)	(0.187)	(1.408)	(0.851)	(-1.211)	(-1.510)						0.222	
	0.169*	0.040	0.363	0.208	-0.216	-0.382	0.034	-0.013	0.108	-0.125	0.057		
	(1.818)	(0.340)	(1.338)	(0.757)	(-1.121)	(-1.490)	(0.344)	(-0.139)	(1.614)	(-1.330)		0.250	0.364

This table reports the results of predictive regressions of tail risk for horizons of 1 to 6 months ahead. The sample period is 1996:01-2012:12. For each horizon the first two rows present the results of a simple model containing only the dependent variable of the current month, term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.15: Predicting equity uncertainty

	$\mathbf{Dep_t}$	**TERM**	$\mathbf{FV^{(1)}}$	$\mathbf{FV^{(2)}}$	$\mathbf{FV^{(3)}}$	$\mathbf{FV^{(4)}}$	$\mathbf{FSC^{(1)}}$	$\mathbf{FSC^{(2)}}$	$\mathbf{FSC^{(3)}}$	$\mathbf{FSC^{(4)}}$	**Adj. $\mathbf{R^2}$**	**Joint FV**	**Joint FSC**
h=1	0.589***	0.036	0.345**	0.107	-0.333**	-0.036					0.483		
	(8.124)	(0.713)	(2.133)	(0.513)	(-2.407)	(-0.378)						0.092	
	0.587***	0.029	0.333*	0.116	-0.331**	-0.034	0.064	-0.005	-0.007	-0.021	0.475		
	(8.368)	(0.627)	(1.909)	(0.528)	(-2.404)	(-0.353)	(1.159)	(-0.100)	(-0.170)	(-0.415)		0.115	0.680
h=2	0.390***	0.046	0.069	0.303	-0.199	-0.143					0.192		
	(3.475)	(0.578)	(0.407)	(1.204)	(-1.497)	(-1.404)						0.473	
	0.387***	0.051	0.033	0.301	-0.166	-0.123	0.008	0.045	0.078	0.028	0.192		
	(3.610)	(0.652)	(0.183)	(1.159)	(-1.333)	(-1.134)	(0.103)	(0.613)	(1.147)	(0.522)		0.538	0.334
h=3	0.315***	0.051	0.136	0.026	-0.269**	0.092					0.090		
	(2.664)	(0.534)	(0.818)	(0.155)	(-2.022)	(0.847)						0.132	
	0.311***	0.050	0.080	0.013	-0.218*	0.142	0.022	0.117	0.106**	0.076	0.137		
	(2.800)	(0.545)	(0.504)	(0.071)	(-1.899)	(1.029)	(0.199)	(1.498)	(2.061)	(1.261)		0.229	0.015
h=4	0.268**	0.061	0.189	-0.024	-0.267**	0.058					0.059		
	(2.260)	(0.564)	(1.222)	(-0.117)	(-2.009)	(0.485)						0.245	
	0.269**	0.045	0.180	-0.078	-0.233**	0.122	0.061	0.188**	0.045	0.034	0.118		
	(2.436)	(0.432)	(1.253)	(-0.355)	(-2.017)	(0.794)	(0.604)	(2.031)	(0.694)	(0.534)		0.244	0.053
h=5	0.274*	0.059	0.218	0.038	-0.356***	0.008					0.069		
	(1.769)	(0.523)	(1.235)	(0.181)	(-2.940)	(0.060)						0.017	
	0.278*	0.046	0.224	-0.019	-0.330***	0.063	0.038	0.176*	0.035	0.019	0.105		
	(1.914)	(0.412)	(1.353)	(-0.085)	(-2.913)	(0.460)	(0.420)	(1.700)	(0.384)	(0.293)		0.044	0.091
h=6	0.179	0.054	0.112	0.285	-0.211	-0.289*					0.038		
	(1.311)	(0.429)	(0.647)	(1.188)	(-1.547)	(-1.717)						0.327	
	0.181	0.041	0.121	0.246	-0.196	-0.246	-0.002	0.153	0.005	0.063	0.056		
	(1.381)	(0.318)	(0.687)	(1.048)	(-1.587)	(-1.644)	(-0.026)	(1.608)	(0.064)	(0.785)		0.421	0.098

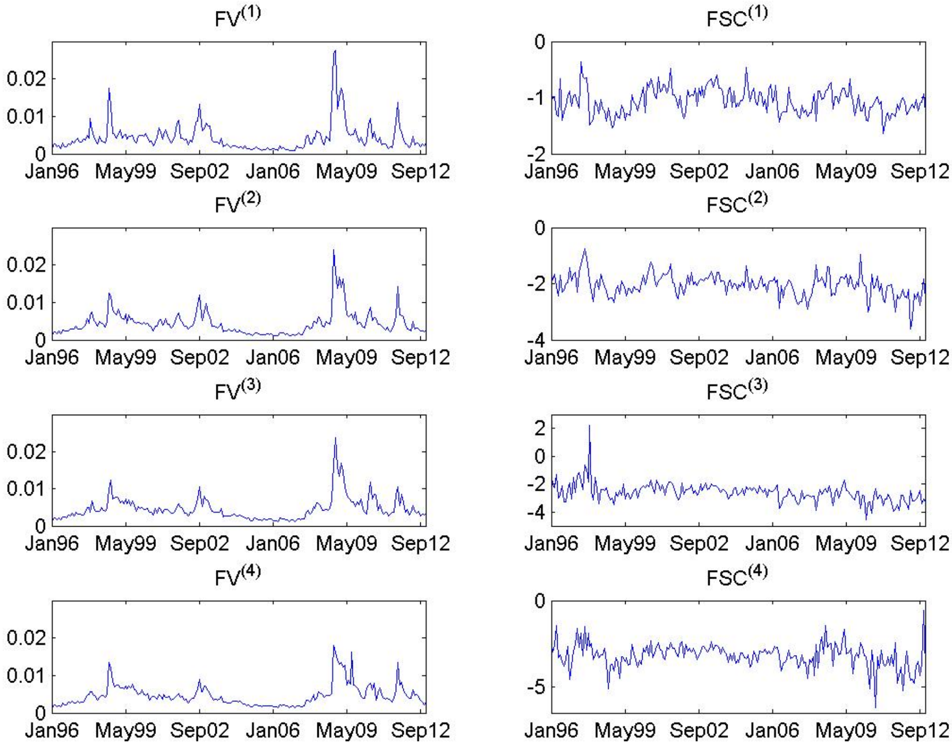
This table reports the results of predictive regressions of equity uncertainty for horizons of 1 to 6 months ahead. The sample period is 1996:01-2012:12. For each horizon the first two rows present the results of a simple model containing only the dependent variable of the current month, term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table 4.16: Predicting economic policy uncertainty

	$\mathbf{Dep_t}$	**TERM**	$\mathbf{FV^{(1)}}$	$\mathbf{FV^{(2)}}$	$\mathbf{FV^{(3)}}$	$\mathbf{FV^{(4)}}$	$\mathbf{FSC^{(1)}}$	$\mathbf{FSC^{(2)}}$	$\mathbf{FSC^{(3)}}$	$\mathbf{FSC^{(4)}}$	**Adj.** $\mathbf{R^2}$	**Joint FV**	**Joint FSC**
h=1	0.833***	0.061*	-0.007	-0.227	0.171*	0.077					0.778		
	(16.202)	(1.830)	(-0.074)	(-1.553)	(1.835)	(0.881)						0.153	
	0.801***	0.075**	-0.063	-0.164	0.184**	0.056	0.075*	-0.121***	0.005	-0.008	0.782		
	(15.100)	(2.138)	(-0.577)	(-1.126)	(2.045)	(0.695)	(1.738)	(-2.818)	(0.178)	(-0.224)		0.153	0.045
h=2	0.697***	0.121**	-0.103	-0.219	0.207	0.135					0.633		
	(9.269)	(2.128)	(-0.869)	(-1.307)	(1.629)	(1.082)						0.102	
	0.672***	0.141**	-0.119	-0.197	0.217*	0.119	0.003	-0.065	0.025	-0.027	0.630		
	(8.724)	(2.373)	(-0.956)	(-1.161)	(1.781)	(1.055)	(0.040)	(-0.813)	(0.744)	(-0.656)		0.081	0.741
h=3	0.619***	0.163**	-0.036	-0.294	0.191	0.170					0.565		
	(7.548)	(2.179)	(-0.289)	(-1.491)	(1.378)	(0.857)						0.229	
	0.608***	0.176**	-0.046	-0.279	0.197	0.159	-0.025	-0.041	0.022	-0.001	0.559		
	(7.245)	(2.338)	(-0.337)	(-1.351)	(1.418)	(0.826)	(-0.300)	(-0.521)	(0.488)	(-0.014)		0.229	0.921
h=4	0.632***	0.163**	0.090	-0.440**	0.095	0.258					0.559		
	(7.033)	(2.028)	(0.699)	(-2.240)	(0.648)	(1.218)						0.221	
	0.631***	0.167**	0.065	-0.415*	0.099	0.250	-0.012	-0.042	0.011	0.028	0.552		
	(6.999)	(2.119)	(0.450)	(-1.966)	(0.664)	(1.186)	(-0.138)	(-0.515)	(0.268)	(0.457)		0.366	0.961
h=5	0.656***	0.158*	0.055	-0.418**	-0.005	0.301*					0.544		
	(6.336)	(1.754)	(0.395)	(-2.082)	(-0.038)	(1.743)						0.249	
	0.649***	0.163*	0.070	-0.426**	-0.009	0.300*	-0.018	0.005	-0.007	-0.006	0.535		
	(6.138)	(1.841)	(0.471)	(-2.003)	(-0.071)	(1.738)	(-0.232)	(0.055)	(-0.150)	(-0.092)		0.249	0.999
h=6	0.616***	0.172	-0.083	-0.312*	0.219**	0.089					0.493		
	(5.559)	(1.579)	(-0.641)	(-1.776)	(2.109)	(0.595)						0.034	
	0.589***	0.184*	-0.022	-0.350*	0.204**	0.091	-0.024	0.041	-0.036	-0.046	0.486		
	(5.264)	(1.671)	(-0.161)	(-1.849)	(2.105)	(0.630)	(-0.262)	(0.411)	(-0.837)	(-0.595)		0.057	0.824

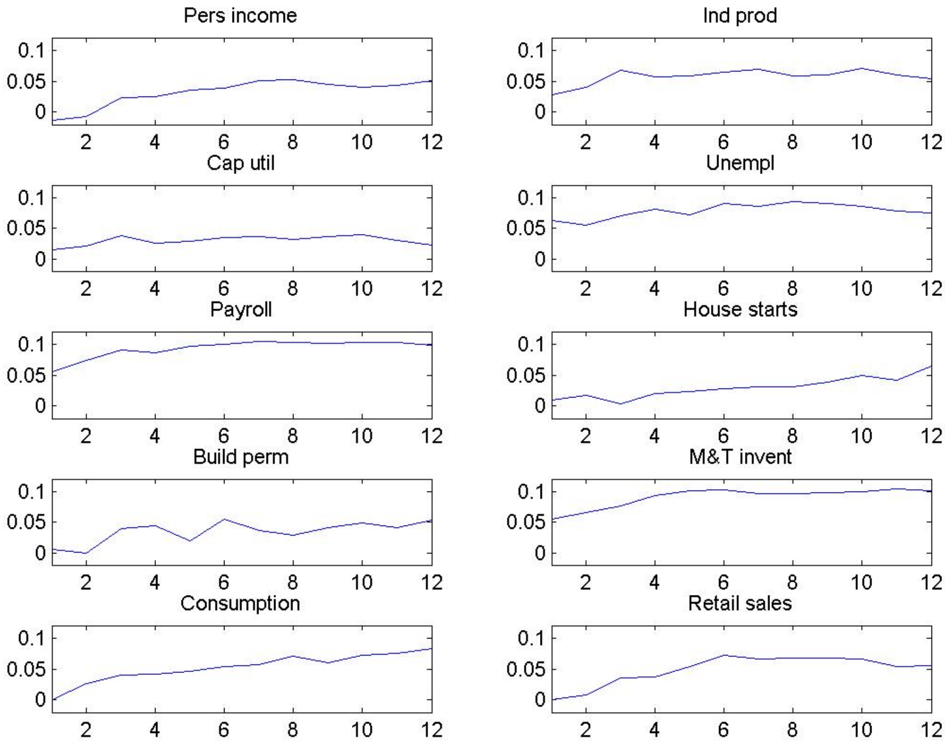
This table reports the results of predictive regressions of economic policy uncertainty for horizons of 1 to 6 months ahead. The sample period is 1996:01-2012:12. For each horizon the first two rows present the results of a simple model containing only the dependent variable of the current month, term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Figure 4.1: Forward variances and skewness coefficients



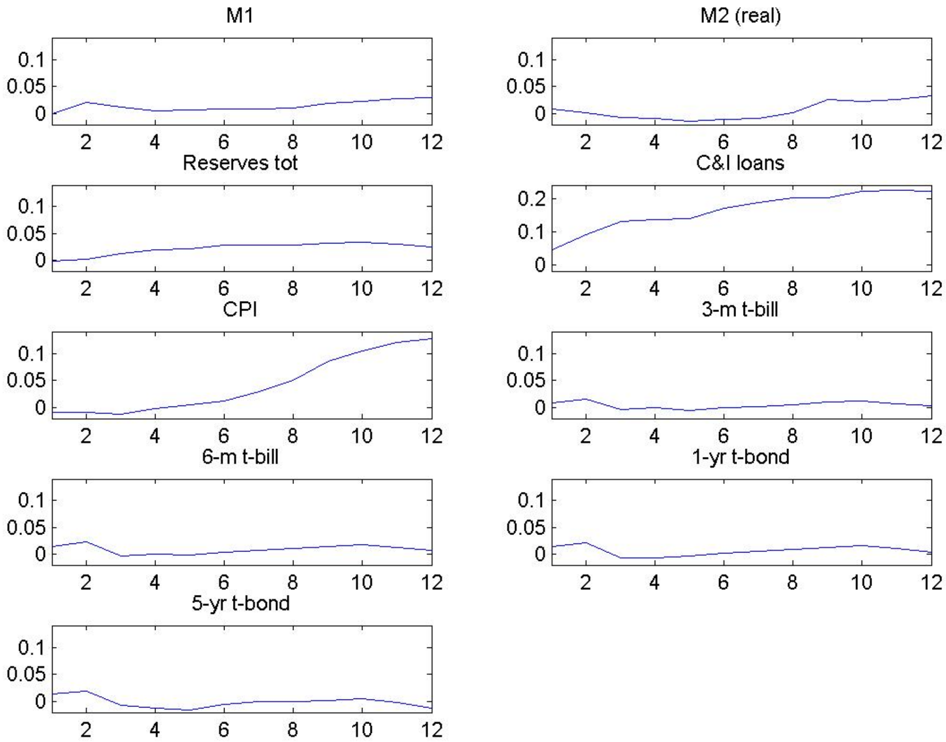
This figure plots the monthly time series of the forward S&P 500 index moments for the period 1996:01-2012:12. The left panels show forward variances while the right panels show forward skewness coefficients.

Figure 4.2: Changes in adjusted R^2 across forecasting horizon for real activity variables



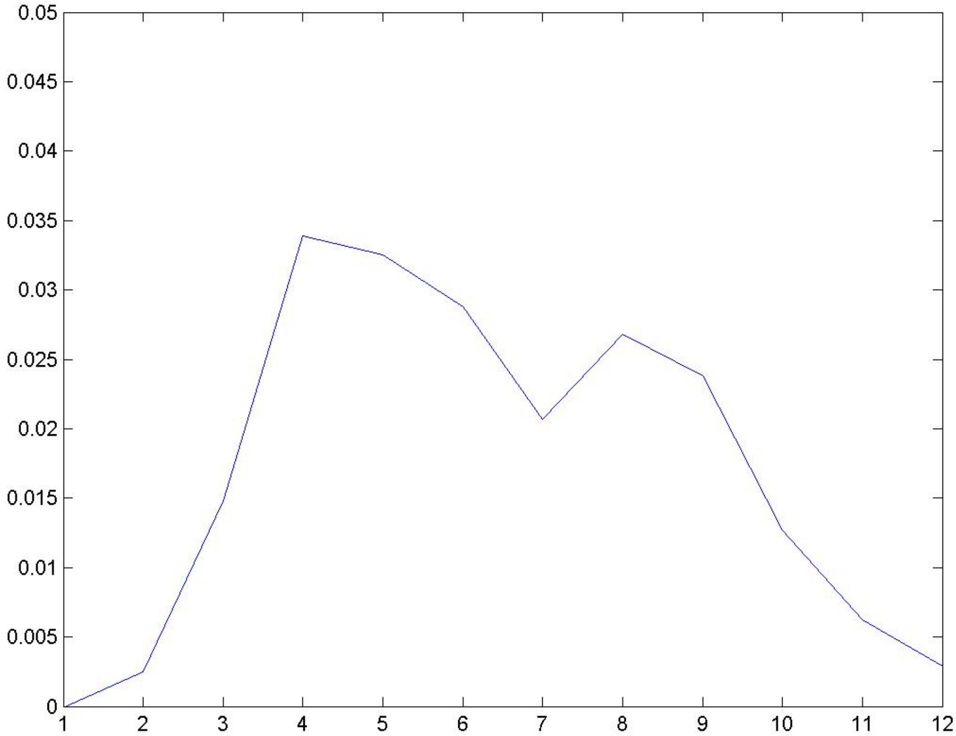
This figure plots the change in adjusted R^2 across horizons when forward skewness coefficients are included into the predictive model of real activity variables. Details about the variables can be found in Table C.1 of Appendix C.

Figure 4.3: Changes in adjusted R^2 across forecasting horizon for money, credit and treasury yield variables



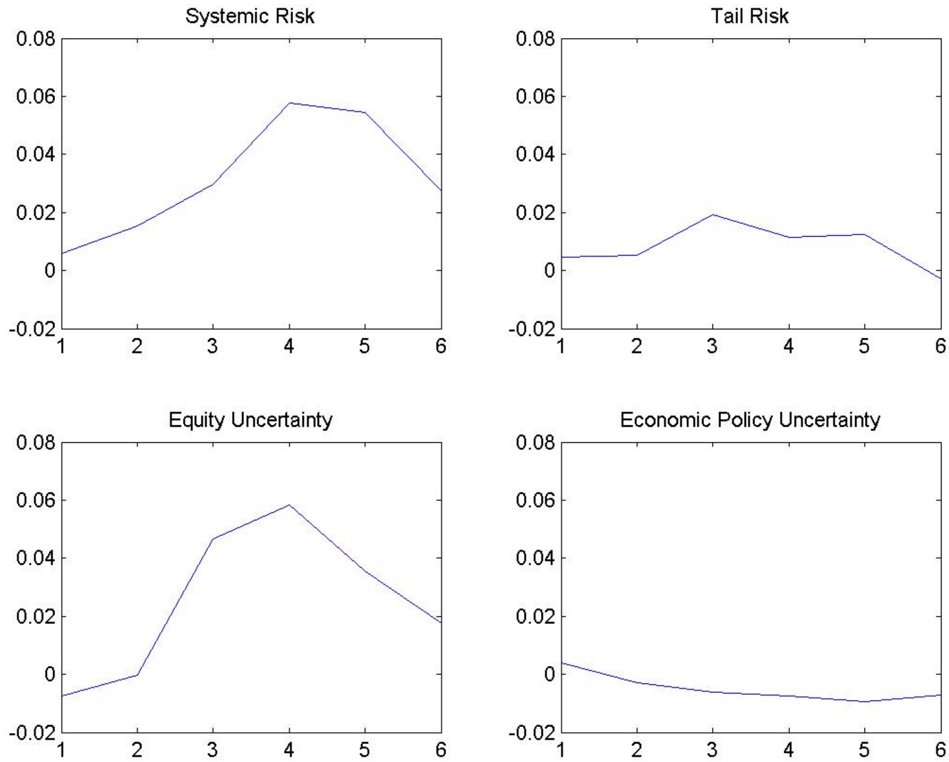
This figure plots the change in adjusted R^2 across horizons when forward skewness coefficients are included into the predictive model of money, credit and treasury yield variables. Details about the variables can be found in Table C.1 of Appendix C.

Figure 4.4: Changes in adjusted R^2 across forecasting horizon for excess stock market returns



This figure plots the change in adjusted R^2 across horizons when forward skewness coefficients are included into the predictive model of stock market excess returns.

Figure 4.5: Changes in adjusted R^2 across forecasting horizon for risk and uncertainty variables



This figure plots the change in adjusted R^2 across horizons when forward skewness coefficients are included into the predictive model of systemic risk, tail risk, equity uncertainty and economic policy uncertainty.

Chapter 5

Dispersion in Options Traders' Expectations and Return Predictability

5.1 Introduction

A growing body of studies has shown that various measures of dispersion in expectations can provide significant stock return predictability at both an individual and an aggregate level. There are two main strands in this literature. A first stream of papers assumes a heterogeneous investors framework and uses dispersion to proxy for the level of disagreement among market participants (e.g. Diether, Malloy and Scherbina, 2002; Yu, 2011; Jiang and Sun, 2014). Disagreement can affect asset returns either due to the existence of trading frictions in the market or by inducing investors to engage into risk-sharing acts that affect asset prices in equilibrium. A second strand of the literature assumes a homogeneous investors framework and uses dispersion to proxy for the level of ambiguity in the market (e.g. Anderson, Ghysels and Juergens, 2009; Drechsler, 2013). Ambiguity can affect asset returns due to the fact that naturally investors exhibit aversion to events with unknown probability distributions.

In this chapter, we explore the information content of the dispersion in options traders' expectations for future market excess returns. In particular, since an option constitutes a direct bet on the future price of the underlying asset, the trades in options of different strike prices can be interpreted as the outcomes of different expected returns. Motivated by the above simple observation we model the dispersion in options traders' expectations via the dispersion in the volume-weighted strike prices of equity index option contracts.

Compared to prior studies that develop dispersion in beliefs measures based on analysts' forecasts or mutual fund and individual investor portfolio holdings,[1] the dispersion in options traders' expectations exhibits several advantageous characteristics. First, it stems from all the trades that take place in a highly liquid options market, thus capturing all the expectations that are considered probable enough to trigger a trade.[2] In contrast, analysts' forecasts constitute only a limited set of opinions,[3] and have been found to be affected by agency issues between firms and investment banks and to be prone to analysts' behavioral biases (Dechow, Hutton and Sloan, 2000; Daniel, Hirshleifer and Teoh, 2002; Cen, Hilary and Wei, 2013). Second, it is directly related to expected returns, while analysts' predictions refer to alternative economic indicators such as corporate earnings and hence supplementary modeling assumptions are needed to derive expectations about returns. Third, unlike dispersion measures constructed from analysts' forecasts or mutual fund holdings data, it can be estimated even on a higher frequency than monthly or quarterly, thus providing a much more realistic picture of the evolution of dispersion in expectations across time. Moreover, the Chicago Board Options Exchange (CBOE) provides freely on its website the intraday trading activity of option contracts, thus making it easy for investors to use the dispersion in options traders' beliefs measure for investment decisions. Fourth, it can equally accommodate optimistic and pessimistic beliefs since it is hardly influenced by the short-sale constraints that are present in the equity market and affect both individual and institutional investor

[1] Diether, Malloy and Scherbina, 2002; Park, 2005; Anderson, Ghysels and Juergens, 2005, 2009; Yu, 2011 and Buraschi, Trojani and Vedolin, 2014 among others utilize the dispersion in analysts' forecasts, while Chen, Hong and Stein, 2002; Goetzmann and Massa, 2005 and Jiang and Sun, 2014 create dispersion measures from mutual fund and individual investor portfolio holdings.

[2] The available range of strike prices for index options is determined by the underlying index fluctuations and the customer requests. CBOE Rule 24.9.04 specifies that typically strike prices for index options should be within 30% of the current index value, but even more extreme strike prices are permitted provided there is demonstrated customer demand.

[3] For example, the average number of forecasters in Anderson, Ghysels and Juergens (2009) is 36.

portfolio holdings.[4] Finally, it can explicitly distinguish between different levels of positive and negative expectations, while this is not straightforward in the case of dispersion measures derived from investors' positions in the equity market.

Our results establish a significant and robust negative relationship between the dispersion in options traders' beliefs and future market returns. This result allows for a dual interpretation: If the dispersion in options trading volume across strikes proxies for the level of disagreement in the underlying asset market, then this finding is in line with the models of Miller (1977) and Scheinkman and Xiong (2003), who show that in the presence of short-sale constraints asset prices reflect only the views of the most optimistic investors since pessimistic investors sit out of the market. Therefore, higher disagreement is accompanied by higher asset prices and lower subsequent returns. In the context of the aggregate market, the above limits-to-arbitrage explanation can be supported by the empirical findings of D'Avolio (2002) and Lamont and Stein (2004) who show that only a very limited fraction of the total stocks is actually sold short.

Alternatively, if we consider a framework wherein the underlying asset market participants have homogeneous beliefs and update their views by observing - to some extent - the trading activity in the options market,[5] then the dispersion in options trading volume across strikes can be regarded as a proxy for the representative investor's ambiguity about the true return generating model. In this respect, the range of strike prices with traded options can be interpreted as the set of all alternative models considered plausible, while the proportion of trading volume attributed to each strike price can be regarded as the probability attached to each model. This framework is intuitive since expected market returns are driven by aggregate macroeconomic factors that are easily observable and therefore the likelihood that different expectations - and hence trades - are induced by information asymmetry is rather low. In such a case, the documented negative relation is in line

[4] Almazan, Brown, Carlson and Chapman (2004) provide evidence showing that approximately only 3% of all mutual funds implement short-selling.

[5] In fact, option-implied sentiment indicators such as the put-call trading volume ratio are widely used by investors for investment allocation decisions.

with the recursive smooth ambiguity model of Hayashi and Miao (2011) and Ju and Miao (2012) in case of a preference for consumption smoothing over time, i.e. an elasticity of intertemporal substitution (EIS) lower than one. In particular, in this setting higher ambiguity increases the pricing kernel but also increases the demand for the risky asset since investors are willing to substitute current consumption with increased future consumption. This positive covariance between the pricing kernel and the risky asset return leads to a decreased equity premium.

The empirical results show that at the 1-month horizon, the suggested measure of dispersion in expectations is a strong predictor of future excess market returns under various model specifications (univariate and multivariate). Moreover, it outperforms in terms of predictive power all other predictors examined in prior literature apart from the variance risk premium (VRP), which explains a higher proportion of the variation in future returns. In addition, it offers additional predictability when combined with VRP in the same forecasting model, thus showing that the two variables have different information content and can be used complementarily for predicting future market returns.[6] The results from long-horizon regression analysis show that our dispersion in beliefs measure remains significant at all horizons and for horizons of 12 and 24 months ahead exhibits an adjusted R^2 higher than 10% outperforming the majority of the alternative predictors. This result is remarkable because unlike the other successful at long horizons predictors, the dispersion measure exhibits a relatively low persistence (about 0.50) hence alleviating potential concerns regarding spurious predictability.

The results of out-of-sample predictive analysis reveal that the dispersion of options traders' beliefs has significantly higher forecasting power than the historical mean and it outperforms all other predictive variables apart from VRP. Following

[6] Compared to the VRP, which has emerged as the primary option-implied return predictor, the dispersion in options traders' expectations is conceptually different. First, it is not extracted from option prices and therefore it is not derived from the risk-neutral distribution. Second, while the level of trading volume could potentially have an effect on option prices and subsequently on VRP in accordance with the limits to arbitrage hypothesis (Bollen and Whaley, 2004), high trading volume is not necessarily associated with high dispersion across different moneyness categories. Therefore, the market forces that influence VRP do not have an apparent effect on the dispersion in options trading volume across strike prices. In fact, the empirical analysis suggests that the correlation between the two measures is close to zero.

Campbell and Thompson (2008), imposing a constraint of positive forecasted equity premia leads to further improvement of the out-of-sample predictability of dispersion in beliefs for the excess market return.

The results are also economically significant since an active trading strategy based on the out-of-sample predictive power of the proposed dispersion in expectations measure offers increased utility to a mean-variance investor that would otherwise follow a passive buy-hold strategy. Moreover, in terms of economic significance dispersion in expectations outperforms almost all alternative predictors and when combined with the VRP it improves the performance of the trading strategy. Therefore, it is confirmed again that the dispersion of options traders' opinions and the VRP act as complementary variables and their joint use for investment decisions can prove very beneficial to an active investor. The performance of rotation strategies that rely on the out-of-sample predictive power of the dispersion in beliefs measure for several equity portfolio excess returns reveal that its information content is economically important not only for the aggregate market but also for the majority of the portfolios sorted on different stock characteristics.

Finally, we compare the dispersion in options trading volume across strike prices with other popular option-implied variables in order to alleviate potential concerns about the information embedded in our measure. More specifically, the alternative option-implied variables include the slope of the implied volatility smirk, the risk-neutral variance, skewness and kurtosis, and the out-of-the-money (OTM) puts to the at-the-money (ATM) calls open interest ratio proxying for investors' hedging pressure. Higher dispersion in options traders' beliefs is associated with higher variance, more negative skewness, higher kurtosis, more negatively sloped volatility smirk, and less hedging pressure. However, the highest correlation coefficient, which is the one between our dispersion in expectations variable and risk-neutral variance, is only 0.29 revealing that the suggested measure does not proxy for any type of variance or tail risk and is not driven by the well-known hedging demand for OTM puts. Bivariate and multivariate regression analysis confirms that in the presence of the alternative option-implied measures, the dispersion in options trading volume across strikes remains highly significant in forecasting subsequent market returns at all horizons.

The remainder of the chapter is structured as follows. Section 5.2 describes the data and the construction of the main variables used in the study. Section 5.3 provides the empirical evidence from in-sample regression analysis. Section 5.4 discusses the results from out-of-sample regression analysis. Section 5.5 presents the economic significance of the out-of-sample empirical evidence. Section 5.6 presents the comparison between the dispersion in options trading volume and other option-related variables. Finally, Section 5.7 concludes.

5.2 Variables Construction and Data

This section first describes the construction of our dispersion in expectations measure, then discusses the alternative predictors used in the study and finally provides some summary statistics.

5.2.1 Dispersion in options traders' beliefs

We construct a measure of dispersion in options traders' expectations by using trading volume information across strike prices. Since a trade on an option contract is a direct bet on the future asset price, the strike price at which the option is traded reveals a specific expectation about the asset return. For example, a trade on a high strike price option is indicative of an optimistic view about the future asset return, while a trade on a low strike price option is indicative of a pessimistic view about the future asset return. Motivated by this observation, we model the dispersion of expected market returns through the dispersion in the volume-weighted strike prices of option contracts on the Standard and Poors (S&P) 500 index. More specifically, we construct the following two measures that proxy for the dispersion of the expected returns distribution:

$$DISP = \sum_{j=1}^{K} w_j \left| X_j - \sum_{j=1}^{K} w_j X_j \right|, \tag{5.1}$$

$$DISP^* = \sqrt{\sum_{j=1}^{K} w_j \left(X_j - \sum_{j=1}^{K} w_j X_j \right)^2}, \tag{5.2}$$

where w_j is the proportion of the total trading volume attributed to the jth strike price X_j. DISP corresponds to the mean absolute deviation of the (volume-weighted) strike prices, while DISP* corresponds to the respective standard deviation. The two measures have the same information content but DISP* is always higher or equal to DISP due to Jensen's inequality.

To construct the above variables, we use S&P 500 index call and put options' volume data. Our sample period is 1996:01 to 2012:12 and for each month we estimate DISP and DISP* using options on the last trading day of the month with moneyness below 0.975 or above 1.025 and maturities between 10 and 360 calendar days.[7] We discard near-the-money options since they are possibly traded as part of straddles and strangles and therefore reflect investors' expectations about future market volatility and not returns (Ni, Pan and Poteshman, 2008). However, keeping such options in our sample provides results of very similar statistical significance with slightly lower coefficients of determination for long horizons. We consider options with maturities up to one year ahead since we want to capture investors' expectations regarding both short-term and long-term market returns. Consistent with this intuition, unreported results show that using a DISP (or DISP*) measure created solely by short-maturity options exhibits similar predictability for short horizons but has limited power for long horizons.

The dispersion in options traders' expectations can have two interpretations. First, it can be considered a proxy of the disagreement among participants in the underlying asset market similarly to Park (2005) and Yu (2011). Such a disagreement can have an impact on future asset returns either due to the existence of

[7]Our sample is restricted to the 1996:01 - 2012:12 period because in the pre-1996 period the relatively low liquidity of options is accompanied by very little variation in DISP and DISP*.

short-sale constraints (Miller, 1977; Scheinkman and Xiong, 2003) or due to risk-sharing effects that impact asset prices in equilibrium (Basak, 2000, 2005; Buraschi and Jiltsov, 2006). Second, it can be regarded as a measure of ambiguity, similarly to Anderson, Ghysels and Juergens (2009) and Drechsler (2013). In particular, considering a framework wherein participants in the underlying asset market have homogeneous beliefs which are driven at least partly by the trading activity in the options market, the dispersion in options traders' expectations can serve as proxy for the set of alternative return generating models that a representative investor is exposed to. In this case, a high (low) dispersion in options traders' opinions implies that it is highly likely that the participants in the underlying asset market exhibit high (low) ambiguity about the true return generating model. Since investors on average exhibit ambiguity aversion, i.e. they are averse to events with unknown probability distributions of all possible outcomes (Ellsberg, 1961), it is apparent that market discount rates should reflect investors' aversion not only to risk but also to ambiguity (Epstein and Wang, 1994; Chen and Epstein, 2002; Ju and Miao, 2012; Drechsler, 2013).

5.2.2 Other variables

We compare the predictive ability of the proposed dispersion in expectations measures with a set of variables that have been found in the literature to predict stock market returns. The main alternative predictor is the variance risk premium (VRP) which was introduced by Bollerslev, Tauchen and Zhou (2009) and has been the key variable in a series of recent studies that examine its predictive power for stock market returns (Drechsler and Yaron, 2011; Bollerslev, Marrone, Xu and Zhou, 2012; Zhou, 2012) or the cross-section of stock returns (Bali and Hovakimian, 2009 and Han and Zhou, 2011). VRP is defined as the difference between the expected 1-month ahead stock return variance under the risk neutral measure and the expected 1-month ahead variance under the physical measure.[8] When investors are more

[8] Following Bollerslev, Tauchen and Zhou (2009), we use the past 1-month realized variance as the expected 1-month ahead variance under the physical measure. Our results remain robust to

averse to future variance risk, they are willing to pay more in order to hedge against variance and therefore increase the VRP. Monthly VRP data are obtained from Hao Zhou's website.[9] Unlike VRP, the dispersion in options traders' expectations does not depend on the risk-neutral distribution extracted from option prices. Moreover, while an increased level of trading volume could be potentially associated with high buying pressure that would increase risk-neutral variance due to limits to arbitrage (Bollen and Whaley, 2004), there exists no obvious link between the dispersion in trading volume across strike prices and the resulting risk-neutral variance.

The rest of the predictor variables include the tail risk (TAIL, Kelly and Jiang, 2014), the aggregate dividend-price ratio (d-p, Fama and French, 1988 and Campbell and Shiller, 1988a,b), the market dividend-payout ratio (d-e, Campbell and Shiller, 1988a and Lamont, 1998), the yield gap (YG, Maio, 2013), the yield term spread (TERM, Campbell, 1987 and Fama and French, 1989), the default spread (DEF, Keim and Stambaugh, 1986 and Fama and French, 1989), the relative short-term risk free rate (RREL, Campbell, 1991) and the realized stock market variance (SVAR, Guo, 2006).[10] TAIL captures the probability of extreme negative market returns and is constructed by applying the Hill's (1975) estimator to the whole NYSE/AMEX/NASDAQ cross-section (share codes 10 and 11) of daily returns within a given month. d-p is the difference between the log aggregate annual dividends and the log level of the S&P 500 index, while d-e is the difference between the log aggregate annual dividends and the log aggregate annual earnings. YG is the difference between the aggregate earnings-price ratio and the 10-year bond yield, both in levels. TERM is the difference between the 10-year bond yield and the 1-year bond yield, while DEF is the difference between BAA and AAA corporate bonds yields from Moody's. Finally, RREL is the difference between the 3-month t-bill rate and its moving average over the preceding twelve months and SVAR is the monthly variance of the S&P 500 index. Data on monthly prices, dividends,

the choice of the VRP proxy.

[9] `https://sites.google.com/site/haozhoupersonalhomepage`

[10] We have also considered as alternative predictors the consumption-wealth ratio of Lettau and Ludvigson (2001) and the stock market illiquidity of Amihud (2002). The respective results are reported in Tables D.1 - D.3 of Appendix D.

and earnings are obtained from Robert Shiller's website.[11] All interest rate data are obtained from the FRED database of the Federal Reserve Bank of St. Louis. SVAR is downloaded from Amit Goyal's website.[12]

As a proxy for stock market returns we use the value-weighted index from the Chicago Center for Research in Security Prices (CRSP). In order to create a series of monthly excess stock market returns we subtract from the monthly log-return the (log of) the 1-month Treasury bill rate obtained from Kenneth French's website.[13] Longer horizons continuously compounded excess market returns are created by taking cumulative sums of monthly excess market returns.

5.2.3 Summary statistics

Figure 5.1 plots DISP along with VIX, a popular investor fear indicator capturing market forward-looking variance risk.[14] Both series are standardized for easier comparison. While the two series exhibit some common variation (the correlation coefficient is 29%) they tend to peak at different times. For example, unlike VIX, DISP is increasing but not very high during the 1997 Asian crisis and the 1998 Russian crisis, showing that there was no much divergence of opinions about the state of the economy during those periods. On the contrary, it exhibits several spikes during the period of the dot-com bubble showing that there were concerns about the very high stock market prices driven by the technology sector. In particular, DISP peaks in 2000:03 when NASDAQ reaches its all-time record high and the U.S. Federal Reserve increases the fed funds rate for a second time within two months, in 2000:09 when NASDAQ slightly recovers before it finally bursts, and finally in 2001:01 when the fed funds rate is decreased twice within one month just before the recession period begins. DISP also peaks in 2001:09 due to the 9/11 terrorist attack, in 2005:01 possibly due to the first concerns expressed by Robert Shiller regarding

[11] http://www.econ.yale.edu/~shiller/data.htm

[12] http://www.hec.unil.ch/agoyal

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

[14] The respective graph for DISP* is very similar and thus omitted.

the existence of a bubble in the US housing market[15] and in 2007:11 just before the beginning of the recent recession period. After the collapse of Lehman Brothers in 2008:09 it increases but not as extremely as VIX showing that given the apparently high risk in the market, there was no extreme dispersion in options traders' expectations. Finally, DISP substantially increases during the latest period of the European sovereign debt crisis and takes its all-time high value in 2012:03 after the Eurogroup agreement regarding the second bailout package for Greece, following the concerns about the success of the Private Sector Involvement (PSI) program.

Table 5.1 Panel A reports descriptive statistics about the dispersion in options traders' expectations measures and the alternative predictive variables, while Panel B presents the respective correlation coefficients. Both dispersion in expectations measures exhibit very similar statistics with slightly positive skewness and excess kurtosis. Unlike the majority of the alternative predictors, they are only moderately persistent with autocorrelation coefficients of 0.488 and 0.505 for DISP and DISP* accordingly. This mitigates the problem of potentially spurious regression results caused by highly persistent regressors (see Valkanov, 2003; Torous, Valkanov, and Yan, 2004; Boudoukh, Richardson, and Whitelaw, 2008). VRP has also a low autocorrelation coefficient of 0.210 and exhibits negative skewness and very large kurtosis. DISP and DISP* are very highly correlated (0.96) and close to uncorrelated with VRP (-0.03 and -0.07 for DISP and DISP* respectively) showing that the dispersion in options traders' beliefs contains different information from VRP. Finally, dispersion in beliefs is negatively correlated with TAIL and to a lesser extent with d-p and RREL, while being weakly positively correlated with YG, TERM, DEF, and SVAR. The respective correlations with d-e are very close to zero. Overall, both measures of dispersion in options traders' beliefs are not highly correlated with any of the alternative predictors, the biggest correlation occurring with TAIL.

[15] In anticipation of the publication of the second edition of Robert Shiller's "Irrational Exuberance", on January 25^{th} 2005 CNN Money publishes a feature on the possibility of a housing bubble in the US market, accompanied by an interview of Robert Shiller expressing his concerns.

5.3 In-Sample Predictability

In order to gauge the predictive power of our proposed dispersion in expectations measures, we run multiple-horizon regressions of excess stock market returns of the following form:

$$re_{t+h,h} = \alpha_h + \boldsymbol{\beta}_h' \mathbf{z}_t + \varepsilon_{t+h,h}, \tag{5.3}$$

where $re_{t+h,h} = \left(\frac{12}{h}\right)\left[re_{t+1} + re_{t+2} + ... + re_{t+h}\right]$ is the annualized h-month excess return of the CRSP value-weighted index and $\mathbf{z}_t$ is the vector of predictors. The regression analysis covers the period 1996:01-2012:12 and for each forecasting horizon we lose h observations. Under the null of no predictability the overlapping nature of the data imposes an $MA\left(h-1\right)$ structure to the error term $\varepsilon_{t+h,h}$ process. To overcome this problem we base our statistical inference on both Newey and West (1987) and Hodrick (1992) standard errors with lag length equal to the forecasting horizon. In general, the Hodrick (1992) standard errors tend to be more conservative, especially in long horizons when the null of no predictability is true (Ang and Bekaert, 2007) but have lower statistical power when the null is false (Bollerslev, Marrone, Xu and Zhou, 2012). The beta coefficients reported in the subsequent tables have been scaled and can be interpreted as the percentage annualized excess market returns caused by a one standard deviation increase in each regressor.

5.3.1 One-month ahead predictability

Table 5.2, Panel A provides the results for 1-month ahead univariate predictive regressions. The results show that the two dispersion in options traders' expectations measures are strong predictors of stock market excess returns as the null hypothesis of no predictability is rejected at 5% level based on both Newey-West and Hodrick standard errors. The slope estimates are negative and economically significant in both cases: a one standard deviation increase in DISP predicts a negative annualized market excess return of 9.68%, while a one standard deviation increase in DISP* leads to a negative annualized market excess return of 9.20%. The adjusted R^2, denoted by $\widetilde{R}^2$, is 2.23% and 1.97% for DISP and DISP* respectively. Turning to the rest of the predictor variables, VRP has a positive slope (16.09), which is significant at the 1% and 5% levels based on Newey-West and Hodrick standard errors, respectively. The corresponding forecasting ratio is relatively large (7.04%). None of the other variables is statistically significant at the 5% level (there is only marginal significance for both RREL and SVAR), a finding which is in line with Goyal and Welch's (2008) conclusion that most of the traditional predictors have performed poorly over the last decades. Moreover, the $\widetilde{R}^2$s of most of the alternative predictors are either negative or below 1% similar to Goyal and Welch (2008) and Campbell and Thompson (2008). Again, the exceptions are RREL and SVAR, which still deliver lower explanatory ratios than DISP.

Next, we assess the robustness of the significant results for DISP and DISP* to the presence of other predictive variables by conducting bivariate regressions. Panel B of Table 5.2 reports the results. The significance of both DISP and DISP* remains intact in all cases, showing that the information content of the dispersion in options traders' beliefs is distinct from that of other variables that have been used in the literature. It is also interesting to note, that the combination of DISP (DISP*) with VRP renders both variables strongly significant and increases $\widetilde{R}^2$ to 9.10% (8.51%), showing that the dispersion in options traders' beliefs and the variance risk premium are complementary measures and capture different features of investors' attitude. Since our dispersion in expectations measures and VRP appear to be the only successful predictors during our sample period, as a final robustness exercise we run trivariate predictive regressions considering combinations of DISP (or DISP*), VRP, and each of the other variables. Results in Panel C of Table 5.2 show that the dispersion in options trading volume across strike prices and VRP continue to be significant at either 1% or 5% level in almost all the cases. Moreover, now YG becomes also strongly significant with a positive predictive slope, in line with Maio (2013).

Overall, the results in Table 5.2 suggest that in our sample period only the dispersion in options trading volume across strikes and VRP are consistently successful in predicting excess market returns, and this predictive power is enhanced when they are combined in the same model.

The negative sign of the predictive slopes for DISP and DISP* shows that higher dispersion in options traders' expectations leads to lower future market excess returns. This can be interpreted in two ways. If the dispersion in options trading volume across strikes proxies for the level of disagreement in the equity market, the negative sign is in line with the models of Miller (1977) and Scheinkman and Xiong (2003). In particular, in the presence of short-sale constraints, the price of the asset is determined by the valuations of the most optimistic investors as pessimistic investors have no means to express their negative views and sit out of the market. Therefore, higher disagreement is associated with higher prices and subsequent lower returns. The above limits-to-arbitrage argument appears empirically well-founded as D'Avolio (2002) reports that only 7% of the total short-sale capacity is actually used and Lamont and Stein (2004) show that during the period 1995-2002 the ratio of the market value of shares sold short to the total value of shares outstanding was always below 4%.

If the dispersion in options trading volume across strikes is regarded as a proxy for the ambiguity of a representative equity market investor, then the negative sign can be explained in the context of the recursive smooth ambiguity model of Klibanoff, Marinacci and Mukerji (2005, 2009), Hayashi and Miao (2011) and Ju and Miao (2012) assuming an EIS lower than one. More specifically, in Ju and Miao's (2012) setting periods of high ambiguity are associated with an increased pricing kernel since investors are concerned about the probability of the true expected growth rate of the economy being lower than the one considered. If EIS is higher than one, ambiguity will lower the demand for the risky asset thus leading to a lower price and current period return. The negative covariance between the pricing kernel and the risky asset return will increase the equity premium. By contrast, if EIS is lower than one, ambiguity will increase the demand for hedging purposes thus leading to a higher price and current period return. The positive covariance between the pricing kernel and the risky asset return will therefore decrease the equity premium. The proper level of the representative investor's EIS constitutes one of the most long-lasting debates in the macro-finance literature (see Beeler and Campbell, 2012 and Bansal, Kiku and Yaron, 2012, for recent discussions on the topic).

5.3.2 Long-horizon predictability

Table 5.3 provides the results for 3-, 6-, 12- and 24-month ahead univariate predictive regressions. Both DISP and DISP* consistently forecast negative excess market returns and are significant at either 1% or 5% level in all cases apart from the 6-month horizon with Hodrick standard errors when both DISP and DISP* are significant at 10% level. The slope estimates continue to be economically significant as a one standard deviation increase in DISP (DISP*) predicts a negative annualized market excess return in the range of 5.19%-7.39% (5.10%-7.10%). In terms of fit, $\widetilde{R}^2$ stays between 3% and 4% for 3- and 6- month horizons but increases substantially for longer horizons and exceeds 10% and 12% for 12- and 24-month horizons respectively. This last result is of particular importance given the relatively low persistence of the proposed dispersion in beliefs variables. The only variables that exhibit higher $\widetilde{R}^2$ at the 24-month horizon are d-p, d-e, and TERM all of which have an autocorrelation coefficient higher than 0.98. Therefore, we conclude that DISP and DISP* can successfully capture divergence of opinions about both short and long horizon market returns.

Turning to the alternative predictors, VRP remains strongly significant for 3- and 6- month horizons with large $\widetilde{R}^2$s, yet its predictive power becomes less significant for longer horizons as in Bollerslev, Tauchen and Zhou (2009), Drechsler and Yaron (2011) and Bollerslev, Marrone, Xu and Zhou (2012). From the rest of the variables d-p, d-e, TERM, DEF, and SVAR become significant as the horizon increases with almost monotonically increasing $\widetilde{R}^2$s. Moreover, their Newey-West t-statistics are always considerably higher than the Hodrick t-statistics implying that in many of these cases the significance may have arisen spuriously due to the high persistence of the predictive variables (Ang and Bekaert, 2007).

Since the results in Table 5.3 suggest that only DISP, DISP*, and VRP exhibit a strong and consistent predictive pattern across all horizons, we proceed by examining trivariate 3-, 6-, 12- and 24-month ahead regressions considering combinations of DISP (or DISP*), VRP, and each of the other variables. The results reported in Table 5.4 suggest that the significance of DISP and DISP* follows the same pattern as in the univariate regressions. In particular, Panels A and B show that DISP and DISP* are significant in all the cases for the 3-month horizon and in all but one case (when dispersion in expectations and VRP are combined with RREL) for the 6-month horizon. The predictive slopes remain economically significant ranging from -3.46 to -10.10. In all the models considered, VRP continues to be strongly significant. Panels C and D show that for 12- and 24-month horizons both DISP and DISP* are again strongly significant in almost all the cases with economically significant slopes ranging from -3.15 to -8.56. As in the univariate analysis, the significance of VRP for 12- and 24-month horizons is weaker.

In summary, the empirical evidence regarding long-horizon predictability confirms that the dispersion in options traders' beliefs embeds important information about future excess market returns that is not included in any of the other variables considered. Moreover, a combination of the dispersion in beliefs and VRP can provide significant long-horizon predictive power for market returns.

5.4 Out-of-Sample Predictability

The results of the previous section provide convincing evidence that the dispersion in options traders' beliefs can significantly predict future excess stock market returns in-sample (IS). In this section, we evaluate the out-of-sample (OS) performance of our dispersion in beliefs measures following Lettau and Ludvigson (2001), Goyal and Welch (2003, 2008), Guo (2006), and Campbell and Thompson (2008) among others. The purpose of this exercise is to assess the usefulness of the dispersion in options trading volume across strike prices for an investor who has access only to real time data when making her forecasts and also to gauge regression parameter instability over time. Following the literature we mainly rely on OS regressions of 1-month horizon but for robustness purposes we also report results for 3- and 6-month horizons, keeping in mind the relatively low statistical power of OS regression analysis compared to IS analysis (Inoue and Kilian, 2004).

As in Goyal and Welch (2008), Campbell and Thompson (2008), Rapach, Strauss and Zhou (2010), and Ferreira and Santa-Clara (2011) we estimate the model in equation (5.3) recursively using the first $s = s_0 ... T-h$ observations and based on the estimated parameters we form our OS forecasts for the expected excess market return using the concurrent values of the predictor variables:

$$\widehat{re}_{s+h,h} = \widehat{\alpha}_{s,h} + \widehat{\boldsymbol{\beta}}'_{s,h}\mathbf{z}_s. \tag{5.4}$$

The initial estimation period is from 1996:01-1999:12 and the first prediction is made for 2000:01. This way we create a series of T_{OS} OS forecasts that is compared to a series of recursively estimated historical averages, which correspond to OS forecasts of a restricted model with only a constant as a regressor. We employ four measures to assess the OS predictability performance of our dispersion in expectations measures.

The first measure is the OS R^2 denoted by R^2_{OS} which takes the form:

$$R^2_{OS} = 1 - \frac{MSE_U}{MSE_R}, \tag{5.5}$$

where $MSE_U = \frac{1}{T_{OS}} \sum_{t=s}^{T-h} \left(re_{t+h,h} - \widehat{re}_{t+h,h}\right)^2$ is the mean square error of the unrestricted model and $MSE_R = \frac{1}{T_{OS}} \sum_{t=s}^{T-h} \left(re_{t+h,h} - \widetilde{re}_{t+h,h}\right)^2$ is the mean square error of the restricted model with $\widetilde{re}_{t+h,h}$ being the recursively estimated historical average. R^2_{OS} takes positive values whenever the unrestricted model outperforms the restricted model in terms of predictive power (i.e. $MSE_U < MSE_R$).

The second measure of OS performance is the F-test from McCracken (2007):

$$MSE - F = (T_{OS} - h + 1)\,\frac{MSE_R - MSE_U}{MSE_U}, \tag{5.6}$$

which tests whether MSE_U is statistically significantly lower than MSE_R.

The third OS performance test is the encompassing test of Clark and McCracken (2001):

$$ENC - NEW = \frac{(T_{OS} - h + 1)}{T_{OS}} \frac{\sum_{t=s}^{T-h}\left[(re_{t+h,h} - \widehat{re}_{t+h,h})^2 - (re_{t+h,h} - \widehat{re}_{t+h,h})\,(re_{t+h,h} - \widetilde{re}_{t+h,h})\right]}{MSE_U}, \tag{5.7}$$

which examines whether the restricted model encompasses the unrestricted model, meaning that the unrestricted model does not improve the forecasting ability of the restricted model. Statistical inference for the $MSE - F$ and the $ENC - NEW$ tests relies on the critical values derived by McCracken (2007) and Clark and McCracken (2001) using Monte Carlo simulations.

The final measure of OS forecasting performance is the constrained OS R^2 denoted by R^2_{C-OS} suggested by Campbell and Thompson (2008). This measure is the same with R^2_{OS} apart from the fact that it sets the OS forecasts of the unrestricted model equal to zero whenever they take negative values. Therefore, an investor's real time equity premium prediction becomes in accordance with standard asset pricing theory.

Table 5.5 presents the results for 1-, 3- and 6-month horizon OS predictability. In the case of 1-month horizon, DISP and DISP* exhibit positive R^2_{OS}s of 1.70% and 1.55% respectively. For both measures, the $MSE - F$ test rejects at 5% level the null hypothesis that the mean square error of the unrestricted model is equal to the mean square error of the restricted model while the $ENC - NEW$ test rejects at 5% level the null hypothesis that the restricted model encompasses the unrestricted model. When we impose the restriction of positive expected equity premium the results are improved for both dispersion in beliefs measures, with R^2_{C-OS} becoming 2.16% for

DISP and 2.44% for DISP*. Turning to the rest of the predictors, only VRP provides a positive R^2_{OS} of 7.96%. Moreover, the $MSE-F$ and $ENC-NEW$ tests strongly reject the respective null hypotheses at 5% level. Since univariate analysis suggests that only the dispersion in options trading volume across strikes and VRP have significant OS forecasting performance, we proceed by combining the two dispersion in options traders' beliefs measures with VRP. The results show that the bivariate models increase the R^2_{OS}, which becomes 9.06% in the regression including DISP and VRP and 8.56% in the case of DISP* and VRP confirming that the information content of the dispersion in options traders' expectations is different from that of VRP. Moreover, the $MSE-F$ and $ENC-NEW$ tests reject the respective null hypotheses even more decisively.

The results for the 3-month horizon are similar to those for the 1-month horizon but stronger for both dispersion in beliefs measures and the VRP. This is in line with the IS regression results presented in the previous section. In particular, DISP (DISP*) has an R^2_{OS} of 3.37% (3.04%) while VRP has an R^2_{OS} of 12.47%. The $MSE-F$ and the $ENC-NEW$ tests prodive even stronger evidence against the respective null hypotheses. As in the 1-month horizon analysis, apart from DISP, DISP* and VRP, none of the other predictors exhibit positive R^2_{OS}s. Furthermore, the bivariate model of the dispersion in options traders' expectations with VRP is even more successful in OS return predictability. The results for the 6-month horizon are in the same vein with the evidence from the other horizons. In particular, the R^2_{OS}s of DISP and DISP* remain positive while the $MSE-F$ and the $ENC-NEW$ tests still reject the respective null hypotheses at 5% level. Moreover, except for VRP none of the other alternative predictors provide a positive R^2_{OS}, while the combination of DISP (or DISP*) with VRP offers even stronger OS predictability.

Overall, the empirical evidence regarding OS return predictability suggests that only the dispersion in options traders' expectations and VRP are successful predictors at short horizons and that their predictive power is enhanced when they are combined in one bivariate model.

5.5 Economic Significance

In this section, we evaluate the economic significance of the information embedded in the dispersion in options traders' beliefs. In particular, following Goyal and Santa-Clara (2003), Campbell and Thompson (2008), Ferreira and Santa-Clara (2011), Rapach, Strauss, Tu and Zhou (2011) and Maio (2013, 2014) we create market-timing and portfolio rotation strategies that rely on the OS forecasting power of the suggested dispersion in expectations measures and the alternative predictors and evaluate their performance.

5.5.1 Market-timing strategy

We assess the economic significance of the dispersion in options traders' beliefs predictability by creating an active trading strategy that is based on its OS predictive power for 1-month ahead stock market excess returns. In particular, we follow the procedure described in the previous section and estimate a series of OS excess market return forecasts.[16] Then we consider two scenarios: one where short-sales are not allowed and one where short-sales are allowed. More specifically, in the first scenario we have:

$$\begin{aligned} a &= 1 \;\; if \;\; \widehat{re}_{t+1} \geq 0 \\ a &= 0 \;\; if \;\; \widehat{re}_{t+1} < 0, \end{aligned} \tag{5.8}$$

where a represents the portfolio weight attributed to the stock market index. In the second scenario we have:

$$\begin{aligned} a &= 1.5 \;\; if \;\; \widehat{re}_{t+1} \geq 0 \\ a &= -0.5 \;\; if \;\; \widehat{re}_{t+1} < 0. \end{aligned} \tag{5.9}$$

[16]In this section, the term return refers to arithmetic return and not to logarithmic return.

The realized return from the active trading strategy can be represented by:

$$R_{p,t+1} = aR_{m,t+1} + (1-a)\,R_{f,t+1}, \tag{5.10}$$

where $R_{m,t+1}$ denotes the arithmetic market return and $R_{f,t+1}$ denotes the return of the riskless asset. Therefore, following this procedure we create a series of realized portfolio returns based on the OS forecasting power of each forecasting variable and we compare the results with those from a buy-hold strategy. This strategy invests only in the market in case of the first scenario and allocates 150% to the market and -50% to the risk-free asset in case of the second scenario.

For each trading strategy, we estimate the mean portfolio return, the standard deviation, and its Sharpe ratio. Moreover, since the Sharpe ratio weights equally the mean and volatility of the portfolio returns, we follow Campbell and Thompson (2008), Ferreira and Santa-Clara (2011), and Maio (2014) and additionally create a certainty equivalent return in excess of the buy-hold strategy (ΔCER), assuming a mean-variance investor with risk aversion coefficient equal to three. ΔCER represents the change in investor's utility resulting from her choice to follow the active instead of the passive trading strategy.[17] As an additional performance measure we also estimate the maximum drawdown (MDD), which represents the maximum loss than an investor can incur if she enters the strategy at any-time during its implementation period. All measures apart from the MDD are in annualized terms. Finally, we also report the percentage of months that each active strategy goes long the stock market index.

The performance results from the strategies are presented in Table 5.6. When short sales are not allowed, the strategy associated with DISP exhibits a mean return of 5.46% while the one associated with DISP* has a relatively low mean return of

[17] ΔCER is more formally defined as:

$$\Delta CER = E\left(R_{p,t+1}\right) - E\left(\bar{R}_{p,t+1}\right) + \frac{\gamma}{2}\left[Var\left(\bar{R}_{p,t+1}\right) - Var\left(R_{p,t+1}\right)\right], \tag{5.11}$$

where γ is the risk aversion coefficient, $R_{p,t+1}$ is the portfolio return of the market-timing strategy and $\bar{R}_{p,t+1}$ is the portfolio return of the buy-hold strategy.

2.62%. Both strategies, however, exhibit remarkably low volatilities of 9.41% and 9.59% for DISP and DISP* respectively. This leads to annualized Sharpe ratios of 0.58 and 0.27 accordingly, both of which outperform the Sharpe ratio of the buy-hold strategy (0.23). Furthermore, the ΔCER of the strategy based on DISP is 4.46% per year and that of the strategy based on DISP* is 1.58%, thus showing that the utility provided by the active strategies related to the dispersion in options traders' beliefs is higher than the utility of the buy-hold strategy. In terms of ΔCER, DISP is only outperformed by VRP, while DISP* is also outperformed by d-e and DEF. This is because the dispersion in options traders' beliefs (and especially DISP*) goes long the risky asset in about only half of the periods thereby avoiding a lot of negative market return realizations, but also ignoring a few large positive spikes. In contrast, both d-e and DEF, despite their poor OS performance at the 1-month horizon, tend to invest much more in the market (in 82.69% and 77.56% of the months respectively), but also go long the riskless rate during the turbulent periods after the dot-com bubble and the Lehman Brothers collapse. Not surprisingly, the strategies based on DISP, DISP*, VRP, d-e and, DEF strategies also exhibit very low MDDs, with the one related to DISP having the lowest cumulative loss (-14.09%).

The most successful variable in terms of ΔCER (4.80%) is VRP. However, when we combine d-e and DEF in bivariate models with VRP the performance of the respective trading strategies deteriorates in comparison to the trading strategy based solely on VRP. The ΔCER in the bivariate model with d-e becomes 4.61%, while the ΔCER when we include DEF is only 1.87%. In contrast, when we combine either DISP or DISP* in bivariate models with VRP the performance of the respective trading strategies improves substantially in comparison to the strategy based solely on VRP. In particular, the ΔCER of the strategy related to the combination of DISP and VRP is 9.37%, while in the case of DISP* and VRP we obtain 7.97%. These results show that unlike d-e and DEF, the information content of the dispersion in options traders' expectations is significantly beneficial in economic terms for an investor who already uses the VRP in her investment decisions.

The pattern in the performance of the different trading strategies is very similar when short-sales are allowed, with the main difference being that extreme realizations (highly positive and highly negative market returns) have now a larger impact on the portfolio returns. The strategy associated with DISP strongly outperforms the passive strategy in terms of both Sharpe ratio and ΔCER, while the strategy associated with DISP* underperforms the buy-hold strategy in terms of Sharpe ratio but clearly outperforms it in terms of ΔCER. As in the first scenario, the ΔCER of the strategy based on DISP (8.93%) is only outperformed by the strategy associated with VRP (9.63%), while the ΔCER of the strategy based on DISP* (3.15%) is also outperformed by the strategies associated with d-e and DEF (8.83% and 4.66% respectively). However, when the trading strategies rely on bivariate models with VRP as the common variable, the strategies based on combinations of VRP and either d-e or DEF offer a lower ΔCER than those based only on VRP (9.24% and 3.74% accordingly), while the opposite is true for combinations of VRP with either DISP or DISP* (18.90% and 16.05% accordingly).

In summary, the empirical evidence associated with a market-timing strategy shows that the OS forecasting ability of the dispersion in options' traders beliefs for future market returns is economically significant, especially for an investor who already considers the information from VRP for her investment decisions.

5.5.2 Portfolio rotation strategies

We further explore the economic importance of the information embedded in the dispersion in options traders' expectations by creating rotation strategies based on its OS predictive power for the 1-month ahead excess returns of portfolios sorted on different stock characteristics. We first create a series of OS excess portfolio return forecasts similar to the previous section. Next, if the highest fitted excess return is positive or equal to zero we allocate 150% to the two portfolios with the highest excess forecasted returns, whilst if the highest fitted excess return is negative we allocate 150% to the risk-free rate. The rotation strategies always go short 50%

the two portfolios with the lowest forecasted excess returns. In essence, an investor following such a rotation strategy short sells the two portfolios that are expected to perform worst and invests either in the two portfolios that are expected to perform best or in the risk-free rate in case none of the portfolios is expected to have a positive excess return in the following month. Therefore, the realized return of the portfolio rotation strategy can be represented by:

$$R_{p,t+1} = \begin{cases} 0.75R_{H1,t+1} + 0.75R_{H2,t+1} - 0.25R_{L1,t+1} - 0.25R_{L2,t+1}, & if \quad \widehat{re}_{H1,t+1} \geq 0 \\ 1.5R_{f,t+1} - 0.25R_{L1,t+1} - 0.25R_{L2,t+1}, & if \quad \widehat{re}_{H1,t+1} < 0, \end{cases} \tag{5.12}$$

where $R_{H1,t+1}$ and $R_{H2,t+1}$ ($R_{L1,t+1}$ and $R_{L2,t+1}$) stand for the realized arithmetic returns of the portfolios with the highest and second highest (lowest and second lowest) fitted excess return and $\widehat{re}_{H1,t+1}$ stands for the highest fitted excess portfolio return.

For the purposes of the rotation strategies, we use decile portfolios sorted on size (Size), book-market ratio (B/M), momentum (Mom), industry (Industry), long-term reversal (LT Reversal) and short-term reversal (ST Reversal). Moreover, we consider a rotation strategy that uses all 60 portfolios simultaneously (Pooled). Data on portfolio returns are obtained from Kenneth French's website. The performance of the rotation strategies is compared to a simple buy-hold strategy that invests 150% in the market and shorts 50% in the risk-free asset. As in the case of the market-timing strategy, for each rotation strategy we estimate the mean return, standard deviation, Sharpe ratio, certainty equivalent return in excess of the buy-hold strategy (ΔCER) and maximum drawdown (MDD). Finally, we also report the percentage of months that each rotation strategy goes long the two highest fitted excess return portfolios.

The performance results of the rotation strategies associated with DISP and DISP* are shown in Table 5.7, Panel A. In terms of Sharpe ratio, DISP and DISP* outperform the buy-hold strategy in the case of Size, LT Reversal, ST Reversal and Pooled portfolios while DISP also outperforms the buy-hold strategy in the case of Mom portfolios. The corresponding Sharpe ratio values vary between 0.27 and

0.54 compared to a ratio of 0.19 associated with the passive strategy. Moreover, in all but one case (B/M portfolios for DISP*) the ΔCER values are positive, ranging from 0.54% to 9.54%, showing that the rotation strategies based on the dispersion in options traders' beliefs increase the utility of an investor who would otherwise follow a passive trading strategy. Finally, the MDDs of the rotation strategies are lower (in absolute value) than the respective buy-hold strategy MDD in all cases, with values varying between -63.31% and -27.02% compared to a value of -67.86% for the passive strategy. Panel B of Table 5.7 presents the results of the rotation strategies associated with the alternative predictors in the case of Pooled portfolios. In terms of Sharpe ratio only VRP and TERM outperform the 0.19 buy-hold strategy Sharpe ratio, exhibiting ratios of 0.64 and 0.27 respectively. In terms of ΔCER, apart from VRP, which offers a large positive ΔCER of 10.74%, none of the other predictors offers additional utility to a mean-variance investor.

Overall, the empirical results of portfolio rotation strategies show that in the majority of the cases the OS forecasting ability of the dispersion in options traders' beliefs for various portfolio excess returns is economically significant. When considering all portfolio categories simultaneously only our dispersion in beliefs measures and VRP yield economically significant results.

5.6 Comparison with Option-Implied Measures

The empirical evidence presented in the previous sections suggests that the dispersion in options trading volume across strike prices has significant IS and OS predictive power for future market excess returns and its information content is distinct from and complementary to that of VRP. However, one might still argue that our dispersion in expectations measures are driven by the well documented "volatility smirk" anomaly (Rubinstein, 1994, Jackwerth and Rubinstein, 1996), the hedging demand for OTM puts (Bollen and Whaley, 2004, Gârleanu, Pedersen and Poteshman, 2009), or that they just proxy for common variance risk captured by VIX. To alleviate such concerns, in this section we compare the dispersion in options traders' beliefs and its predictive power with a set of popular option-implied variables. The first variable is the slope of the implied volatility curve measured as the difference between the (volume-weighted) implied volatility of OTM puts and that of ATM calls (Slope; Xing, Zhang and Zhao, 2010, Atilgan, Bali and Demirtas, 2014). The second variable is the ratio of open interest of OTM puts to the open interest of ATM calls which proxies for hedging pressure in the S&P 500 index options market (HP; Han, 2008). The last three variables are the second, third, and fourth risk-neutral moments (VIX, Skewness and Kurtosis; Ang, Hodrick, Xing and Zhang, 2006, Chang, Christoffersen and Jacobs, 2013).[18]

Panel A of Table 5.8 reports the correlation coefficients between the two proposed dispersion in beliefs measures and the other option-implied variables. While DISP (DISP*) displays some common variation with all the other variables, the maximum (absolute value) correlation is 0.29 (0.25) showing that the information embedded in the dispersion in options traders' expectations is unique and is not subsumed by any other option-implied measure studied in the literature. In general, higher DISP and DISP* values are related to higher implied volatility, more negative skewness, higher kurtosis, and a more negatively sloped implied volatility curve. Moreover, DISP and DISP* are negatively correlated with HP showing that in periods of high demand for portfolio insurance there is less divergence of opinions about expected returns since the majority of the traders anticipate negative jumps.

Panel B of Table 5.8 shows the results of 1-month ahead bivariate predictive regressions with DISP or DISP* as the main variables. In all the models considered, both DISP and DISP* remain significant at either 5% or 1% level based on both Newey-West and Hodrick standard errors. Moreover, none of the other option-related measures exhibits significant predictive ability at the 1-month horizon. Panel C of Table 5.8 considers the case of multivariate regressions with all the option-implied variables being included into the predictive regression. The results show

[18] Risk-neutral moments are calculated using the model-free method of Bakshi, Kapadia and Madan (2003). The estimated implied volatility has a correlation of 99.7% with VIX and thus for better comparability with other studies we proceed by keeping VIX as our proxy for risk-neutral volatility.

that DISP and DISP* are strongly significant while again all the other variables remain highly insignificant.

The multivariate analysis is extended in Table 5.9 for horizons of 3, 6, 12, and 24 months ahead. The results reveal that both DISP and DISP* exhibit significant forecasting power for all horizons examined at either 5% or 1% level irrespectively of which standard errors are considered. Turning to the rest of the predictors, only VIX appears to be consistently and strongly significant, predicting positive excess market returns for all horizons longer than a quarter ahead. In the case of the 24-month horizon, Slope becomes significant at the 5% level, while Skewness and Kurtosis appear significant at the 5% level only when statistical inference is based on Newey-West standard errors.

Overall, the results of this section suggest that the dispersion in options trading volume across strikes is not highly related to other well-established option-implied variables that proxy for hedging demand, crash risk, or variance risk and its predictive power for excess market returns remains intact in the presence of such measures.

5.7 Conclusion

In this chapter, we develop a measure of dispersion in options traders' expectations about future stock returns by utilizing dispersion in trading volume information across various strike prices. A high dispersion implies that there is little consensus in the options market about the future underlying asset return, whereas a low dispersion suggests that options traders' beliefs are similar. Our dispersion in beliefs measure relies on the expectations about future returns represented by the trading activity in highly liquid options markets, is associated directly with asset prices (and not with a related indicator such as corporate earnings), can be estimated even on a higher frequency than monthly or quarterly, and is by construction able to capture different levels of both optimistic and pessimistic views.

We provide empirical evidence for a strong and robust negative relation between the dispersion in S&P 500 index options trading volume across strike prices and subsequent market returns. Moreover, the relatively low autocorrelation coefficient of our measure alleviates the common concern of potentially spurious regression results stemming from a highly persistent predictive variable. In-sample analysis shows that at the 1-month horizon, the dispersion in options traders' expectations compares favorably to the well-established variance risk premium (VRP) and clearly outperforms all other alternative predictors examined. At longer horizons, it remains significant and exhibits a high adjusted R^2, outperformed only by highly persistent variables. Most importantly, the forecasting power of the dispersion in options traders' expectations remains intact at all horizons when VRP and other predictors are added into the predictive model. It is therefore evident that its information content is different from that of VRP and the two variables can be used complementarily for forecasting purposes.

The results of out-of-sample analysis reveal that the dispersion in options traders' beliefs has significantly higher predictive power than the simple historical average and its forecasting ability can be enhanced by imposing a restriction of positive forecasted equity premia. Apart from VRP, none of the other alternative predictors examined can improve the simple historical average model. The out-of-sample forecasting power of the dispersion in options traders' beliefs is also economically significant, as indicated by the additional utility offered to an investor who follows an active trading strategy associated with its predictive ability. Unlike other variables, the suggested dispersion in beliefs measure also improves the performance of a market-timing strategy based solely on VRP, when it is added into the predictive model. Furthermore, the results of portfolio rotation strategies reveal that it exhibits economically significant out-of-sample forecasting power for the excess returns of portfolios sorted on several stock characteristics.

We also investigate the relationship of the dispersion in options traders' expectations with other popular option-implied variables and show that it does not proxy for any of them. More specifically, dispersion in options traders' expectations is associated with higher implied volatility, more negative skewness, higher kurtosis, a more negatively sloped implied volatility curve, and less hedging pressure. However, its correlation with these variables lies between 2% and 29% showing that the information content of the proposed dispersion in beliefs measure is largely distinct. Most importantly, a regression analysis confirms that dispersion in options traders' expectations remains highly significant at all horizons when combined with the other option-related variables.

The documented significant and negative relationship between the dispersion in options trading volume across strike prices and subsequent market returns allows for a dual interpretation: It is possible that the divergence of options traders' beliefs proxies for the disagreement among investors in the equity market. In such a case, a higher disagreement leads to a higher current price and lower subsequent returns due to the existence of short-sale constraints that prevent pessimistic investors from taking negative positions. Alternatively, it is possible that the dispersion of options traders' expectations proxies for the level of ambiguity of a representative equity market participant whose expectations are affected by the trading activity in the options market. In such a case, the negative relationship implies that when ambiguity is high, the representative investor increases her hedging demand for the risky asset and lowers her expected return due to her preference for consumption smoothing.

Table 5.1: Descriptive statistics and correlation coefficients of forecasting variables

Panel A: Descriptive Statistics											
	DISP	**DISP***	**VRP**	**TAIL**	**d-p**	**d-e**	**YG**	**TERM**	**DEF**	**RREL**	**SVAR**
Mean	99.827	132.985	19.306	0.412	-4.047	-0.860	0.000	0.014	0.010	-0.002	0.004
Median	100.098	130.693	16.601	0.415	-4.039	-0.983	-0.003	0.014	0.009	-0.001	0.002
Maximum	220.223	319.493	116.521	0.501	-3.324	1.379	0.054	0.034	0.034	0.011	0.058
Minimum	26.898	34.010	-180.679	0.290	-4.503	-1.244	-0.032	-0.004	0.006	-0.027	0.000
Std. Dev.	32.659	42.940	22.697	0.038	0.229	0.491	0.022	0.011	0.005	0.008	0.006
Skewness	0.376	0.504	-2.579	-0.528	0.276	2.943	0.736	0.096	2.779	-0.818	5.910
Kurtosis	3.634	4.490	33.137	3.366	3.595	12.093	2.755	1.651	12.374	3.762	48.803
Auto	0.488	0.505	0.210	0.612	0.983	0.983	0.998	0.984	0.961	0.970	0.695

Panel B: Correlation Coefficients											
	DISP	**DISP***	**VRP**	**TAIL**	**d-p**	**d-e**	**YG**	**TERM**	**DEF**	**RREL**	**SVAR**
DISP	1.00										
DISP*	0.96	1.00									
VRP	-0.03	-0.07	1.00								
TAIL	-0.43	-0.41	0.09	1.00							
d-p	-0.13	-0.12	-0.07	-0.06	1.00						
d-e	0.05	0.01	0.21	-0.09	0.51	1.00					
YG	0.15	0.22	-0.19	-0.11	0.41	-0.40	1.00				
TERM	0.14	0.14	0.05	-0.02	0.38	0.37	0.22	1.00			
DEF	0.26	0.23	-0.03	-0.29	0.66	0.74	0.09	0.41	1.00		
RREL	-0.17	-0.12	-0.07	0.00	-0.18	-0.44	0.21	-0.35	-0.41	1.00	
SVAR	0.24	0.22	-0.36	-0.43	0.34	0.38	-0.08	0.14	0.59	-0.33	1.00

This table reports descriptive statistics (Panel A) and correlation coefficients (Panel B) of the forecasting variables used in the study. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), tail risk (TAIL), dividend-price ratio (d-p), dividend payout ratio (d-e), yield gap (YG), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL) and realized stock market variance (SVAR). The sample period is 1996:01-2012:12.

Table 5.2: 1-month horizon predictability

	Panel A: Univariate		Panel B: Bivariate						Panel C: Trivariate							
		$\tilde{R}^2$ (%)	**DISP**	**Z**	$\tilde{R}^2$ (%)	**DISP***	**Z**	$\tilde{R}^2$ (%)	**DISP**	**VRP**	**Z**	$\tilde{R}^2$ (%)	**DISP***	**VRP**	**Z**	$\tilde{R}^2$ (%)
DISP	-9.68	2.23														
	(-2.36)**															
	[-2.44]**															
DISP*	-9.20	1.97														
	(-2.29)**															
	[-2.35]**															
VRP	16.09	7.04	-9.27	15.85	9.10	-8.14	15.53	8.51								
	(4.66)***		(-2.47)**	(4.34)***		(-2.24)**	(4.28)***									
	[2.49]**		[-2.35]**	[2.46]**		[-2.11]**	[2.41]**									
TAIL	4.18	0.01	-9.72	-0.10	1.75	-9.02	0.43	1.48	-10.01	15.98	-1.67	8.70	-8.44	15.58	-0.74	8.06
	(0.93)		(-2.28)**	(-0.02)		(-2.21)**	(0.09)		(-2.46)**	(4.30)***	(-0.40)		(-2.17)**	(4.23)***	(-0.18)	
	[0.85]		[-2.51]**	[-0.02]		[-2.39]**	[0.09]		[-2.57]**	[2.47]**	[-0.33]		[-2.25]**	[2.42]**	[-0.15]	
d-p	6.14	0.60	-9.01	4.91	2.44	-8.56	5.05	2.21	-8.43	16.29	6.10	9.71	-7.30	16.02	6.30	9.19
	(1.04)		(-2.09)**	(0.82)		(-2.04)**	(0.85)		(-2.21)**	(4.53)***	(1.12)		(-2.00)**	(4.45)***	(1.15)	
	[1.12]		[-2.19]**	[0.88]		[-2.11]**	[0.90]		[-2.06]**	[2.52]**	[1.09]		[-1.83]*	[2.48]**	[1.13]	
d-e	1.87	-0.40	-9.81	2.37	1.91	-9.22	1.93	1.59	-9.22	16.06	-1.00	8.67	-8.11	15.82	-1.37	8.10
	(0.31)		(-2.39)**	(0.39)		(-2.28)**	(0.32)		(-2.46)**	(4.51)***	(-0.17)		(-2.24)**	(4.48)***	(-0.23)	
	[0.34]		[-2.47]**	[0.43]		[-2.36]**	[0.35]		[-2.34]**	[2.40]**	[-0.17]		[-2.11]**	[2.37]**	[-0.24]	
YG	4.30	0.04	-10.54	5.85	2.73	-10.60	6.57	2.68	-10.57	17.53	9.14	10.96	-10.09	17.21	9.68	10.60
	(1.11)		(-2.60)***	(1.46)		(-2.64)***	(1.61)		(-2.88)***	(4.51)***	(2.46)**		(-2.82)***	(4.47)***	(2.58)**	
	[1.05]		[-2.63]***	[1.42]		[-2.64]***	[1.56]		[-2.64]***	[2.62]***	[2.08]**		[-2.53]**	[2.58]**	[2.17]**	
TERM	0.82	-0.48	-10.01	2.27	1.89	-9.51	2.17	1.61	-9.49	15.78	1.45	8.70	-8.32	15.46	1.27	8.10
	(0.21)		(-2.40)**	(0.58)		(-2.34)**	(0.55)		(-2.45)**	(4.36)***	(0.38)		(-2.23)**	(4.29)***	(0.33)	
	[0.21]		[-2.45]**	[0.57]		[-2.37]**	[0.55]		[-2.33]**	[2.44]**	[0.36]		[-2.10]**	[2.39]**	[0.32]	
DEF	-2.31	-0.34	-9.76	0.28	1.75	-9.17	-0.15	1.48	-9.43	15.86	0.59	8.65	-8.14	15.53	0.00	8.05
	(-0.33)		(-2.28)**	(0.04)		(-2.21)**	(-0.02)		(-2.33)**	(4.32)***	(0.09)		(-2.10)**	(4.26)***	(0.00)	
	[-0.37]		[-2.34]**	[0.04]		[-2.26]**	[-0.02]		[-2.27]**	[2.46]**	[0.09]		[-2.03]**	[2.41]**	[0.00]	
RREL	7.35	1.08	-8.70	5.90	2.74	-8.45	6.35	2.64	-8.06	16.41	7.22	10.12	-7.18	16.16	7.69	9.75
	(1.57)		(-2.15)**	(1.30)		(-2.15)**	(1.40)		(-2.14)**	(4.18)***	(1.70)*		(-1.99)**	(4.13)***	(1.82)*	
	[1.77]*		[-2.18]**	[1.41]		[-2.17]**	[1.53]		[-2.04]**	[2.52]**	[1.69]*		[-1.87]*	[2.48]**	[1.81]*	
SVAR	-9.40	2.08	-7.86	-7.49	3.29	-7.48	-7.74	3.15	-8.87	15.24	-1.72	8.71	-7.67	14.71	-2.36	8.19
	(-1.66)*		(-1.99)**	(-1.29)		(-1.99)**	(-1.36)		(-2.38)**	(3.22)***	(-0.29)		(-2.14)**	(3.09)***	(-0.41)	
	[-1.23]		[-2.05]**	[-0.97]		[-2.03]**	[-1.01]		[-2.32]**	[2.44]**	[-0.22]		[-2.08]**	[2.36]**	[-0.31]	

This table reports the results of 1-month ahead predictive regressions for the excess return on the CRSP value-weighted index. The sample period is 1996:01-2012:12. Panel A reports the results of univariate regressions, Panel B the results of bivariate regressions and Panel C the results of trivariate regressions. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), tail risk (TAIL), dividend-price ratio (d-p), dividend payout ratio (d-e), yield gap (YG), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL) and realized stock market variance (SVAR). Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance in 1%, 5% and 10% level.

Table 5.3: Univariate long-horizon predictability

	h=3		h=6		h=12		h=24	
		$\widetilde{\mathbf{R}}^2$ (%)		$\widetilde{\mathbf{R}}^2$ (%)		$\widetilde{\mathbf{R}}^2$ (%)		$\widetilde{\mathbf{R}}^2$ (%)
DISP	-7.39	3.61	-5.47	3.54	-6.67	10.89	-5.19	12.72
	(-2.59)**		(-2.25)**		(-2.90)***		(-2.35)**	
	[-2.33]**		[-1.83]*		[-2.38]**		[-2.02]**	
DISP*	-7.10	3.30	-5.10	3.00	-6.65	10.80	-5.20	12.74
	(-2.65)***		(-2.24)**		(-3.03)***		(-2.49)**	
	[-2.31]**		[-1.81]*		[-2.40]**		[-2.11]**	
VRP	13.05	12.32	8.34	8.90	3.94	3.46	2.90	3.58
	(5.25)***		(3.96)***		(2.40)**		(2.29)**	
	[3.62]***		[3.33]***		[2.06]**		[1.77]*	
TAIL	-1.21	-0.39	-1.48	-0.22	2.22	0.74	3.58	5.75
	(-0.38)		(-0.49)		(0.74)		(1.25)	
	[-0.32]		[-0.43]		[0.77]		[1.32]	
d-p	6.70	2.88	7.67	7.44	8.23	16.84	8.32	33.56
	(1.28)		(1.84)*		(2.72)***		(4.46)***	
	[1.25]		[1.50]		[1.76]*		[1.96]*	
d-e	3.42	0.38	4.38	2.09	4.28	4.17	6.10	17.79
	(0.73)		(1.23)		(2.26)**		(3.25)***	
	[0.67]		[0.99]		[1.17]		[2.18]**	
YG	3.79	0.58	2.96	0.67	2.28	0.81	0.77	-0.27
	(1.13)		(0.96)		(0.77)		(0.22)	
	[0.95]		[0.76]		[0.61]		[0.23]	
TERM	1.01	-0.43	1.82	-0.06	4.03	3.64	7.55	27.54
	(0.28)		(0.54)		(1.38)		(3.27)***	
	[0.26]		[0.46]		[1.05]		[2.35]**	
DEF	-0.27	-0.50	2.07	0.07	3.30	2.27	4.22	8.22
	(-0.04)		(0.45)		(1.23)		(2.00)**	
	[-0.05]		[0.38]		[0.78]		[1.27]	
RREL	7.63	3.89	7.65	7.39	6.53	10.41	-0.58	-0.40
	(1.95)*		(1.94)*		(1.59)		(-0.33)	
	[1.82]*		[1.71]*		[1.42]		[-0.18]	
SVAR	-4.88	1.29	0.49	-0.48	2.37	0.92	2.48	2.48
	(-1.21)		(0.20)		(2.05)**		(2.64)***	
	[-0.93]		[0.10]		[0.77]		[1.19]	

This table reports the results of 3-, 6-, 12- and 24-month ahead univariate predictive regressions for the excess return on the CRSP value-weighted index. The sample period is 1996:01-2012:12. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), tail risk (TAIL), dividend-price ratio (d-p), dividend payout ratio (d-e), yield gap (YG), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL) and realized stock market variance (SVAR). Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance in 1%, 5% and 10% level.

Table 5.4: Trivariate long-horizon predictability

	Panel A: 3-month horizon								Panel B: 6-month horizon							
	DISP	**VRP**	**Z**	$\widetilde{R}^2$ (%)	**DISP***	**VRP**	**Z**	$\widetilde{R}^2$ (%)	**DISP**	**VRP**	**Z**	$\widetilde{R}^2$ (%)	**DISP***	**VRP**	**Z**	$\widetilde{R}^2$ (%)
TAIL	-10.10	13.42	-6.93	18.12	-8.75	13.01	-6.08	16.68	-7.94	8.72	-5.93	15.52	-6.82	8.39	-5.26	13.82
	(-3.74)***	(5.87)***	(-2.61)***		(-3.43)***	(5.70)***	(-2.34)**		(-3.27)***	(4.55)***	(-2.07)**		(-3.02)***	(4.39)***	(-1.89)*	
	[-3.31]***	[3.70]***	[-1.78]*		[-2.99]***	[3.61]***	[-1.56]		[-2.95]***	[3.58]***	[-1.83]*		[-2.74]***	[3.46]***	[-1.61]	
d-p	-6.09	13.32	6.72	18.56	-5.30	13.13	6.87	17.88	-4.15	8.72	7.62	19.47	-3.46	8.59	7.74	18.76
	(-2.34)**	(6.10)***	(1.67)*		(-2.25)**	(6.07)***	(1.70)*		(-2.21)**	(5.00)***	(2.22)**		(-2.04)**	(4.97)***	(2.25)**	
	[-1.85]*	[3.70]***	[1.22]		[-1.69]*	[3.65]***	[1.26]		[-1.36]	[3.63]***	[1.47]		[-1.21]	[3.61]***	[1.50]	
d-e	-7.12	12.61	1.18	15.31	-6.25	12.43	0.90	14.43	-5.45	7.56	3.13	13.01	-4.63	7.43	2.90	11.88
	(-2.99)***	(5.37)***	(0.29)		(-2.77)***	(5.31)***	(0.21)		(-2.40)**	(3.85)***	(0.88)		(-2.14)**	(3.82)***	(0.81)	
	[-2.27]**	[3.44]***	[0.23]		[-2.06]**	[3.40]***	[0.17]		[-1.85]*	[2.70]***	[0.67]		[-1.66]*	[2.66]***	[0.62]	
YG	-8.11	14.23	7.59	19.36	-7.84	13.99	8.04	18.93	-6.01	9.20	5.49	15.64	-5.68	9.02	5.79	15.02
	(-3.49)***	(4.37)***	(2.51)**		(-3.55)***	(4.31)***	(2.63)***		(-3.04)***	(3.47)***	(2.02)**		(-2.97)***	(3.43)***	(2.10)**	
	[-2.49]**	[3.87]***	[1.82]*		[-2.46]**	[3.82]***	[1.90]*		[-1.95]*	[3.61]***	[1.34]		[-1.93]*	[3.57]***	[1.40]	
TERM	-7.26	12.78	1.45	15.36	-6.42	12.54	1.32	14.50	-5.59	8.09	2.24	12.41	-4.86	7.91	2.13	11.38
	(-3.03)***	(4.83)***	(0.46)		(-2.84)***	(4.81)***	(0.41)		(-2.50)**	(3.74)***	(0.74)		(-2.24)**	(3.70)***	(0.69)	
	[-2.23]**	[3.55]***	[0.36]		[-2.05]**	[3.49]***	[0.33]		[-1.82]*	[3.22]***	[0.54]		[-1.69]*	[3.16]***	[0.52]	
DEF	-7.59	12.90	2.06	15.51	-6.61	12.63	1.60	14.56	-6.29	8.27	3.91	13.68	-5.36	8.06	3.49	12.35
	(-3.35)***	(5.26)***	(0.51)		(-3.02)***	(5.14)***	(0.39)		(-2.79)***	(4.52)***	(1.15)		(-2.43)**	(4.38)***	(1.01)	
	[-2.33]**	[3.59]***	[0.33]		[-2.08]**	[3.53]***	[0.26]		[-2.08]**	[3.37]***	[0.70]		[-1.87]*	[3.28]***	[0.63]	
RREL	-5.75	13.45	7.65	19.51	-5.23	13.27	7.97	19.11	-3.96	8.79	7.61	19.38	-3.55	8.67	7.84	18.98
	(-2.60)**	(3.99)***	(2.60)***		(-2.45)**	(3.95)***	(2.66)***		(-1.56)	(3.06)***	(2.15)**		(-1.53)	(3.04)***	(2.23)**	
	[-1.83]*	[3.69]***	[1.79]*		[-1.73]*	[3.65]***	[1.87]*		[-1.32]	[3.38]***	[1.65]		[-1.26]	[3.34]***	[1.71]*	
SVAR	-7.49	13.54	1.89	15.43	-6.50	13.09	1.34	14.49	-6.68	10.35	5.94	15.64	-5.66	9.95	5.43	14.09
	(-3.19)***	(4.43)***	(0.59)		(-2.90)***	(4.17)***	(0.41)		(-2.85)***	(5.81)***	(2.99)***		(-2.53)**	(5.65)***	(2.76)***	
	[-2.39]**	[3.40]***	[0.33]		[-2.13]**	[3.32]***	[0.23]		[-2.26]**	[3.99]***	[1.13]		[-2.03]**	[3.90]***	[1.04]	

(to be continued)

(continued)

	Panel C: 12-month horizon								Panel D: 24-month horizon							
	DISP	**VRP**	**Z**	$\tilde{R}^2$ (%)	**DISP***	**VRP**	**Z**	$\tilde{R}^2$ (%)	**DISP**	**VRP**	**Z**	$\tilde{R}^2$ (%)	**DISP***	**VRP**	**Z**	$\tilde{R}^2$ (%)
TAIL	-7.46	3.97	-1.78	14.34	-7.13	3.62	-1.50	13.52	-4.55	2.64	1.21	16.04	-4.40	2.42	1.36	15.62
	(-3.88)***	(2.43)**	(-0.68)		(-3.98)***	(2.21)**	(-0.59)		(-2.69)***	(1.85)*	(0.51)		(-2.72)***	(1.69)*	(0.58)	
	[-3.11]***	[2.19]**	[-0.73]		[-3.03]***	[2.00]**	[-0.61]		[-2.25]**	[1.70]*	[0.55]		[-2.37]**	[1.54]	[0.62]	
d-p	-5.21	4.30	7.53	27.87	-4.99	4.07	7.55	27.28	-3.33	3.27	7.76	43.82	-3.15	3.13	7.78	43.25
	(-3.18)***	(2.86)***	(3.06)***		(-3.05)***	(2.71)***	(2.99)***		(-3.55)***	(2.16)**	(6.80)***		(-3.52)***	(2.04)**	(6.40)***	
	[-1.82]*	[2.36]**	[1.59]		[-1.77]*	[2.28]**	[1.59]		[-1.28]	[2.15]**	[1.80]*		[-1.27]	[2.09]**	[1.80]*	
d-e	-7.00	2.93	4.29	18.26	-6.66	2.68	4.05	17.12	-6.02	1.33	6.61	35.89	-5.76	1.11	6.45	34.41
	(-3.27)***	(1.70)*	(1.95)*		(-3.14)***	(1.55)	(1.80)*		(-3.72)***	(1.01)	(3.69)***		(-3.47)***	(0.82)	(3.67)***	
	[-2.52]**	[1.39]	[1.13]		[-2.44]**	[1.28]	[1.06]		[-2.37]**	[0.75]	[2.26]**		[-2.37]**	[0.63]	[2.21]**	
YG	-6.75	4.47	3.45	16.69	-6.78	4.22	3.80	16.65	-5.01	2.97	0.98	15.93	-4.93	2.78	1.17	15.54
	(-3.17)***	(2.18)**	(1.36)		(-3.31)***	(2.08)**	(1.46)		(-2.46)**	(1.82)*	(0.33)		(-2.47)**	(1.72)*	(0.39)	
	[-2.39]**	[2.43]**	[0.91]		[-2.42]**	[2.33]**	[1.00]		[-2.00]**	[2.43]**	[0.30]		[-2.05]**	[2.32]**	[0.35]	
TERM	-7.25	3.57	4.83	19.64	-7.06	3.25	4.79	18.90	-6.01	2.16	8.06	47.31	-5.71	1.91	7.90	45.59
	(-3.53)***	(2.14)**	(2.05)**		(-3.55)***	(1.93)*	(1.93)*		(-4.18)***	(1.61)	(4.35)***		(-3.89)***	(1.39)	(4.16)***	
	[-2.52]**	[1.87]*	[1.23]		[-2.51]**	[1.72]*	[1.23]		[-2.29]**	[1.32]	[2.46]**		[-2.29]**	[1.17]	[2.42]**	
DEF	-7.95	3.90	5.38	20.74	-7.52	3.55	5.02	19.29	-6.58	2.86	5.93	31.93	-6.24	2.58	5.66	30.10
	(-3.82)***	(2.61)***	(2.78)***		(-3.62)***	(2.36)**	(2.57)**		(-3.64)***	(2.06)**	(4.31)***		(-3.40)***	(1.85)*	(4.11)***	
	[-2.84]***	[2.08]**	[1.25]		[-2.73]***	[1.91]*	[1.17]		[-2.59]**	[1.78]*	[1.78]*		[-2.58]**	[1.60]	[1.71]*	
RREL	-5.49	4.25	5.78	22.03	-5.50	4.01	6.01	22.15	-5.42	2.63	-1.50	16.55	-5.22	2.41	-1.26	15.65
	(-2.01)**	(1.93)*	(1.46)		(-2.17)**	(1.84)*	(1.54)		(-2.50)**	(1.92)*	(-0.88)		(-2.47)**	(1.74)*	(-0.74)	
	[-1.99]**	[2.09]**	[1.23]		[-2.03]**	[1.98]**	[1.28]		[-2.16]**	[1.51]	[-0.47]		[-2.17]**	[1.39]	[-0.39]	
SVAR	-8.56	6.42	7.14	24.16	-8.07	5.90	6.71	22.45	-6.71	4.94	6.12	30.38	-6.41	4.55	5.86	28.75
	(-3.90)***	(3.82)***	(4.23)***		(-3.79)***	(3.81)***	(4.31)***		(-3.85)***	(3.05)***	(3.77)***		(-3.57)***	(3.04)***	(3.91)***	
	[-3.00]***	[3.02]***	[2.12]**		[-2.87]***	[2.84]***	[1.99]**		[-2.64]***	[2.55]**	[2.66]***		[-2.65]***	[2.37]**	[2.55]**	

This table reports the results of 3- (Panel A), 6- (Panel B), 12- (Panel C) and 24-month (Panel D) ahead trivariate predictive regressions for the excess return on the CRSP value-weighted index. The sample period is 1996:01-2012:12. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), tail risk (TAIL), dividend-price ratio (d-p), dividend payout ratio (d-e), yield gap (YG), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL) and realized stock market variance (SVAR). Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance in 1%, 5% and 10% level.

Table 5.5: Out-of-sample predictability

	1-month horizon				3-month horizon				6-month horizon			
	R^2_{OS} (%)	MSE-F	ENC-NEW	R^2_{C-OS} (%)	R^2_{OS} (%)	MSE-F	ENC-NEW	R^2_{C-OS} (%)	R^2_{OS} (%)	MSE-F	ENC-NEW	R^2_{C-OS} (%)
DISP	1.70	2.69**	5.25**	2.16	3.37	5.29**	6.16**	4.93	1.39	2.05**	5.16**	7.19
DISP*	1.55	2.46**	5.04**	2.44	3.04	4.77**	5.99**	4.82	1.09	1.61**	4.76**	6.36
VRP	7.96	13.50**	11.27**	5.61	12.47	21.66**	22.44**	13.22	8.76	14.02**	14.59**	13.05
TAIL	-1.39	-2.14	-0.36	-0.47	-5.43	-7.83	-2.89	-4.44	-9.99	-13.26	-4.72	-7.62
d-p	-0.71	-1.10	0.88	-0.54	-3.81	-5.58	0.59	-3.90	-2.88	-4.09	3.72**	-2.93
d-e	-7.60	-11.01	-0.63	-2.13	-35.49	-39.81	-8.54	-8.24	-83.24	-66.33	-17.59	-42.80
YG	-1.79	-2.74	4.68**	-1.90	-4.23	-6.17	9.62**	-4.99	-7.11	-9.69	9.12**	-7.86
TERM	-2.71	-4.12	-0.73	-1.70	-12.27	-16.61	-4.85	-9.96	-32.81	-36.07	-9.07	-28.44
DEF	-4.91	-7.30	2.46**	-0.34	-24.50	-29.91	0.58	1.78	-66.88	-58.51	-8.17	5.79
RREL	-0.68	-1.05	0.32	-0.41	-3.83	-5.61	0.38	-3.99	-14.19	-18.14	-1.91	-13.57
SVAR	-5.39	-7.97	1.97*	-4.75	-13.05	-17.55	-5.83	-10.32	-7.21	-9.81	-3.51	-3.57
DISP & VRP	9.06	15.54**	17.14**	4.07	15.15	27.14**	30.56**	12.36	11.27	18.55**	22.24**	15.78
DISP* & VRP	8.56	14.60**	16.54**	4.10	14.34	25.45**	29.78**	12.13	10.36	16.87**	21.33**	15.12

This table reports the results of 1-, 3- and 6-month ahead out-of-sample predictability for the excess return on the CRSP value-weighted index. The total sample period is 1996:01-2012:12 and the forecasting period begins in 2000:01. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), tail risk (TAIL), dividend-price ratio (d-p), dividend payout ratio (d-e), yield gap (YG), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL) and realized stock market variance (SVAR). R^2_{OS} is the out-of-sample coefficient of determination, MSE-F is the McCracken (2007) F-statistic, ENC-NEW is the encompassing test of Clark and McCracken (2001) and R^2_{C-OS} is the out-of-sample coefficient of determination when the out-of-sample prediction is restricted to be positive. ** and * denote significance in 5% and 10% level.

Table 5.6: Market-timing strategy

	Mean (%)	**St. Dev. (%)**	**Sharpe**	**ΔCER (%)**	**MDD (%)**	**Long (%)**
			Panel A: No Short Sales			
Buy & Hold	3.88	16.75	0.23		-51.44	
DISP	5.46	9.41	0.58	4.46	-14.09	55.77
DISP*	2.62	9.59	0.27	1.58	-24.06	52.56
VRP	6.93	12.81	0.54	4.80	-32.24	55.77
TAIL	2.50	13.26	0.19	0.20	-46.24	83.97
d-p	3.99	16.44	0.24	0.27	-51.44	96.79
d-e	6.39	12.44	0.51	4.40	-27.70	82.69
YG	3.39	15.23	0.22	0.25	-51.44	89.10
TERM	1.63	14.31	0.11	-1.11	-47.30	84.62
DEF	4.43	12.75	0.35	2.33	-33.35	77.56
RREL	3.53	14.38	0.25	0.76	-52.82	88.46
SVAR	3.99	15.03	0.27	0.94	-44.87	91.03
DISP & VRP	10.98	11.38	0.96	9.37	-19.95	56.41
DISP* & VRP	9.57	11.34	0.84	7.97	-19.95	51.92
VRP & d-e	6.17	11.23	0.55	4.61	-25.48	69.87
VRP & DEF	3.34	10.98	0.30	1.87	-31.20	50.64
			Panel B: Short Sales			
Buy & Hold	4.76	25.15	0.19		-67.86	
DISP	7.91	15.74	0.50	8.93	-29.02	55.77
DISP*	2.25	15.98	0.14	3.15	-48.50	52.56
VRP	10.85	19.93	0.54	9.63	-41.02	55.77
TAIL	2.00	20.57	0.10	0.39	-64.12	83.97
d-p	4.98	24.73	0.20	0.54	-69.24	96.79
d-e	9.78	19.47	0.50	8.83	-40.59	82.69
YG	3.79	23.14	0.16	0.50	-73.00	89.10
TERM	0.27	21.93	0.01	-2.21	-68.21	84.62
DEF	5.87	19.92	0.29	4.66	-48.66	77.56
RREL	4.06	22.02	0.18	1.52	-72.65	88.46
SVAR	4.99	22.87	0.22	1.88	-62.26	91.03
DISP & VRP	18.97	17.89	1.06	18.90	-22.74	56.41
DISP* & VRP	16.14	17.93	0.90	16.05	-22.74	51.92
VRP & d-e	9.34	17.95	0.52	9.24	-35.94	69.87
VRP & DEF	3.69	17.68	0.21	3.74	-50.54	50.64

This table reports the results of market-timing strategies based on the 1-month ahead out-of-sample predictability for the excess return on the CRSP value-weighted index. The total sample period is 1996:01-2012:12 and the forecasting period begins in 2000:01. Panel A shows the results when short sales are not allowed and Panel B when short sales are allowed. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), tail risk (TAIL), dividend-price ratio (d-p), dividend payout ratio (d-e), yield gap (YG), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL) and realized stock market variance (SVAR). Buy & Hold refers to a passive strategy that goes long the market portfolio. Mean denotes the average return, St. Dev. denotes the standard deviation of returns, Sharpe stands for the Sharpe ratio, ΔCER is the certainty equivalent return in excess of the buy-hold strategy, MDD stands for the maximum drawdown and Long is the percentage of months that the strategy goes long the market index. All measures of performance apart from MDD are in annualized terms.

Table 5.7: Portfolio rotation strategies

	Mean (%)	St. Dev. (%)	Sharpe	ΔCER (%)	MDD (%)	Long (%)
	Panel A: Portfolio Strategies for DISP and DISP*					
Buy & Hold	4.76	25.15	0.19		-67.86	
DISP						
Size	10.02	20.56	0.49	8.41	-39.47	76.92
B/M	2.00	17.97	0.11	1.90	-51.42	87.18
Mom	5.31	20.02	0.27	4.03	-43.06	86.54
Industry	3.53	22.70	0.16	0.54	-62.41	100.00
LT Reversal	8.51	21.72	0.39	6.17	-43.69	92.95
ST Reversal	8.62	18.07	0.48	8.46	-45.72	84.62
Pooled	8.11	24.93	0.33	3.52	-51.61	100.00
DISP*						
Size	8.90	20.62	0.43	7.26	-34.01	75.64
B/M	-0.58	17.36	-0.03	-0.36	-50.25	86.54
Mom	3.45	19.76	0.17	2.33	-40.77	83.97
Industry	3.24	21.98	0.15	0.73	-63.31	100.00
LT Reversal	6.73	22.24	0.30	4.04	-46.64	93.59
ST Reversal	9.47	17.64	0.54	9.54	-27.02	82.69
Pooled	8.12	23.82	0.34	4.34	-43.69	100.00
	Panel B: Pooled Portfolio Strategies for Alternative Predictors					
VRP	18.51	28.87	0.64	10.74	-57.34	100.00
TAIL	1.54	25.81	0.06	-3.72	-73.21	100.00
d-p	3.38	33.60	0.10	-8.83	-76.54	100.00
d-e	5.50	30.43	0.18	-3.65	-75.11	96.79
YG	4.22	29.39	0.14	-4.00	-72.34	100.00
TERM	8.07	29.57	0.27	-0.31	-64.15	100.00
DEF	2.57	28.56	0.09	-4.93	-76.33	96.15
RREL	-3.08	27.85	-0.11	-9.98	-82.69	100.00
SVAR	4.09	29.22	0.14	-3.99	-68.04	96.15

This table reports the results of portfolio rotation strategies based on the 1-month ahead out-of-sample predictability for the excess stock portfolio returns. The total sample period is 1996:01-2012:12 and the forecasting period begins in 2000:01. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), tail risk (TAIL), dividend-price ratio (d-p), dividend payout ratio (d-e), yield gap (YG), yield term spread (TERM), default spread (DEF), relative short-term risk-free rate (RREL) and realized stock market variance (SVAR). Buy & Hold refers to a passive strategy that goes long the market portfolio. Mean denotes the average return, St. Dev. denotes the standard deviation of returns, Sharpe stands for the Sharpe ratio, ΔCER is the certainty equivalent return in excess of the buy-hold strategy, MDD stands for the maximum drawdown and Long is the percentage of months that the strategy goes long the two winner portfolios. All measures of performance apart from MDD are in annualized terms. The rotation strategies use decile portfolios sorted on size (Size), book-market ratio (B/M), momentum (Mom), industry (Industry), long-term reversal (LT Reversal) and short-term reversal (ST Reversal). Pooled refers to a rotation strategy that uses all 60 portfolios. Panel A shows the results of the rotation strategies based on the forecasting performance of the dispersion in options' traders expectations, while Panel B shows the results of the rotation strategies using all 60 portfolios based on the forecasting performance of the alternative predictors.

Table 5.8: Comparison with other option-implied measures - 1-month horizon

Panel A: Correlation Coefficients

	Slope	HP	VIX	Skewness	Kurtosis
DISP	0.18	-0.24	0.29	-0.02	0.05
DISP*	0.19	-0.25	0.23	-0.09	0.12

Panel B: Bivariate Regressions

	Slope	HP	VIX	Skewness	Kurtosis
DISP	-9.95	-9.80	-10.89	-9.67	-9.68
	(-2.39)**	(-2.27)**	(-2.73)***	(-2.35)**	(-2.34)**
	[-2.44]**	[-2.37]**	[-2.78]***	[-2.43]**	[-2.42]**
Z	1.53	-0.51	4.14	0.72	0.00
	(0.34)	(-0.12)	(0.64)	(0.17)	(0.00)
	[0.33]	[-0.12]	[0.70]	[0.18]	[0.00]
$\widetilde{\mathbf{R}}^2$ **(%)**	1.81	1.75	2.20	1.76	1.75
DISP*	-9.48	-9.33	-9.97	-9.20	-9.28
	(-2.30)**	(-2.17)**	(-2.59)**	(-2.24)**	(-2.24)**
	[-2.34]**	[-2.26]**	[-2.63]***	[-2.30]**	[-2.29]**
Z	1.52	-0.52	3.30	0.07	0.64
	(0.33)	(-0.12)	(0.52)	(0.02)	(0.16)
	[0.33]	[-0.12]	[0.56]	[0.02]	[0.17]
$\widetilde{\mathbf{R}}^2$ **(%)**	1.55	1.49	1.78	1.48	1.49

Panel C: Multivariate Regressions

DISP	Slope	HP	VIX	Skewness	Kurtosis	$\widetilde{\mathbf{R}}^2$ (%)
-11.51	3.07	0.25	3.90	9.75	8.54	0.52
(-2.65)***	(0.50)	(0.06)	(0.55)	(0.86)	(0.87)	
[-2.72]***	[0.50]	[0.06]	[0.59]	[0.84]	[0.90]	
DISP*	**Slope**	**HP**	**VIX**	**Skewness**	**Kurtosis**	$\widetilde{\mathbf{R}}^2$ **(%)**
-10.59	2.48	0.15	3.27	8.26	8.00	0.05
(-2.41)**	(0.41)	(0.03)	(0.47)	(0.73)	(0.82)	
[-2.49]**	[0.41]	[0.03]	[0.50]	[0.72]	[0.85]	

This table reports the results of 1-month ahead predictive regressions for the excess return on the CRSP value-weighted index. The sample period is 1996:01-2012:12. Panel A reports the correlation coefficients, Panel B the results of bivariate regressions and Panel C the results of multivariate regressions. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), slope of the implied volatility curve (Slope), hedging pressure (HP), implied volatility (VIX), risk-neutral skewness (Skewness) and risk-neutral kurtosis (Kurtosis). Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance in 1%, 5% and 10% level.

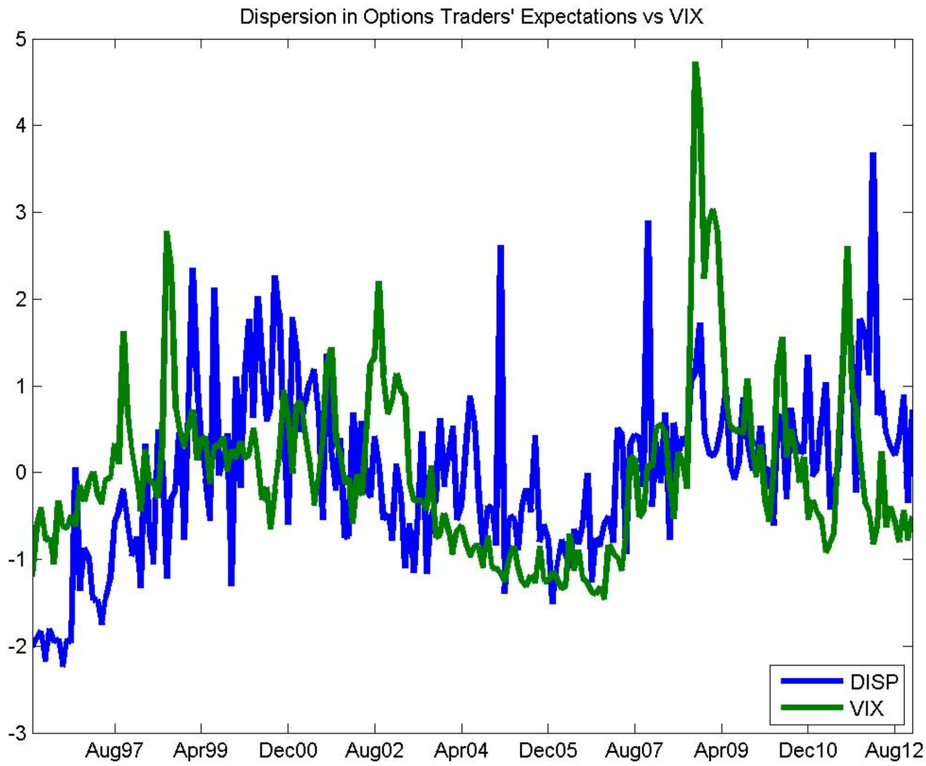
Table 5.9: Comparison with other option-implied measures - long horizons

	DISP	**Slope**	**HP**	**VIX**	**Skewness**	**Kurtosis**	$\widetilde{\mathbf{R}}^2$ (%)
h=3	-8.66	2.29	3.40	6.06	-5.93	-4.16	6.89
	(-3.09)***	(0.66)	(1.36)	(1.04)	(-0.67)	(-0.45)	
	[-2.58]**	[0.56]	[1.20]	[1.15]	[-0.71]	[-0.56]	
h=6	-7.63	1.64	2.59	7.95	-0.34	2.87	11.48
	(-3.00)***	(0.71)	(1.34)	(2.89)***	(-0.06)	(0.44)	
	[-2.53]**	[0.54]	[1.09]	[1.77]*	[-0.06]	[0.53]	
h=12	-8.68	1.33	2.06	7.07	2.86	5.67	23.11
	(-4.01)***	(1.11)	(1.56)	(3.29)***	(0.52)	(0.86)	
	[-3.14]***	[0.67]	[1.06]	[2.36]**	[0.58]	[1.21]	
h=24	-7.65	3.23	0.82	6.08	9.73	8.34	33.23
	(-4.35)***	(2.54)**	(0.49)	(2.82)***	(2.57)**	(2.38)**	
	[-3.06]***	[2.42]**	[0.53]	[2.47]**	[1.63]	[1.46]	

	DISP*	**Slope**	**HP**	**VIX**	**Skewness**	**Kurtosis**	$\widetilde{\mathbf{R}}^2$ (%)
h=3	-8.01	1.87	3.32	5.60	-7.03	-4.55	6.25
	(-2.88)***	(0.55)	(1.31)	(0.97)	(-0.79)	(-0.49)	
	[-2.43]**	[0.46]	[1.16]	[1.07]	[-0.84]	[-0.61]	
h=6	-6.88	1.24	2.56	7.50	-1.33	2.47	10.29
	(-2.73)***	(0.54)	(1.26)	(2.75)***	(-0.22)	(0.38)	
	[-2.43]**	[0.42]	[1.06]	[1.67]*	[-0.22]	[0.46]	
h=12	-8.33	0.93	1.84	6.70	1.79	5.34	22.19
	(-3.98)***	(0.77)	(1.35)	(3.22)***	(0.32)	(0.81)	
	[-3.05]***	[0.48]	[0.95]	[2.24]**	[0.36]	[1.14]	
h=24	-7.20	2.94	0.67	5.70	8.94	8.01	30.93
	(-4.43)***	(2.52)**	(0.39)	(2.75)***	(2.28)**	(2.25)**	
	[-3.06]***	[2.33]**	[0.43]	[2.32]**	[1.51]	[1.40]	

This table reports the results of 3-, 6-, 12- and 24-month ahead multivariate predictive regressions for the excess return on the CRSP value-weighted index. The sample period is 1996:01-2012:12. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), slope of the implied volatility curve (Slope), hedging pressure (HP), implied volatility (VIX), risk-neutral skewness (Skewness) and risk-neutral kurtosis (Kurtosis). Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance in 1%, 5% and 10% level.

Figure 5.1: Dispersion in options traders' expectations vs VIX



This figure plots the monthly time series of DISP versus VIX for the period 1996:01-2012:12. Both variables have been standardized to have zero mean and variance one.

Chapter 6

Conclusion

6.1 Limitations and Future Research

This thesis is related to a large literature that investigates the properties of the options market and its relationship with the rest of the economy. A common characteristic of this strand of the literature is that the conclusions drawn are subject to the limited availability of options data. In particular, unlike equity, fixed income or macroeconomic data which are largely available for the last fifty years, options data are typically available only for half of that period. In our case, the available options data begin in 1990 and therefore our analysis cannot be extended beyond that year. Furthermore, some of the option-related measures require a high amount of traded options across both the moneyness and the time-to-maturity dimension in order to be constructed accurately. Thus, the relatively low liquidity of the S&P 500 index options market in the early nineties has restricted some parts of the analysis to the post-1996 period. It is important to note, however, that even our truncated dataset spans a quite large period of seventeen years (1996-2012) encompassing both bull and bear markets and two major crises.

Despite the thorough analysis conducted in the previous chapters and the respective appendices, additional investigation may shed more light to the empirical findings of the thesis. Such a supplementary analysis can examine further the robustness of our results as well as provide a deeper understanding of the main conclusions drawn. In that respect, some of the issues discussed below can serve as ideas for future research.

Chapter 3 documents a significant relationship between the economic fundamentals sentiment component and the S&P 500 index risk-neutral skewness. The direction of this relationship is consistent across all three sentiment proxies in the first period. However, in the second period the way the sentiment proxy associated with large speculators responds to the macroeconomic conditions changes completely and hence the relationship of its economic fundamentals component with skewness becomes opposite to the relationship documented for the respective components of the other two sentiment proxies. Understanding the causes of the changing behavior of this particular sentiment proxy is beyond the scope of the thesis but constitutes an interesting research question that can be examined in subsequent projects. Furthermore, while our conclusions are based on a mature options market with a popular equity index as its underlying asset, it would be intriguing to investigate whether we draw the same conclusions if we examine less developed index options markets or options markets on different underlying assets such as treasuries or currencies.

In Chapter 4, we show that forward skewness coefficients are jointly important for predicting future macroeconomic and financial conditions. One concern that can arise from our analysis is that the high cross-correlations between the forward moments reduces our ability to provide a meaningful interpretation of the individual regression coefficients. In principle this is true but we need to underline that the aim of this chapter is to examine whether it is valuable to take into consideration the information embedded in the whole term structure of the risk-neutral skewness or not and hence it focuses on the joint significance of the combined forward skewness coefficients. In fact all variables of each forward moments group share a strong common component that is related to the cross-section of options across moneyness. It is possible to isolate this common component by orthogonalizing the forward moments of two, three and four months ahead to the respective forward moment of one month ahead. Such an exercise changes the regression coefficients of all forward moments and the standard error of the one month ahead forward moments. However, the rest of the individual and joint significance results remain unaltered. Therefore, the results presented in this thesis are those with the original forward moments and the reported regression coefficients show the effect of each forward moment controlling for the effect of the rest of the variables. Furthermore, given that that the empirical results of Chapter 4 are based on in-sample regressions, it would be particularly interesting to extend the analysis to an out-of-sample setting and also investigate whether the joint predictive power of forward skewness coefficients for future market returns is economically significant.

Finally, Chapter 5 implicitly assumes that all option trades reflect investors' expectations about future market returns. To this end, we remove from the analysis near-the-money options that can be more easily related to investors' beliefs about future volatility. It is possible, however, that still some portion of the trading activity is driven by some sort of "clientele effect". While we cannot rule out this possibility, it is important to note that there is no evidence in the literature relating empirically the trading activity in the options market with motives other than investors' expectations about the future distribution of the underlying asset returns. As further analysis, it would be interesting to elaborate more on the effect of the dispersion in options traders' expectations measure by investigating whether it depends on the underlying economic conditions. Furthermore, while our analysis is focused on the relationship between the suggested dispersion measure and future market returns, it would be interesting to complete the empirical evidence by investigating also the relationship between dispersion and contemporaneous returns.

6.2 Summary and Implications

Overall, this thesis is mainly related to two strands of the options markets literature. First it contributes to the literature that investigates the determinants of the shape the risk-neutral distribution. Second, it contributes to the literature that explores the information content of option prices and options' trading volume or open interest. Furthermore, it has implications about the consumption-based asset pricing literature in general, while many of the results can be considered useful for regulators and investors as well.

More specifically, Chapter 3 provides evidence that in recent years option prices are mainly settled according to investors' expectations stemming from the economic conditions and not to investors' errors in beliefs. This is a remarkable result for the asset pricing literature as it implies that the pricing kernel, which incorporates investors' risk preferences and can be considered the unifying link for all asset markets, is not driven by any sort of irrational beliefs. Second, it is important for the option pricing literature as it demonstrates that modelling investors' irrationality is not likely to improve the performance of the existing option pricing models. Moreover, the result that the economic fundamentals sentiment component has an opposite impact on calls and puts gives further credence to the notion that call and put options markets are segmented.

Chapter 4 suggests the usage of measures of forward skewness coefficients for predicting future macroeconomic and stock market conditions as well as systemic risk and equity uncertainty. In that respect, the reported results are of interest not only to academics but also to regulators and investors. In particular, the information embedded in the time-to-maturity dimension of option prices can be used to provide signals about required policy actions to be taken such as a looser monetary policy or a more relaxed regulation towards banks' capital requirements. Moreover, it can be used in the context of a market-timing strategy together with more traditional predictors of future market returns such as market valuation ratios.

Finally, Chapter 5 proposes a novel, easy-to-implement, yet theoretically founded measure of dispersion in expectations and shows that it is a strong predictor of future market returns both in-sample and out-of-sample. The above result has two main implications. First, it reveals a new dimension of the information embedded in the trading activity in the options market. It is also shown empirically that the predictability of dispersion in options traders' beliefs can be remarkably beneficial for investment strategies. Second, it provides a new measure of dispersion in expectations that exhibits several advantageous characteristics compared to previously proposed measures.

Bibliography

Abken, P., D. Madan, and S. Ramamurtie. 1996. Estimation of risk-neutral and statistical densities by hermite polynomial approximation: with an application to eurodollar futures options. *Working Paper, Federal Reserve Bank of Atlanta.*

Ai, H. 2010. Information quality and long-run risk: asset pricing implications. *Journal of Finance* 65, 1333–1367.

Äijö, J. 2008. Impact of US and UK macroeconomic news announcements on the return distribution implied by FTSE-100 index options. *International Review of Financial Analysis* 17, 242–258.

Aït-Sahalia, Y. and A. W. Lo. 1998. Nonparametric estimation of state-price densities implied in financial asset prices. *Journal of Finance* 53, 499–547.

Aït-Sahalia, Y., Y. Wang, and F. Yared. 2001. Do option markets correctly price the probabilities of movement of the underlying asset? *Journal of Econometrics* 102, 67–110.

Allen, L., T. G. Bali, and Y. Tang. 2012. Does systemic risk in the financial sector predict future economic downturns? *Review of Financial Studies* 25, 3000–3036.

Almazan, A., K. C. Brown, M. Carlson, and D. A. Chapman. 2004. Why constrain your mutual fund manager? *Journal of Financial Economics* 73, 289–321.

Amihud, Y. 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5, 31–56.

Amin, K., J. D. Coval, and H. N. Seyhun. 2004. Index option prices and stock market momentum. *Journal of Business* 77, 835–874.

Anagnou-Basioudis, I., M. Bedendo, S. Hodges, and R. Tompkins. 2005. Forecasting accuracy of implied and GARCH-based probability density functions. *Review of Futures Markets* 11, 41–66.

Anderson, E. W., E. Ghysels, and J. L. Juergens. 2005. Do heterogeneous beliefs matter for asset pricing? *Review of Financial Studies* 18, 875–924.

Anderson, E. W., E. Ghysels, and J. L. Juergens. 2009. The impact of risk and uncertainty on expected returns. *Journal of Financial Economics* 94, 233–263.

Andreou, P. C. 2013. Effects of market default risk on index options risk-neutral moments. *Working Paper, Cyprus University of Technology.*

Ang, A. and G. Bekaert. 2007. Stock return predictability: is it there? *Review of Financial Studies* 20, 651–707.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61, 259–299.

Aparicio, S. D. and S. Hodges. 1998. Implied risk-neutral distribution: a comparison of estimation methods. *Working Paper, University of Warwick.*

Atilgan, Y., T. G. Bali, and K. O. Demirtas. 2014. Implied volatility spreads and expected market returns. *Journal of Business and Economic Statistics*, forthcoming.

Bahra, B. 1997. Implied risk-neutral probability density functions from option prices: theory and application. *Working Paper, Bank of England.*

Bai, J. and S. Ng. 2002. Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.

Bailey, W., L. Zheng, and Y. Zhou. 2014. What makes the VIX tick? *Working Paper, Cornell University.*

Baker, M. and J. Wurgler. 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61, 1645–1680.

Baker, S. R., N. Bloom, and S. J. Davis. 2013. Measuring economic policy uncertainty. *Working Paper, Stanford University.*

Bakshi, G., C. Cao, and Z. Chen. 1997. Empirical performance of alternative option pricing models. *Journal of Finance* 52, 2003–2049.

Bakshi, G., C. Cao, and Z. Chen. 2000. Do call prices and the underlying stock always move in the same direction? *Review of Financial Studies* 13, 549–584.

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

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

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

Bakshi, G., G. Panayotov, and G. Skoulakis. 2011. Improving the predictability of real economic activity and asset returns with forward variances inferred from option portfolios. *Journal of Financial Economics* 100, 475–495.

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

Bansal, R., D. Kiku, and A. Yaron. 2012. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review* 1, 183–221.

Basak, S. 2000. A model of dynamic equilibrium asset pricing with extraneous risk. *Journal of Economic Dynamics and Control* 24, 63–95.

Basak, S. 2005. Asset pricing with heterogeneous beliefs. *Journal of Banking and Finance* 29, 2849–2881.

Bates, D. S. 1991. The crash of '87: was it expected? The evidence from options markets. *Journal of Finance* 46, 1009–1044.

Bates, D. S. 2000. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics* 94, 181–238.

Beber, A. and M. W. Brandt. 2006. The effect of macroeconomic news on beliefs and preferences: evidence from the options market. *Journal of Monetary Economics* 53, 1997–2039.

Beber, A. and M. W. Brandt. 2009. Resolving macroeconomic uncertainty in stock and bond markets. *Review of Finance* 13, 1–45.

Beeler, J. and J. Y. Campbell. 2012. The long-run risks model and aggregate asset prices: an empirical assessment. *Critical Finance Review* 1, 141–182.

Bekaert, G. and M. Hoerova. 2013. The VIX, the variance premium and stock market volatility. *Working Paper, Columbia University.*

Bekaert, G., M. Hoerova, and M. L. Duca. 2013. Risk, uncertainty and monetary policy. *Journal of Monetary Economics* 60, 771–788.

Ben-Rephael, A., S. Kandel, and A. Wohl. 2012. Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics* 104, 363–382.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein. 2011. Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics* 101, 552–573.

Birru, J. and S. Figlewski. 2012. Anatomy of a meltdown: the risk neutral density for the S&P 500 in the fall of 2008. *Journal of Financial Markets* 15, 151–180.

Black, F. 1976. The pricing of commodity contracts. *Journal of Financial Economics* 3, 167–179.

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

Bliss, R. R. and N. Panigirtzoglou. 2002. Testing the stability of implied probability density functions. *Journal of Banking and Finance* 26, 381–422.

Bollen, N. P. B. and R. E. Whaley. 2004. Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance* 59, 711–753.

Bollerslev, T., J. Marrone, L. Xu, and H. Zhou. 2012. Stock return predictability and variance risk premia: statistical inference and international evidence. *Journal of Financial and Quantitative Analysis*, forthcoming.

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

Bondarenko, O. 2003. Estimation of risk-neutral densities using positive convolution approximation. *Journal of Econometrics* 116, 85–112.

Bondarenko, O. 2014. Variance trading and market price of variance risk. *Journal of Econometrics*, forthcoming.

Bookstaber, R. M. and J. B. McDonald. 1987. A general distribution for describing security price returns. *Journal of Business* 60, 401–424.

Boudoukh, J., M. Richardson, and R. F. Whitelaw. 2008. The myth of long-horizon predictability. *Review of Financial Studies* 21, 1577–1605.

Bradshaw, M. T., A. P. Hutton, A. Marcus, and H. Tehranian. 2010. Opacity, crash risk, and the option smirk curve. *Working Paper, Boston College.*

Brandt, M. W. and Q. Kang. 2004. On the relationship between the conditional mean and volatility of stock returns: a latent VAR approach. *Journal of Financial Economics* 72, 217–257.

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

Britten-Jones, M. and A. Neuberger. 2000. Option prices, implied price processes, and stochastic volatility. *Journal of Finance* 55, 839–866.

Broadie, M. and A. Jain. 2008. The effect of jumps and discrete sampling on volatility and variance swaps. *International Journal of Theoretical and Applied Finance* 11, 761–797.

Brown, C. A. and D. M. Robinson. 2002. Skewness and kurtosis implied by option prices: a correction. *Journal of Financial Research* 25, 279–282.

Brown, G. W. and M. T. Cliff. 2004. Investor sentiment and the near-term stock market. *Journal of Empirical Finance* 11, 1–27.

Brown, G. W. and M. T. Cliff. 2005. Investor sentiment and asset valuation. *Journal of Business* 78, 405–440.

Buchen, P. W. and M. Kelly. 1996. The maximum entropy distribution of an asset inferred from option prices. *Journal of Finance and Quantitative Analysis* 31, 143–159.

Buraschi, A. and J. Jackwerth. 2001. The price of a smile: hedging and spanning in option markets. *Review of Financial Studies* 14, 495–527.

Buraschi, A. and A. Jiltsov. 2006. Model uncertainty and option markets with heterogeneous agents. *Journal of Finance* 61, 2841–2897.

Buraschi, A., F. Trojani, and A. Vedolin. 2014. Economic uncertainty, disagreement, and credit markets. *Management Science* 60, 1281–1296.

Byun, S. J. and J. S. Kim. 2013. Improving the predictability of stock market returns with the growth of options open interest. *Working Paper, Korea Advanced Institute of Science and Technology.*

Campa, J. M., P. H. K. Chang, and R. L. Reider. 1998. Implied exchange rate distributions: evidence from OTC options markets. *Journal of International Money and Finance* 17, 117–160.

Campbell, J. Y. 1987. Stock returns and the term structure. *Journal of Financial Economics* 18, 373–399.

Campbell, J. Y. 1991. A variance decomposition for stock returns. *Economic Journal* 101, 157–179.

Campbell, J. Y. and R. J. Shiller. 1988a. Stock prices, earnings and expected dividends. *Journal of Finance* 43, 661–676.

Campbell, J. Y. and R. J. Shiller. 1988b. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1, 195–228.

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

Carr, P. and R. Lee. 2009. Robust replication of volatility derivatives. *Working Paper, New York University.*

Carr, P. and D. Madan. 1998. Towards a theory of volatility trading. In: *Volatility.* Ed. by R. Jarrow. Risk Publications, 417–427.

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

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

Cen, L., G. Hilary, and K. C. J. Wei. 2013. The role of anchoring bias in the equity market: evidence from analysts' earnings forecasts and stock returns. *Journal of Financial and Quantitative Analysis* 48, 47–76.

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

Chang, E. C., J. Ren, and Q. Shi. 2009. Effects of the volatility smile on exchange settlement practices: the Hong Kong case. *Journal of Banking and Finance* 33, 98–112.

Chen, H., S. Joslin, and S. Ni. 2013. Demand for crash insurance, intermediary constraints, and stock return predictability. *Working Paper, MIT.*

Chen, J., H. Hong, and J. C. Stein. 2002. Breadth of ownership and stock returns. *Journal of Financial Economics* 66, 171–205.

Chen, N.-F. 1991. Financial investment opportunities and the macroeconomy. *Journal of Finance* 46, 529–554.

Chen, Z. and L. G. Epstein. 2002. Ambiguity, risk and asset returns in continuous time. *Econometrica* 70, 1403–1445.

Chernov, M. and E. Ghysels. 2000. A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics* 56, 407–458.

Christoffersen, P., S. Heston, and K. Jacobs. 2013. Capturing option anomalies with a variance-dependent pricing kernel. *Review of Financial Studies* 26, 1963–2006.

Clark, T. E. and M. W. McCracken. 2001. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics* 105, 85–110.

Coakley, J., G. Dotsis, X. Liu, and J. Zhai. 2013. Investor sentiment and value and growth stock index options. *European Journal of Finance*, forthcoming.

Cochrane, J. H. 2005. *Asset pricing, revised edition.* Princeton, NJ: Princeton University Press.

Cochrane, J. H. and M. Piazzesi. 2005. Bond risk premia. *American Economic Review* 95, 138–160.

Connor, G. and R. A. Korajczyk. 1986. Performance measurement with the arbitrage pricing theory: a new framework for analysis. *Journal of Financial Economics* 15, 373–394.

Constantinides, G. M., M. Czerwonko, J. C. Jackwerth, and S. Perrakis. 2011. Are options on index futures profitable for risk-averse investors? empirical evidence. *Journal of Finance* 66, 1407–1437.

Constantinides, G. M., J. C. Jackwerth, and S. Perrakis. 2009. Mispricing of S&P 500 index options. *Review of Financial Studies* 22, 1247–1277.

Corrado, C. J. 2001. Option pricing based on the generalized lambda distribution. *Journal of Futures Markets* 21, 213–236.

Corrado, C. J. and T. Su. 1996. Skewness and kurtosis in S&P 500 index returns implied by option prices. *Journal of Financial Research* 19, 175–192.

Corrado, C. J. and T. Su. 1997. Implied volatility skews and stock index skewness and kurtosis implied by S&P 500 index option prices. *Journal of Derivatives* 4, 8–19.

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

Cox, J. C. and S. A. Ross. 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3, 145–166.

David, A. and H. P. Veronesi. 2014. Investors' and central bank's uncertainty embedded in index options. *Review of Financial Studies* 27, 1661–1716.

D'Avolio, G. 2002. The market for borrowing stock. *Journal of Financial Economics* 66, 171–205.

Dechow, P. M., A. Hutton, and R. G. Sloan. 2000. The relation between analysts' long-term earnings growth forecasts and stock price performance following equity offerings. *Contemporary Accounting Research* 17, 1–32.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou. 1999. More than you ever wanted to know about volatility swaps. *Quantitative Strategies Research Notes, Goldman Sachs.*

Dennis, P. and S. Mayhew. 2002. Risk-neutral skewness: evidence from stock options. *Journal of Financial and Quantitative Analysis* 37, 471–493.

Deuskar, P., A. Gupta, and M. G. Subrahmanyam. 2008. The economic determinants of interest rate option smiles. *Journal of Banking and Finance* 32, 714–728.

Diether, K. B., C. J. Malloy, and A. Scherbina. 2002. Differences of opinion and the cross section of stock returns. *Journal of Finance* 57, 2113–2141.

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

Drechsler, I. 2013. Uncertainty, time-varying fear, and asset prices. *Journal of Finance* 68, 1843–1889.

Drechsler, I. and A. Yaron. 2011. What's vol got to do with it. *Journal of Financial and Quantitative Analysis* 24, 1–45.

Driessen, J., P. Maenhout, and G. Vilkov. 2013. Option-implied correlations and the price of correlation risk. *Working Paper, Tilburg University.*

Du, D. 2011. General equilibrium pricing of options with habit formation and event risks. *Journal of Financial Economics* 99, 400–426.

Du, J. and N. Kapadia. 2012. Tail and volatility indices from option prices. *Working Paper, University of Massachusetts, Amherst.*

Duan, J.-C. and J. Wei. 2009. Systematic risk and the price structure of individual equity options. *Review of Financial Studies* 22, 1981–2006.

Dutta, K. K. and D. F. Babbel. 2005. Extracting probabilistic information from the prices of interest rate options: tests of distributional assumptions. *Journal of Business* 78, 841–870.

Easley, D., M. O'Hara, and P. S. Srinivas. 1998. Option volume and stock prices: evidence on where informed traders trade. *Journal of Finance* 53, 431–465.

Ederington, L. H. and J. H. Lee. 1996. The creation and resolution of market uncertainty: the impact of information releases on implied volatility. *Journal of Financial and Quantitative Analysis* 31, 513–539.

Ellsberg, D. 1961. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics* 75, 643–669.

Epstein, L. G. and T. Wang. 1994. Intertemporal asset pricing under Knightian uncertainty. *Econometrica* 62, 283–322.

Estrella, A. and G. A. Hardouvelis. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46, 555–576.

Fabozzi, F. J., R. Tunaru, and G. Albota. 2009. Estimating risk-neutral density with parametric models in interest rate markets. *Quantitative Finance* 9, 55–70.

Fama, E. F. 1990. Term-structure forecasts of interest rates, inflation and real returns. *Journal of Monetary Economics* 25, 59–76.

Fama, E. F. and K. R. French. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22, 3–25.

Fama, E. F. and K. R. French. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25, 23–49.

Ferreira, M. A. and P. Santa-Clara. 2011. Forecasting stock market returns: the sum of the parts is more than the whole. *Journal of Financial Economics* 100, 514–537.

Feunou, B., J.-S. Fontaine, A. Taamouti, and R. Tédongap. 2014. Risk premium, variance premium, and the maturity structure of uncertainty. *Review of Finance* 18, 219–269.

Figlewski, S. 2010. Estimating the implied risk neutral density for the U.S.market portfolio. In: *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*. Ed. by T. Bollerslev, J. R. Russell, and M. Watson. Oxford, UK: Oxford University Press.

Fodor, A., K. Krieger, and J. S. Doran. 2011. Do option open-interest changes foreshadow future equity returns? *Financial Markets and Portfolio Management* 25, 265–280.

Friesen, G. C., Y. Zhang, and T. S. Zorn. 2012. Heterogeneous beliefs and risk neutral skewness. *Journal of Financial and Quantitative Analysis* 47, 851–872.

Gârleanu, N., L. H. Pedersen, and A. M. Poteshman. 2009. Demand-based option pricing. *Review of Financial Studies* 22, 4259–4299.

Gemmill, G. and A. Saflekos. 2000. How useful are implied distributions? Evidence from stock index options. *Journal of Derivatives* 7, 83–91.

Glatzer, E. and M. Scheicher. 2005. What moves the tail? The determinants of the option-implied probability density function of the DAX index. *Journal of Futures Markets* 25, 515–536.

Goetzmann, W. N. and M. Massa. 2005. Dispersion of opinion and stock returns. *Journal of Financial Markets* 8, 325–350.

Golez, B. 2014. Expected returns and dividend growth rates implied by derivative markets. *Review of Financial Studies* 27, 790–822.

Goyal, A. and P. Santa-Clara. 2003. Idiosyncratic risk matters! *Journal of Finance* 58, 975–1008.

Goyal, A. and I. Welch. 2003. Predicting the equity premium with dividend ratios. *Management Science* 49, 639–654.

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

Graham, M., J. Nikkinen, and P. Sahlström. 2003. Relative importance of scheduled macroeconomic news for stock market investors. *Journal of Economics and Finance* 27, 153–165.

Guo, H. 2006. On the out-of-sample predictability of stock returns. *Journal of Business* 79, 645–670.

Han, B. 2008. Investor sentiment and option prices. *Review of Financial Studies* 21, 387–414.

Han, B. and Y. Zhou. 2011. Variance risk premium and cross-section of stock returns. *Working Paper, University of Texas at Austin.*

Hayashi, T. and J. Miao. 2011. Intertemporal substitution and recursive smooth ambiguity preferences. *Theoretical Economics* 6, 423–475.

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

Hill, B. M. 1975. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* 3, 1163–1174.

Hodrick, R. J. 1992. Dividend yields and expected stock returns: alternative procedures for inference and measurement. *Review of Financial Studies* 2, 357–386.

Inoue, A. and L. Kilian. 2004. In-sample or out-of-sample tests of predictability: which one should we use? *Econometric Reviews* 23, 371–402.

Jackwerth, J. C. 1999. Option implied risk-neutral distributions and implied binomial trees: a literature review. *Journal of Derivatives* 7, 66–82.

Jackwerth, J. C. 2000. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* 13, 433–451.

Jackwerth, J. C. 2004. *Option-implied risk-neutral distributions and risk aversion.* Charlottesville, VA: Research Foundation of AIMR. ISBN: 0-943205-66-2.

Jackwerth, J. C. and M. Rubinstein. 1996. Recovering probability distributions from options prices. *Journal of Finance* 51, 1611–1631.

Jarrow, R. and A. Rudd. 1982. Approximate valuation for arbitrary stochastic processes. *Journal of Financial Economics* 10, 347–369.

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

Jiang, H. and Z. Sun. 2014. Dispersion in beliefs among active mutual funds and the cross-section of stock returns. *Journal of Financial Economics* 114, 341–365.

Jondeau, E. and M. Rockinger. 2001. Gram-Charlier densities. *Journal of Economic Dynamics and Control* 25, 1457–1483.

Jones, C. S. 2003. The dynamics of stochastic volatility: evidence from underlying and options markets. *Journal of Econometrics* 116, 181–224.

Jones, C. S. 2006. A nonlinear factor analysis of S&P 500 index option returns. *Journal of Finance* 61, 2325–2363.

Ju, N. and J. Miao. 2012. Ambiguity, learning, and asset returns. *Econometrica* 80, 559–591.

Karoui, M. 2012. Option-implied equity premia and the predictability of stock market returns. *Working Paper, McGill University.*

Keim, D. and R. F. Stambaugh. 1986. Predicting returns in the stock and bond markets. *Journal of financial Economics* 17, 357–390.

Kelly, B. T. and H. Jiang. 2014. Tail risk and asset prices. *Review of Financial Studies* 27, 2841–2871.

Kim, M. and M. Kim. 2003. Implied volatility dynamics in the foreign exchange markets. *Journal of International Money and Finance* 22, 511–528.

Kim, S. and G. Lee. 2011. Effects of macroeconomic news announcements on risk-neutral distribution: evidence from KOSPI200 intraday options data. *Asia-Pacific Journal of Financial Studies* 40, 403–432.

Klibanoff, P., M. Marinacci, and S. Mukerji. 2005. A smooth model of decision making under ambiguity. *Econometrica* 73, 1849–1892.

Klibanoff, P., M. Marinacci, and S. Mukerji. 2009. Recursive smooth ambiguity preferences. *Journal of Economic Theory* 144, 930–976.

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

Lamont, O. 1998. Earnings and expected returns. *Journal of Finance* 53, 1563–1587.

Lamont, O. and J. C. Stein. 2004. Aggregate short interest and market valuations. *American Economic Review* 94, 29–32.

Lemmon, M. and S. X. Ni. 2011. The effects of investor sentiment on speculative trading and prices of stock and index options. *Working Paper, University of Utah.*

Lemmon, M. and E. Portniaguina. 2006. Consumer confidence and asset prices: Some empirical evidence. *Review of Financial Studies* 19, 1499–1529.

Lettau, M. and S. Ludvigson. 2001. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56, 815–849.

Liu, J., J. Pan, and T. Wang. 2005. An equilibrium model of rare-event premia and its implication for option smirks. *Review of Financial Studies* 18, 131–164.

Longstaff, F. A. 1995. Option pricing and the martingale restriction. *Review of Financial Studies* 8, 1091–1124.

Ludvigson, S. C. and S. Ng. 2007. The empirical risk-return relation: a factor analysis approach. *Journal of Financial Economics* 83, 171–222.

Ludvigson, S. C. and S. Ng. 2009. Macro factors in bond risk premia. *Review of Financial Studies* 22, 5027–5067.

Ludvigson, S. C. and S. Ng. 2011. A factor analysis of bond risk premia. In: *Handbook of Empirical Economics and Finance*. Ed. by A. Ullah and D. E. A. Giles. Chapman and Hall, 313–372.

Luo, X. and J. E. Zhang. 2012. Expected stock returns and forward variance. *Working Paper, Zhejiang University.*

Lynch, D. and N. Panigirtzoglou. 2008. Summary statistics of option-implied probability density functions and their properties. *Working Paper, Bank of England.*

Madan, D. and F. Milne. 1994. Contingent claims valued and hedged by pricing and investing in a basis. *Mathematical Finance* 4, 223–245.

Maio, P. 2013. The "FED model"and the predictability of stock returns. *Review of Finance* 17, 1489–1533.

Maio, P. 2014. Don't fight the Fed! *Review of Finance* 18, 623–679.

Maio, P. and D. Philip. 2014a. Macro factors and the cross-section of stock returns. *Working Paper, Hanken School of Economics.*

Maio, P. and D. Philip. 2014b. Macro variables and the components of stock returns. *Working Paper, Hanken School of Economics.*

Malz, A. M. 1996. Using option prices to estimate realignment probabilities in the European monetary system: the case of sterling- mark. *Journal of International Money and Finance* 15, 717–748.

Malz, A. M. 1997. Estimating the probability distribution of the future exchange rate from option prices. *Journal of Derivatives* 5, 18–36.

Markose, S. and A. Alentorn. 2011. The generalized extreme value (GEV) distribution, implied tail index and option pricing. *Journal of Derivatives* 18, 35–60.

McCracken, M. W. 2007. Asymptotics for out of sample tests of Granger causality. *Journal of Econometrics* 140, 719–752.

Melick, W. R. and C. P. Thomas. 1997. Recovering an asset's implied PDF from option prices: an application to crude oil during the Gulf crisis. *Journal of Financial and Quantitative Analysis* 32, 91–115.

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

Miller, E. M. 1977. Risk, uncertainty, and divergence of opinion. *Journal of Finance* 32, 1151–1168.

Mixon, S. 2002. Factors explaining movements in the implied volatility surface. *Journal of Futures Markets* 22, 915–937.

Mixon, S. 2011. What does implied volatility skew measure? *Journal of Derivatives* 18, 9–25.

Mueller, P., A. Vedolin, and Y.-M. Yen. 2013. Bond variance risk premiums. *Working Paper, London School of Economics.*

Mueller, P., A. Vedolin, and H. Zhou. 2011. Short-run bond risk premia. *Working Paper, London School of Economics.*

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

Neumann, M. and G. Skiadopoulos. 2013. Predictable dynamics in higher order risk-neutral moments: evidence from the S&P 500 options. *Journal of Financial and Quantitative Analysis* 48, 947–977.

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

Ni, S. X., J. Pan, and A. M. Poteshman. 2008. Volatility information trading in the option market. *Journal of Finance* 63, 1059–1091.

Nikkinen, J. and P. Sahlström. 2004. Impact of the federal open market committee's meetings and scheduled macroeconomic news on stock market uncertainty. *International Review of Financial Analysis* 13, 1–12.

Nofsinger, J. R. and B. Prucyk. 2003. Option volume and volatility response to scheduled economic news releases. *Journal of Futures Markets* 23, 315–345.

Nordén, L. and C. Xu. 2012. Option happiness and liquidity: is the dynamics of the volatility smirk affected by relative option liquidity? *Journal of Futures Markets* 32, 47–74.

Pan, J. 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50.

Pan, J. and A. M. Poteshman. 2006. The information in option volume for future stock prices. *Review of Financial Studies* 19, 871–908.

Pan, X. 2012. Investor beliefs and state price densities in the crude oil market. *Working Paper, Tulane University.*

Park, C. 2005. Stock return predictability and the dispersion in earnings forecasts. *Journal of Business* 78, 2351–2376.

Peña, I., G. Rubio, and G. Serna. 1999. Why do we smile? On the determinants of the implied volatility function. *Journal of Banking and Finance* 23, 1151–1179.

Qiu, L. and I. Welch. 2004. Investor sentiment measures. *Working Paper, NBER.*

Rapach, D. E., J. K. Strauss, J. Tu, and G. Zhou. 2011. Out-of-sample industry return predictability: evidence from a large number of predictors. *Working Paper, Washington University in St. Louis.*

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

Rebonato, R. 2004. *Volatility and correlation, 2nd edition: the perfect hedger and the fox.* Wiley.

Ritchey, R. J. 1990. Call option valuation for discrete normal mixtures. *Journal of Financial Research* 13, 285–296.

Rompolis, L. S. and E. Tzavalis. 2008. Recovering risk neutral densities from option prices: a new approach. *Journal of Financial and Quantitative Analysis* 43, 1037–1054.

Rompolis, L. S. and E. Tzavalis. 2013. Retrieving risk neutral moments and expected quadratic variation from option prices. *Working Paper, Athens University of Economics and Business.*

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

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

Rubinstein, M. 1998. Edgeworth binomial trees. *Journal of Derivatives* 5, 20–27.

Ruf, T. 2012. Limits to market making and the skewness risk premium in options markets. *Working Paper, University of New South Wales.*

Sarantopoulou-Chiourea, S. and G. Skiadopoulos. 2014. Index option prices as a predictor of real economic activity. *Working Paper, University of Piraeus.*

Savickas, R. 2002. A simple option-pricing formula. *The Financial. Review* 37, 207–226.

Scheinkman, J. A. and W. Xiong. 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* 111, 1183–1220.

Shefrin, H. 2008. *A behavioral approach to asset pricing, second edition.* Burlington, MA: Elsevier Academic Press.

Sherrick, B. J., P. Garcia, and V. Tirupattur. 1996. Recovering probabilistic information from option markets: tests of distributional assumptions. *Journal of Futures Markets* 16, 545–560.

Sherrick, B. J., S. H. Irwin, and D. L. Forster. 1992. Option-based evidence of the nonstationarity of expected S&P 500 futures price distributions. *Journal of Futures Markets* 12, 275–290.

Sherrick, B. J., S. H. Irwin, and D. L. Forster. 1996. An examination of option-implied S&P 500 futures price distributions. *Financial Review* 31, 667–694.

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

Sibley, S. E., Y. Xing, and X. Zhang. 2013. Is sentiment sentimental? *Working Paper, Purdue University.*

Söderlind, P. 2000. Market Expectations in the UK before and after the ERM Crisis. *Economica* 67, 1–18.

Söderlind, P. and L. Svensson. 1997. New techniques to extract market expectations from financial instruments. *Journal of Monetary Economics* 40, 383–429.

Steeley, J. M. 2004. Stock price distributions and news: evidence from index options. *Review of Quantitative Finance and Accounting* 23, 229–250.

Stock, J. H. and M. W. Watson. 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20, 147–162.

Stock, J. H. and M. W. Watson. 2005. Implications of dynamic factor models for VAR analysis. *Working Paper, NBER.*

Stutzer, M. 1996. A simple nonparametric approach to derivative security valuation. *Journal of Finance* 51, 1633–1652.

Sun, P. and C. Sutcliffe. 2003. Scheduled announcements and volatility patterns: the effects of monetary policy committee announcements on LIBOR and short sterling futures and options. *Journal of Futures Markets* 23, 773–797.

Taylor, S. J., P. K. Yadav, and Y. Zhang. 2009. Cross-sectional analysis of risk-neutral skewness. *Journal of Derivatives* 16, 38–52.

Torous, W., R. Valkanov, and S. Yan. 2004. On predicting stock returns with nearly integrated explanatory variables. *Journal of Business* 77, 937–966.

Vähämaa, S. and J. Äijö. 2011. The fed's policy decisions and implied volatility. *Journal of Futures Markets* 31, 995–1009.

Vähämaa, S., S. Watzka, and J. Äijö. 2005. What moves option-implied bond market expectations? *Journal of Futures Markets* 25, 817–843.

Valkanov, R. 2003. Long-horizon regressions: theoretical results and applications. *Journal of Financial Economics* 68, 201–232.

Vilkov, G. and Y. Xiao. 2013. Option-implied information and predictability of extreme returns. *Working Paper, Goethe University.*

Xing, Y., X. Zhang, and R. Zhao. 2010. What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis* 45, 641–662.

Yu, J. 2011. Disagreement and return predictability of stock portfolios. *Journal of Financial Economics* 99, 162–183.

Zhou, H. 2012. Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty. *Working Paper, Tsinghua University*.

Appendix A

Appendix to Chapter 2

A.1 Parametric methods for Extracting Risk-Neutral Densities

The risk-neutral density (RND) can be extracted by integrating the stochastic process governing the underlying asset price dynamics, assuming that this is known. Making use of this argument, Bates (1991) and Malz (1996) fit different stochastic processes in the observed option prices, estimate the necessary parameters and finally obtain the respective RNDs. In contrast to this approach, several alternative techniques extract the RND directly from the observed option prices, remaining silent about the stochastic process followed by the underlying asset price. As pointed out by Melick and Thomas (1997), such techniques cannot provide any information about the evolution of the asset price throughout the life of the option, but have the advantage that they are more flexible in capturing the shape of the implied distribution. This is important since a given RND can be consistent with several stochastic processes, while a given stochastic process can only be associated with one RND.

Therefore, the aim of this appendix is to provide an overview of the parametric methods that can be used in order to extract risk-neutral densities from option prices. The parametric techniques can be further divided into three subgroups: the *expansion* methods, the *generalized distribution* methods and the *mixture* methods.

A.1.1 Expansion methods

Expansion techniques add correction terms to a reference probability distribution which is usually either the normal or the lognormal one. As pointed out by Jackwerth (1999) the idea is similar to that of the Taylor series expansion for the approximation of an analytic function.

Many researchers make use of the Gram-Charlier series - sometimes they refer to it as Edgeworth series - in order to approximate the risk-neutral density of the underlying asset. Jarrow and Rudd (1982) state that the RND function $g(S_t)$ of the asset price can be approximated by a lognormal distribution $f(S_t)$ as follows:

$$\begin{aligned} g(S_t) &= f(S_t) + \frac{\kappa_2(g) - \kappa_2(f)}{2!}\frac{d^2 f(S_t)}{dS_t^2} - \frac{\kappa_3(g) - \kappa_3(f)}{3!}\frac{d^3 f(S_t)}{dS_t^3} + \\ &+ \frac{(\kappa_4(g) - \kappa_4(f)) + 3\,(\kappa_2(g) - \kappa_2(f))^2}{4!}\frac{d^4 f(S_t)}{dS_t^4}, \end{aligned} \tag{A.1}$$

assuming that $\kappa_1(g) = \kappa_1(f) = S_0 e^{rt}$, where $\kappa_i(.)$ is the i^{th} cumulant of the respective probability distribution. Jarrow and Rudd (1982) show that the pricing formula for a call option with strike price X becomes:

$$\begin{aligned} C_g(X) &= C_f(X) + e^{-rt}\frac{\kappa_2(g)-\kappa_2(f)}{2!}f(X) - e^{-rt}\frac{\kappa_3(g)-\kappa_3(f)}{3!}\frac{df(X)}{dS_t} + \\ &+ e^{-rt}\frac{(\kappa_4(g)-\kappa_4(f)) + 3(\kappa_2(g)-\kappa_2(f))^2}{4!}\frac{d^2f(X)}{dS_t^2}, \end{aligned} \tag{A.2}$$

where $C_f(X)$ corresponds to the call price given by the Black-Scholes formula.

Corrado and Su (1997) simplify the formula by assuming that $\kappa_2(g) = \kappa_2(f)$. In this case, for a lognormal distribution $f(S_t)$ it is true that $\kappa_2(g) = \kappa_1^2(f)(e^{\sigma^2 t} - 1)$, with σ being the volatility parameter. Therefore, the call option pricing formula becomes:

$$C_g(X) = C_f(X) + \lambda_1 Q_3 + \lambda_2 Q4, \tag{A.3}$$

where:

$$\begin{aligned} \lambda_1 &= \gamma_1(g) - \gamma_1(f) \\ \lambda_2 &= \gamma_2(g) - \gamma_2(f) \\ Q_3 &= -(S_0 e^{rt})^3 \left(e^{\sigma^2 t} - 1\right)^{3/2} \frac{e^{-rt}}{3!}\frac{df(X)}{dS_t} \\ Q_4 &= (S_0 e^{rt})^4 \left(e^{\sigma^2 t} - 1\right)^{2} \frac{e^{-rt}}{4!}\frac{df^2(X)}{dS_t^2} \\ \gamma_1(g) &= \frac{\kappa_3(g)}{\kappa_2^{3/2}(g)} \\ \gamma_2(g) &= \frac{\kappa_4(g)}{\kappa_2^2(g)} \\ \gamma_1(f) &= 3q + q^3 \\ \gamma_2(f) &= 16q^2 + 15q^4 + 6q^6 + q^8, \end{aligned} \tag{A.4}$$

and $\left(e^{\sigma^2 t} - 1\right)$ is defined as q^2.

Corrado and Su (1996) use the normal distribution as the reference probability distribution and therefore apply an A-Type Gram-Charlier series to approximate the risk-neutral distribution of the log-price of the undering asset. More specifically, they show that after standardizing for a zero mean and unit variance, the RND can be expressed as the sum of the normal density plus correction terms adjusting for skewness and kurtosis:

$$\begin{aligned} g(z) &= n(z)\left[1 - \frac{\mu_3}{3!}H_3(z) + \frac{\mu_4 - 3}{4!}H_4(z)\right] = \\ &= n(z)\left[1 - \frac{\mu_3}{3!}\left(z^3 - 3z\right) + \frac{\mu_4 - 3}{4!}\left(z^4 - 6z^2 + 3\right)\right], \end{aligned} \tag{A.5}$$

where $n(z)$ is the standard normal density function, μ_i is the i^{th} central moment of $g(z)$ and $H_i(z)$ is the i^{th} Hermite polynomial of the standardized value of $\ln(S_t)$.[1]

[1] The Hermite polynomial can be defined by the relation: $H(z)\,n(z) = (-1)^n \frac{d^n n(z)}{dz^n}$. However,

Now the call option pricing formula becomes:

$$C_g(X) = C_f + \mu_3 Q_3 + (\mu_4 - 3) Q_4, \tag{A.6}$$

where:

$$Q_3 = \frac{1}{3!} S_0 \sigma \sqrt{t} \left[\left(2\sigma\sqrt{t} - d \right) n(d) + \sigma^2 t N(d) \right] \tag{A.7}$$

$$Q_4 = \frac{1}{4!} S_0 \sigma \sqrt{t} \left[\left(d^2 - 1 - 3\sigma\sqrt{t} \left(d - \sigma\sqrt{t} \right) \right) n(d) + \sigma^3 t^{3/2}(d) \right], \tag{A.8}$$

d is defined as in the Black-Scholes formula and $n(.)$, $N(.)$ are the probability density function (PDF) and the cumulative distribution function (CDF) of the normal distribution accordingly.[2] The A-Type Gram-Charlier expansion for approximating the RND of the underlying asset is also used by Longstaff (1995). Jondeau and Rockinger (2001) improve the method by creating an algorithm that guarantees positive probabilities over the whole distribution.

Abken et al. (1996) provide a similar expansion technique based on Madan and Milne's (1994) suggestion that any contigent claim can be seen as an element of a seperable Hilbert space. The Hilbert space is assumed to be a one-dimensional Gaussian reference space and its basis can be formed with Hermite polynomials. Therefore, each polynomial $H_i(z)$ can be seen as a risk coefficient and the risk-neutral density can be estimated as a linear combination of those risk elements. The authors define $\pi_i = \beta_i e^{-rt}$ as the implicit price of polynomial risk $H_i(z)$. The RND function takes the form:

$$g(z) = n(z) \left[1 + \frac{\beta_3}{\sqrt{3!}} H_3(z) + \frac{\beta_4}{\sqrt{4!}} H_4(z) \right], \tag{A.9}$$

if we assume that $\pi_0 = e^{-rt}$, $\pi_1 = \pi_2 = 0$. $H_i(z)$ denotes again the Hermite polynomial and the term $\sqrt{i!}$ normalizes each polynomial to unit variance.

Rubinstein (1998) assumes that the log-returns of the asset follow a binomial distribution $b(x)$ and applies an Edgeworth expansion to approximate the risk-neutral density function. Therefore, if it is assumed that the distribution is standardized to zero mean and unit variance, the RND function becomes:

$$g(z) = b(z) \left[\begin{array}{c} 1 + \frac{\mu_3}{3!}(z^3 - 3z) + \frac{\mu_4 - 3}{4!}(z^4 - 6z^2 + 3) + \\ + \frac{\mu_3^2}{6!}(z^6 - 15z^4 + 45z^2 - 15) \end{array} \right], \tag{A.10}$$

where z is the standardized log-return of the underlying asset and μ_i is the i^{th} central moment of $g(z)$.

Finally, Rompolis and Tzavalis (2008) suggest a methodology which belongs to the class of Gram-Charlier series but does not rely on any reference distribution. As a result, it can be considered nonparametric and it is much more flexible than the simple A-Type Gram-Charlier expansion. The authors apply a C-Type Gram-Charlier expansion in order to approximate the RND function of the log-returns (x)

Corrado and Su (1996) ignore the term $(-1)^n$ in their definition.

[2] The original equations presented in Corrado and Su (1996) contain two typos. The equations presented here are corrected as suggested by Brown and Robinson (2002).

of the underlying asset:

$$g(x) = Q \exp\left[\sum_{i=1}^{\infty} \frac{1}{i} \delta_i H_i(z)\right], \tag{A.11}$$

where $Q = \left\{\int \exp\left[\sum_{i=1}^{\infty} \frac{1}{i}\delta_i H_i(z)\right] dx\right\}^{-1}$, z denotes the standardized value of x and δ_i is the i^{th} order series coefficient of the expansion. The expansion is truncated up to an optimal order (m) which can be found with an information criterion like Akaike's (AIC) or Schwarz (SC). The estimation of the coefficients δ_i for $i = 1, 2, ..., m$ needs estimates of the noncentral moments of x which can be obtained following the methodologies suggested by Bakshi, Kapadia and Madan (2003) for $i = 1, 2, 3, 4$ and Rompolis and Tzavalis (2013) for $i > 4$. A great advantage of the methodology proposed by Rompolis and Tzavalis (2008) is that the exponential form of the RND function guarantees that there will not exist any negative probabilities.

A.1.2 Generalized distribution methods

Generalized distributions include more parameters than the typical two representing the mean and the variance. As a result, they can offer more flexibility and can be useful for extracting risk-neutral distributions from option prices. Moreover, common distributions like the lognormal can be seen as special cases of a generalized one.

The most widely used generalized distribution is the Generalized Beta distribution of the second kind (GB2). The GB2 distribution, introduced by Bookstaber and McDonald (1987), is highly flexible since it has four parameters and encompasses many other distributions used in the literature as special or limiting cases (e.g. Weibull, Burr III and XII, Generalized Gamma). In the context of options markets, it is used by Aparicio and Hodges (1998), Anagnou-Basioudis et al. (2005) and Rebonato (2004) among others. It can be written as:

$$g(S_t) = \frac{|a|\, S_t^{ap-1}}{b^{ap} B(p,q) \left[1 + \left(\frac{S_t}{b}\right)^a\right]^{p+q}}, \tag{A.12}$$

where $B(p,q)$ denotes the beta function defined as follows:

$$B(p,q) = \int_0^1 t^{p-1} (1-t)^{q-1}\, dt, \tag{A.13}$$

and α, b, p and q are the parameters of the distribution. Parameter b is a scale parameter whereas the other parameters determine the shape of the distribution. Specifically, parameter a affects the kurtosis, the interaction of a and q determines the number of existing higher moments and the interaction of parameters p and q drives the skewness. Rebonato (2004) also derives closed form solutions for the option prices. The resulting call option pricing formula is:[3]

$$C_{GB2}(X) = e^{-rt}\left\{\begin{array}{c}\frac{X\left(\frac{b}{X}\right)^{aq}F_1^2\left[q-1/a,p+q;1+q-1/a;-\left(\frac{b}{X}\right)^a\right]}{(q-1/a)B(p,q)} \\ -\frac{X\left(\frac{b}{X}\right)^{aq}F_1^2\left[q,p+q;1+q;-\left(\frac{b}{X}\right)^a\right]}{qB(p,q)}\end{array}\right\}, \tag{A.14}$$

where $F_1^2\left[.,.;.;.\right]$ denotes the hypergeometric function defined as:

$$F_1^2\left[k_1,k_2;l_1;m\right] = \sum_{n=0}^{\infty}\frac{\left(k_1\right)_n\left(k_2\right)_n}{\left(l_1\right)_n}\frac{m^n}{n!}$$

$$\left(k\right)_n = k\left(k+1\right)\left(k+2\right)...\left(k+n-1\right) \text{ and } \left(k\right)_0 = 1, \tag{A.15}$$

with the series terminating when either k_1 or k_2 is a non-positive integer.

Sherrick, Garcia and Tirupattur (1996) use a Burr III distribution to approximate the RND. Burr III is a special case of the GB2 for $q = 1$. The density function in this case can be described by:

$$g\left(S_t\right) = \frac{apS_t^{ap-1}b^a}{\left(b^a + S_t^a\right)^{p+1}}, \tag{A.16}$$

where $a > 0$, $b > 0$ and $p > 0$ are the parameters of the distribution.[4] Following a similar approach, Sherrick, Irwin and Forster (1992, 1996) use the Burr XII distribution to model the RND as:

$$g\left(S_t\right) = \frac{aqS_t^{a-1}b^{aq}}{\left(b^a + S_t^a\right)^{q+1}}, \tag{A.17}$$

where $a > 0$, $b > 0$ and $q > 0$ are again the parameters of the distribution. Burr XII is another special case of the GB2 for $p = 1$. While the aforementioned three studies use American options, none of them explains how the possibility of an early exercise is incorporated into the RND estimations.

Fabozzi et al. (2009) propose the Generalized Gamma distribution (GG) for modelling the implied distribution. GG can be seen as a limiting case of the GB2 when $b = \beta q^{1/a}$ and $q \to \infty$. The respective RND function is:

$$g\left(S_t\right) = \frac{1}{\Gamma\left(p\right)}\left(\frac{\beta}{a}\right)\left(\frac{S_t}{a}\right)^{\beta p-1}\exp\left(-\left(\frac{S_t}{a}\right)^{\beta}\right), \tag{A.18}$$

where $a > 0$, $\beta > 0$ and $p > 0$ are the parameters and $\Gamma\left(p\right)$ is the gamma function defined as:

$$\Gamma\left(p\right) = \int_0^{\infty} t^{p-1}e^{-t}dt. \tag{A.19}$$

[3] The discounting term does not appear in the original formula derived by Rebonato (2004) as he assumes that there are no interest rates in the economy.

[4] Sherrick, Garcia and Tirupattur (1996) as well as other researchers use futures options to extract the risk-neutral density function. However, since this chapter does not investigate the empirical applications of each study but only the theoretical contributions, the notation for the risk-neutral density function remains $g\left(S_t\right)$ even if the underlying asset under investigation in the original study is the futures price.

Using the GG density function, they derive the following option pricing formula for a call:

$$C_{GG}(X) = e^{-rt}\left\{\begin{array}{c} a\frac{\Gamma\left(p+\frac{1}{\beta}\right)}{\Gamma(p)} - X - a\frac{\Gamma\left(p+\frac{1}{\beta}\right)}{\Gamma(p)} I\left[p+\frac{1}{\beta}, \left(\frac{X}{a}\right)^{\beta}\right] + \\ +XI\left[p, \left(\frac{X}{a}\right)^{\beta}\right] \end{array}\right\}, \tag{A.20}$$

where $I(.,.)$ is the incomplete gamma function:

$$I(p,t) = \frac{1}{\Gamma(p)}\int_0^t u^{p-1}e^{-u}du. \tag{A.21}$$

A special case of the GG distribution for $p = 1$ is the Weibull distribution. Savickas (2002) suggests that the RND of the underlying asset can be described by the formula:

$$g(S_t) = k\beta S_t^{\beta-1}e^{-kS_t^{\beta}}, \tag{A.22}$$

where $k > 0$ and $\beta > 0$ are the distribution parameters and $k = a^{-\beta}$ from the GG distribution. Savickas (2002) shows that despite the fact that the Weibull distribution has only two parameters similarly to the lognormal distribution, it strongly outperforms the lognormal in terms of pricing ability. The reason is that, unlike the lognormal distribution, it is possible for the Weibull distribution to exhibit negative skewness. When the risk-neutral distribution is assumed to be of Weibull form, the call option pricing formula becomes (Fabozzi et al., 2009):[5]

$$C_W(X) = e^{-rt}\frac{1}{\beta k^{1/\beta}}\left[\Gamma\left(\frac{1}{\beta}\right) - I\left(\frac{1}{\beta}, kX^{\beta}\right)\right], \tag{A.23}$$

where $\Gamma(.)$ and $I(.)$ denote again the gamma and the incomplete gamma function respectively.

Dutta and Babbel (2005) use the g-and-h distribution to capture the implied distribution of the asset price. In this case, the RND function can be expressed as:

$$g(z) = a + b\left(e^{gz} - 1\right)\frac{\exp\left(hz^2/2\right)}{g}, \tag{A.24}$$

where z is the standard normal variable and a, b, g, h are the parameters which refer to location, scale, skewness and kurtosis respectively. The respective call option pricing formula takes the following form:

$$C_{g-h}(X) = e^{-rt}\left\{\begin{array}{c} (a-X)\left[1-N(X)\right] - \frac{b}{g\left(\sqrt{1-h}\right)} \\ \left[1-N\left(X\sqrt{1-h}\right)\right] + \frac{b}{g\left(\sqrt{1-h}\right)}e^{g^2/2(1-h)} \\ \left[1-N\left(\sqrt{1-h}\right)X - g/\left(\sqrt{1-h}\right)\right] \end{array}\right\}, \tag{A.25}$$

where $N(.)$ is the CDF of the normal distribution.

Corrado (2001) suggests the usage of the Generalized Lamda distribution (GL) for modelling the risk-neutral distribution of an asset. Since the GL distribution is defined by its percentile function, Corrado (2001) first transforms the formula of the

[5]A similar formula is also derived by Dutta and Babbel (2005). However, Fabozzi et al. (2009) claim that there is a mistake in the formula derivation.

call option's expected payoff as follows:

$$\int_{X}^{\infty}(S_t - X)\,dg\,(S_t) = \int_{g(X)}^{1}(S_t\,(g) - X)\,dg. \tag{A.26}$$

By setting the mean equal to S_0e^{rt} and the variance to $(S_0e^{rt})^2\left(e^{\sigma^2 t}-1\right)$, the call price is estimated by:

$$C_{GL}\,(X) = S_0G_1 - e^{-rt}XG_2 \tag{A.27}$$

where:

$$\begin{aligned}
G_1 &= 1-g\,(X)+\frac{\sqrt{e^{\sigma^2 t}-1}}{\lambda_2\,(\lambda_3,\lambda_4)}\left(\begin{array}{c}\frac{g(X)-g(X)^{\lambda_3+1}}{\lambda_3+1}+\\ +\frac{1-g(X)-\left(1-g(X)^{\lambda_4+1}\right)}{\lambda_4+1}\end{array}\right)\\
G_2 &= 1-g\,(X)\\
\lambda_2\,(\lambda_3,\lambda_4) &= sign\,(\lambda_3)\,\sqrt{b-a^2}\\
a &= \frac{1}{\lambda_3+1}-\frac{1}{\lambda_4+1}\\
b &= \frac{1}{2\lambda_3+1}-\frac{1}{2\lambda_4+1}-2B\,(\lambda_3+1,\lambda_4+1)\,.
\end{aligned} \tag{A.28}$$

Therefore, the three parameters to be estimated are σ, λ_3 and λ_4, where λ_3 and λ_4 affect the shape of the distribution and $B\,(.,.)$ stands again for the beta function.

Finally, Markose and Alentorn (2011) propose the Generalized Extreme Value (GEV) distribution for capturing the risk neutral distribution of an asset's returns. The authors model the returns in terms of losses (L_t) and derive the RND function of the asset's price at maturity as:

$$g\,(S_t) = \frac{1}{S_0\sigma}\left[1+\xi\frac{(L_t-\mu)}{\sigma}\right]^{-1(-1/\xi)}\exp\left\{-\left[1+\xi\frac{(L_t-\mu)}{\sigma}\right]^{-1/\xi}\right\}, \tag{A.29}$$

where μ, σ and ξ are the parameters that determine the location, scale and shape of the distribution respectively. The corresponding call option pricing formula is:

$$C_{GEV}\,(X) = e^{-rt}\left\{S_0\left(\begin{array}{c}(1-\mu+\sigma/\xi)\,e^{-H^{-1/\xi}}-\\ \text{-}\frac{\sigma}{\xi}I\left(1-\xi,H^{-1/\xi}\right)\end{array}\right)-Xe^{-H^{-1/\xi}}\right\}, \tag{A.30}$$

where:

$$H = 1+\frac{\xi}{\sigma}\left(1-\frac{X}{S_0}-\mu\right). \tag{A.31}$$

and $I\,(.,.)$ is the incomplete gamma function.

A.1.3 Mixture methods

Mixture methods incorporate the idea of capturing the RND by the weighted average of two or more distributions. Ritchey (1990) first points out that the well documented leptokurtic stock return distributions can be explained if it is assumed that returns follow a nonstationary normal distribution, i.e. that they are normally distributed for short time periods but the distributional parameters change over time. In this case, a mixture of normal distributions can approximate such a non-stationary process. Then, Ritchey (1990) derives an option pricing model which is composed by a weighted sum of Black-Scholes prices.

Melick and Thomas (1997) assume that the RND of the underlying asset at the maturity of the option can be described by a mixture of three lognormal distributions:

$$g(S_t) = \pi_1 f_1(S_t) + \pi_2 f_2(S_t) + \pi_3 f_3(S_t), \tag{A.32}$$

where $f(S_t)$ stands for the lognormal distribution. The reason they mix three distributions is because they examine a period where the market was anticipating three possible outcomes. Since Melick and Thomas (1997) use American style options they cannot derive a closed form solution for the option prices and surpass this problem by using upper and lower boundaries:

$$\begin{aligned} C^u &= E_0\left[\max\left(0, S_t - X\right)\right] \\ C^l &= \max\left\{E_0\left[S_t\right] - X, e^{-rt} E_0\left[\max\left(0, S_t - X\right)\right]\right\}, \end{aligned} \tag{A.33}$$

Therefore, for the price of a call option they derive the following formula:

$$C_M(X) = w_j C^u(X, \theta) + (1 - w_j) C^l(X, \theta), \tag{A.34}$$

where θ denotes the nine main parameters (π_i, μ_i, σ_i for $i = 1, 2, 3$) to be estimated, and w_j - with $j = 1$ when the call option is in-the-money and $j = 2$ when the call option is out-of-the-money - is the weight that expresses the relative position of the option price in respect to the bounds (C^u, C^l).

Bahra (1997) argues that Melick and Thomas' (1997) methodology may be impossible to be implemented in cases where few options are traded in the market due to the large number of parameters that have to be estimated. As a result, he proposes a similar approach which assumes that the RND of the terminal price is a mixture of two lognormal distributions. Given that in his research he examines only European options, the pricing formula of a call option takes the form:

$$C_M(X) = e^{-rt} \int_X^{\infty} \left[\pi f_1(S_t) + (1 - \pi) f_2(S_t)\right] (S_t - X) dS_t. \tag{A.35}$$

Hence, Bahra's method requires the estimation of only five parameters (μ_1, σ_1, μ_2, σ_2, π). The same technique is also used by Gemmill and Saflekos (2000). A variant of the aforementioned methodologies is suggested by Söderlind and Svensson (1997) and Söderlind (2000) where the joint distribution of the logarithm of the asset price and the discount factor corresponding to the lifetime of the option can be modeled as a mixture of bivariate normal densities.

Appendix B

Appendix to Chapter 3

B.1 Macroeconomic Dataset

Table B.1: Macroeconomic variables
This table lists the name of each macroeconomic variable along with its mnemonic label, brief description of the series and the transformation applied to ensure stationarity. In the *transf* column, 1 denotes using levels, 2 denotes taking first-differences, 3 denotes taking second-differences, 4 denotes taking logs, 5 denotes taking log-differences and 6 denotes taking second-log-differences. All the series are from Global Insights Basic Economics database unless specified as TCB (The Conference Board) or AC (Author calculation). The sample period is 1990:01 to 2011:06.

Series no.	Name	Mnemonic	Description	transf
Output and Income				
1	PI	ypr	Personal Income (AR, Bil. Chain 2000 $)	5
2	PI less transfers	a0m051	Personal Income less Transfer Payments (AR, Bil. Chain 2000 $)	5
3	IP: total	ips10	Industrial Production Index - Total Index	5
4	IP:products	ips11	Industrial Production Index - Products, Total	5
5	IP:final prod	ips299	Industrial Production Index - Final Products	5
6	IP:consgds	ips12	Industrial Production Index - Consumer Goods	5
7	IP: cons dble	ips13	Industrial Production Index - Durable Consumer	5
8	IP: cons nondble	ips18	Industrial Production Index - Nondurable Consumer	5
9	IP: bus eqpt	ips25	Industrial Production Index - Business Equipment	5
10	IP: matls	ips32	Industrial Production Index - Materials	5
11	IP: dble matls	ips34	Industrial Production Index - Durable Goods	5
12	IP: nondble matls	ips38	Industrial Production Index - Nondurable Goods	5
13	IP: mfg	ips43	Industrial Production Index - Manufacturing	5
14	IP: res util	ips307	Industrial Production Index - Residential Utilities	5
15	IP: fuels	ips306	Industrial Production Index - Fuels	5
16	NAPM prodn	pmp	Napm Production Index (Percent)	1
17	Cap util	utl11	Capacity Utilization (SIC-Mfg)	2
Employment				
18	Help wanted indx	lhel	Index of Help-Wanted Advertising In Newspapers (1967=100;Sa)	2
19	Help wanted/unemp	lhelx	Help-Wanted Ads:No. Unemployed Clf	2
20	Emp CPS total	lhem	Civilian Labor Force: Employed, Total (Thous.,Sa)	5
21	Emp CPS nonag	lhnag	Civilian Labor Force: Employed, Nonagric.Industries (Thous.,Sa)	5
22	U: all	lhur	Unemployment Rate: All Workers, 16 Years	2
23	U: mean duration	lhu680	Unemploy.By Duration: Average(Mean)Duration In Weeks (Sa)	2
24	U ¡ 5 wks	lhu5	Unemploy.By Duration: Persons Unempl.Less Than 5 Wks (Thous.,Sa)	5
25	U 5-14 wks	lhu14	Unemploy.By Duration: Persons Unempl.5 To 14 Wks (Thous.,Sa)	5
26	U 15+ wks	lhu15	Unemploy.By Duration: Persons Unempl.15 Wks + (Thous.,Sa)	5

Series no.	Name	Mnemonic	Description	transf
27	U 15-26 wks	lhu26	Unemploy.By Duration: Persons Unempl.15 To 26 Wks (Thous.,Sa)	5
28	U 27+ wks	lhu27	Unemploy.By Duration: Persons Unempl.27 Wks + (Thous,Sa)	5
29	UI claims	luinc	Average Weekly Initial Claims, Unemploy. Insurance (Thous. Sa)	5
30	Emp: total	ces002	Employees On Nonfarm Payrolls: Total Private	5
31	Emp: gds prod	ces003	Employees On Nonfarm Payrolls - Goods-Producing	5
32	Emp: mining	ces006	Employees On Nonfarm Payrolls - Mining	5
33	Emp: const	ces011	Employees On Nonfarm Payrolls - Construction	5
34	Emp: mfg	ces015	Employees On Nonfarm Payrolls - Manufacturing	5
35	Emp: dble gds	ces017	Employees On Nonfarm Payrolls - Durable Goods	5
36	Emp: nondbles	ces033	Employees On Nonfarm Payrolls - Nondurable Goods	5
37	Emp: services	ces046	Employees On Nonfarm Payrolls - Service-Providing	5
38	Emp: TTU	ces048	Employees On Nonfarm Payrolls - Trade, Transportation, & Utilities	5
39	Emp: wholesale	ces049	Employees On Nonfarm Payrolls - Wholesale Trade.	5
40	Emp: retail	ces053	Employees On Nonfarm Payrolls - Retail Trade	5
41	Emp: FIRE	ces088	Employees On Nonfarm Payrolls - Financial Activities	5
42	Emp: Govt	ces140	Employees On Nonfarm Payrolls - Government	5
43	Agg wkly hours		Index of Aggregate Weekly Hours	2
44	Avg hrs	ces151	Avg Weekly Hrs of Prod or Nonsup Workers On Private Nonfarm Payrolls - Goods-Producing	2
45	Overtime: mfg	ces155	Avg Weekly Hrs of Prod or Nonsup Workers On Private Nonfarm Payrolls - Mfg Overtime Hours	2
46	Avg hrs: mfg	a0m001	Average Weekly Hours, Mfg. (Hours)	2
47	NAPM empl	pmemp	Napm Employment Index (Percent)	1
48	AHE: goods	ces275	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls - Goods-Producing	6
49	AHE: const	ces277	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls - Construction	6
50	AHE: mfg	ces278	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls - Manufacturing	6
Housing				
51	Starts: nonfarm	hsfr	Housing Starts:Nonfarm(1947-58);Total Farm & Nonfarm(1959-)(Thous.,Saar)	5
52	Starts: NE	hsne	Housing Starts:Northeast (Thous.U.)S.A.	5
53	Starts: MW	hsmw	Housing Starts:Midwest(Thous.U.)S.A.	5
54	Starts: South	hssou	Housing Starts:South (Thous.U.)S.A.	5
55	Starts: West	hswst	Housing Starts:West (Thous.U.)S.A.	5
56	BP: total	hsbr	Housing Authorized: Total New Priv Housing Units (Thous.,Saar)	5
57	BP: NE	hsbne	Houses Authorized By Build. Permits:Northeast(Thou.U.)S.A	5
58	BP: MW	hsbmw	Houses Authorized By Build. Permits:Midwest(Thou.U.)S.A.	5
59	BP: South	hsbsou	Houses Authorized By Build. Permits:South(Thou.U.)S.A.	5
60	BP: West	hsbwst	Houses Authorized By Build. Permits:West(Thou.U.)S.A.	5
Consumption, Orders and Inventories				
61	PMI	pmi	Purchasing Managers' Index (Sa)	1
62	NAPM new ordrs	pmno	Napm New Orders Index (Percent)	1
63	NAPM vendor del	pmdel	Napm Vendor Deliveries Index (Percent)	1
64	NAPM Invent	pmnv	Napm Inventories Index (Percent)	1
65	Orders: cons gds	a1m008	Mfrs' New Orders, Consumer Goods & Materials (Mil. Chain 1982 $) (TCB)	5
66	Orders: dble gds	a0m007	Mfrs' New Orders, Durable Goods Industries (Bil. Chain 2000 $) (TCB)	5
67	Orders: cap gds	a0m027	Mfrs' New Orders, Nondefense Capital Goods (Mil. Chain 1996 $) (TCB)	5
68	Unf orders: dble	a1m092	Mfrs' Unfilled Orders, Durable Goods Indus. (Bil. Chain 2000 $) (TCB)	5
69	M&T invent	a0m070	Manufacturing & Trade Inventories (Bil. Chain 2005 $) (TCB)	5
70	M&T invent/sales	a0m077	Ratio, Mfg. & Trade Inventories To Sales (Based On Chain 2005 $) (TCB)	2
71	Real Consumption	cons-r	Real Personal Consumption Expenditures (AC) (Bil. $) pi031 / gmdc	5
72	M&Tsales	mtq	Manufacturing & Trade Sales (Mil. Chain 1996 $)	5
73	Retail sales	a0m059	Sales Of Retail Stores (Mil. Chain 2000 $) (TCB)	5

Series no.	Name	Mnemonic	Description	transf
74	Consumer expect	hhsntn	U. Of Mich. Index Of Consumer Expectations(Bcd-83)	2
Money and Credit				
75	M1	fm1	Money Stock: M1(Curr,Trav.Cks,Dem Dep,Other Ck'able Dep)(Bil. $,Sa)	6
76	M2	fm2	Money Stock:M2(M1+O'nite Rps,Euro$,G/P&B/D & Mmmfs&Sav&Sm Time Dep)(Bil. $,Sa)	6
77	Currency	fmscu	Money Stock: Currency held by the public (Bil $,Sa)	6
78	M2 (real)	fm2-r	Money Supply: Real M2, fm2 / gmdc (AC)	5
79	MB	fmfba	Monetary Base, Adj For Reserve Requirement Changes(Mil. $,Sa)	6
80	Reserves tot	fmrra	Depository Inst Reserves:Total, Adj For Reserve Req Chgs(Mil. $,Sa)	6
81	C&I loans	fclnbw	Commercial & Industrial Loans Outstanding + NonFin Comm. Paper (Mil. $, SA) (Bci)	6
82	C&I loans	fclbmc	Wkly Rp Lg Com'l Banks:Net Change Com'l & Indus Loans(Bil$,Saar)	1
83	Cons credit	ccinrv	Consumer Credit Outstanding - Nonrevolving(G19)	6
84	Inst cred/PI	crdpi	Ratio, Consumer Installment Credit To Personal Income (Pct.) (TCB)	2
Interest Rates, Exchange Rates and Spreads				
85	Fed Funds	fyff	Interest Rate: Federal Funds (Effective) (% Per Annum,Nsa)	2
86	Comm paper	cpf3m	3-Month AA Financial Commercial Paper Rate (FRED)	2
87	3 mo T-bill	fygm3	Interest Rate: U.S.Treasury Bills,Sec Mkt,3-Mo.(% Per Ann,Nsa)	2
88	6 mo T-bill	fygm6	Interest Rate: U.S.Treasury Bills,Sec Mkt,6-Mo.(% Per Ann,Nsa)	2
89	1 yr T-bond	fygt1	Interest Rate: U.S.Treasury Const Maturities,1-Yr.(% Per Ann,Nsa)	2
90	5 yr T-bond	fygt5	Interest Rate: U.S.Treasury Const Maturities,5-Yr.(% Per Ann,Nsa)	2
91	10 yr T-bond	fygt10	Interest Rate: U.S.Treasury Const Maturities,10-Yr.(% Per Ann,Nsa)	2
92	Aaa bond	fyaaac	Bond Yield: Moody's Aaa Corporate (% Per Annum)	2
93	Baa bond	fybaac	Bond Yield: Moody's Baa Corporate (% Per Annum)	2
94	CP-FF spread	scp90F	cp90-fyff (AC)	1
95	3 mo-FF spread	sfygm3	fygm3-fyff (AC)	1
96	6 mo-FF spread	sfygm6	fygm6-fyff (AC)	1
97	1 yr-FF spread	sfygt1	fygt1-fyff (AC)	1
98	5 yr-FF spread	sfygt5	fygt5-fyff (AC)	1
99	10 yr-FF spread	sfygt10	fygt10-fyff (AC)	1
100	Aaa-FF spread	sfyaaac	fyaaac-fyff (AC)	1
101	Baa-FF spread	sfybaac	fybaac-fyff (AC)	1
102	Eff ex rate: US	exrus	United States;Effective Exchange Rate (Merm)(Index No.)	5
103	Ex rate: Switz	exrsw	Foreign Exchange Rate: Switzerland (Swiss Franc Per U.S.$)	5
104	Ex rate: Japan	exrjan	Foreign Exchange Rate: Japan (Yen Per U.S.$)	5
105	Ex rate: UK	exruk	Foreign Exchange Rate: United Kingdom (Cents Per Pound)	5
106	EX rate: Canada	exrcan	Foreign Exchange Rate: Canada (Canadian $Per U.S.$)	5
Prices				
107	PPI: fin gds	pwfsa	Producer Price Index: Finished Goods (82=100,Sa)	6
108	PPI: cons gds	pwfcsa	Producer Price Index: Finished Consumer Goods (82=100,Sa)	6
109	PPI: int materials	pwimsa	Producer Price Index: Intermed Mat.Supplies & Components(82=100,Sa)	6
110	PPI: crude materials	pwcmsa	Producer Price Index: Crude Materials (82=100,Sa)	6
111	Spot market price	psccom	Spot market price index: bls & crb: all commodities(1967=100)	6
112	PPI: nonferrous materials	pw102	Producer Price Index: Nonferrous Materials (1982=100, Nsa)	6
113	NAPM com price	pmcp	Napm Commodity Prices Index (Percent)	1
114	CPI-U: all	punew	Cpi-U: All Items (82-84=100,Sa)	6
115	CPI-U: apparel	pu83	Cpi-U: Apparel & Upkeep (82-84=100,Sa)	6
116	CPI-U:transp	pu84	Cpi-U: Transportation (82-84=100,Sa)	6
117	CPI-U: medical	pu85	Cpi-U: Medical Care (82-84=100,Sa)	6
118	CPI-U: comm.	puc	Cpi-U: Commodities (82-84=100,Sa)	6
119	CPI-U:dbles	pucd	Cpi-U: Durables (82-84=100,Sa)	6
120	CPI-U:services	pus	Cpi-U: Services (82-84=100,Sa)	6
121	CPI-U:exfood	puxf	Cpi-U: All Items Less Food (82-84=100,Sa)	6
122	CPI-U:exshelter	puxhs	Cpi-U: All Items Less Shelter (82-84=100,Sa)	6

Series no.	**Name**	**Mnemonic**	**Description**	**transf**
123	CPI-U:exmed	puxm	Cpi-U: All Items Less Midical Care (82-84=100,Sa)	6
124	PCEdefl	gmdc	Pce, Impl Pr Defl:Pce (2005=100, Sa) (BEA)	6
125	PCEdefl: dlbes	gmdcd	Pce, Impl Pr Defl:Pce; Durables (2005=100, Sa) (BEA)	6
126	PCEdefl: nondble	gmdcn	Pce, Impl Pr Defl:Pce; Nondurables (2005=100, Sa) (BEA)	6
127	PCEdefl: service	gmdcs	Pce, Impl Pr Defl:Pce; Services (2005=100, Sa) (BEA)	6
Stock Market				
128	S&P 500	fspcom	S&P's Common Stock Price Index: Composite (1941-43=10)	5
129	S&P: indust	fspin	S&P's Common Stock Price Index: & Industrials (1941-43=10)	5
130	S&P div yield	fsdxp	S&P's Composite Common Stock: Dividend Yield (% Per Annum)	2
131	S&P PE ratio	fspxe	S&P's Composite Common Stock: & Price-Earnings Ratio (%,Nsa)	5

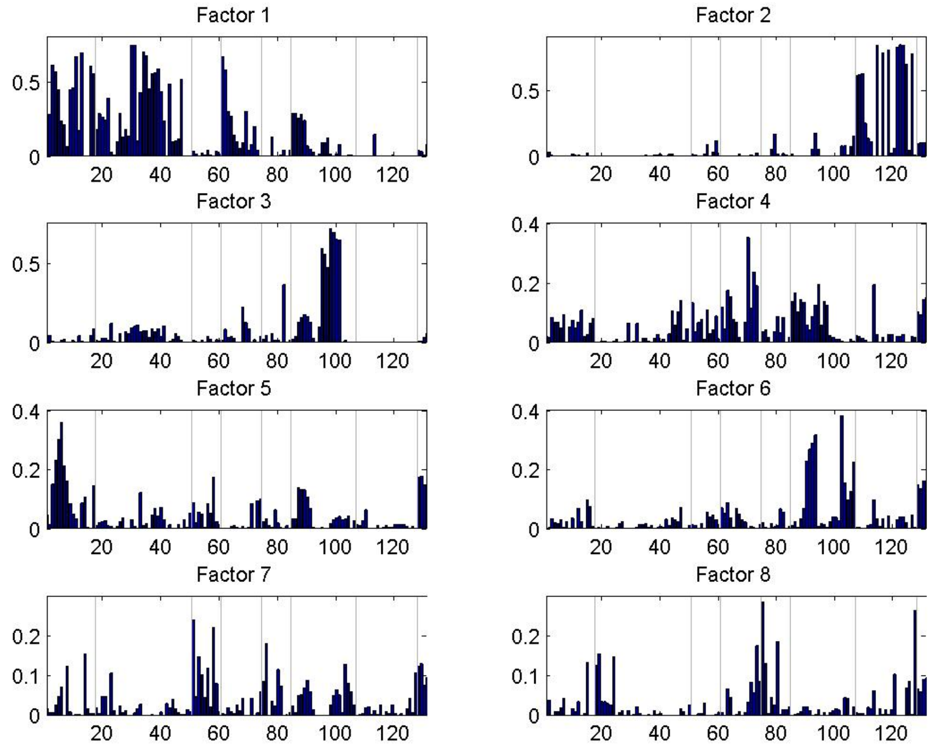
B.2 Estimation of Common Factors

Before proceeding to the common factors estimation, all the data are transformed appropriately in order to become stationary and are standardized. The procedure we follow for the estimation of the pervasive macroeconomic factors is the asymptotic principal component analysis (APCA) introduced by Connor and Korajczyk (1986) and widely used for summarizing latent information from large macroeconomic panels. Let N be the number of observed variables, T the number of time series observations and K the number of latent common factors. For $i = 1, ..., N$, $t = 1, ..., T$, and assuming a static factor model with approximate structure, a variable x_{it} can be written as:

$$x_{it} = \lambda_{i1} f_{1t} + \lambda_{i2} f_{2t} + ... + \lambda_{iK} f_{Kt} + e_{it}, \tag{B.1}$$

where x_{it} is the variable i at time t, λ_{ik} is the factor loading of variable i corresponding to the k^{th} factor, f_{kt} is the value of the k^{th} factor at time t and e_{it} is the idiosyncratic error of variable i at time t. The $T \times K$ factor matrix $\widehat{F}$ is calculated as $\sqrt{T}$ multiplied by the eigenvectors corresponding to the first K eigenvalues of the $T \times T$ matrix XX'. The normalization $F'F/T = I$ gives the solution for the factor loadings matrix as $\widehat{\Lambda}' = \widehat{F}'X/T$. The number of significant factors r is specified using the second information criterion proposed by Bai and Ng (2002) as it has been found to be the most stable. In our case, eight factors are found to explain sufficiently the macroeconomic variations.

Figure B.1: R^2s of common factors



This figure depicts the R^2s from simple univariate regressions of the eight common factors against each of the 131 macroeconomic variables. The variables' categories are output and income (series 1 to 17); employment (18-50); housing (51-60); consumption, orders and inventories (61-74); money and credit (75-84); interest rates, exchange rates and spreads (85-106); prices (107-127) and stock market (128-131). The sample is from 1990:01 to 2011:06.

B.3 Results with Alternative Macroeconomic Variables

Table B.2: Risk-neutral skewness and EF, EB sentiment components using alternative macroeconomic variables

	Adv-Sent			**Spec-Sent**			**Ind-Sent**		
				Panel A: 1990:01 - 1997:06					
LagRNS	0.249***	0.261***	0.242***	0.376***	0.369***	0.366***	0.434***	0.436***	0.445***
	(3.163)	(3.325)	(2.924)	(3.849)	(3.696)	(3.595)	(5.914)	(5.817)	(5.795)
RelDem		0.008	0.005		-0.002	-0.003		0.003	0.004
		(1.198)	(0.866)		(-0.270)	(-0.352)		(0.424)	(0.575)
TrVlm		0.040	0.031		-0.017	-0.019		0.004	0.008
		(0.577)	(0.449)		(-0.268)	(-0.300)		(0.055)	(0.115)
Vol			-0.766			-0.248			0.373
			(-1.160)			(-0.289)			(0.458)
EF	1.734***	1.785***	1.739***	4.013**	4.062**	3.820*	0.870***	0.876***	0.918***
	(4.984)	(5.204)	(5.122)	(2.031)	(2.057)	(1.681)	(3.909)	(3.866)	(3.431)
EB	0.817***	0.836***	0.844***	0.974	0.953	0.932	0.243	0.235	0.261
	(2.886)	(2.822)	(2.987)	(1.267)	(1.252)	(1.254)	(1.591)	(1.457)	(1.537)
$\widetilde{\mathbf{R}}^2$	0.409	0.400	0.401	0.253	0.236	0.227	0.327	0.311	0.304
				Panel B: 1997:07 - 2011:06					
LagRNS	0.465***	0.402***	0.401***	0.482***	0.415***	0.402***	0.485***	0.433***	0.420***
	(6.300)	(4.938)	(4.947)	(7.133)	(5.484)	(5.220)	(6.289)	(5.126)	(4.972)
RelDem		-0.038	-0.035		-0.066	-0.053		-0.043	-0.035
		(-0.939)	(-0.859)		(-1.637)	(-1.300)		(-1.034)	(-0.859)
TrVlm		-0.140***	-0.138***		-0.132***	-0.137***		-0.127**	-0.128**
		(-2.903)	(-2.879)		(-2.770)	(-2.856)		(-2.538)	(-2.558)
Vol			0.285			0.492			0.533*
			(0.663)			(1.580)			(1.714)
EF	-0.788***	-0.955***	-0.811**	2.842**	3.901***	3.334**	-0.670**	-0.796***	-0.559*
	(-3.166)	(-3.587)	(-2.289)	(2.167)	(3.050)	(2.362)	(-2.514)	(-2.906)	(-1.733)
EB	-0.017	-0.131	-0.048	-0.805	-0.729	-0.579	-0.007	-0.113	0.001
	(-0.062)	(-0.489)	(-0.146)	(-1.167)	(-1.147)	(-0.874)	(-0.023)	(-0.414)	(0.002)
$\widetilde{\mathbf{R}}^2$	0.277	0.302	0.299	0.274	0.302	0.304	0.267	0.288	0.290

This table reports the results of monthly regressions of S&P 500 index risk-neutral skewness on the EF and EB components of the sentiment proxies used in the study and a set of control variables. A constant term is included in all the regressions but omitted for brevity. Panel A reports the results for the period 1990:1 - 1997:6, while Panel B reports the results for the period 1997:07 - 2011:06. Risk-neutral skewness is estimated using the model-free method of Bakshi, Kapadia and Madan (2003). LagRNS is the lagged skewness value. RelDem is the relative demand pressure as captured by the ratio of the open interest of OTM puts to the open interest of NTM calls and puts. TrVlm is the heterogeneity of beliefs, proxied by the detrended logarithm of options trading volume. Vol is the index instantaneous volatility as proxied by VIX. Adv-Sent is the bull-bear spread based on Investors Intelligence's advisors sentiment index. Spec-Sent is the net position of non-commercial traders on S&P 500 index futures scaled by the total open interest. Ind-Sent is the normalized aggregate net exchanges of the equity funds. EF and EB are the two components of each sentiment proxy estimated as described in Section 3.3.3. Newey-West t-statistics are reported in parentheses below the coefficients. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Appendix C

Appendix to Chapter 4

C.1 Macroeconomic Variables

Table C.1: Description of macroeconomic variables

Name	Description	Source
	Real Activity	
Pers income	Personal Income Account, Overall, Total (Current Prices, AR, SA, Billions $)	US BEA
Ind prod	Industrial Production, Overall, Total (Volume, SA, 2007=100)	US FED
Cap util	Capacity Utilization, Total index (SA, %)	US FED
Unempl	Unemployment Rate, Total (SA,%)	US BLS
Payroll	Employment, Overall, Nonfarm payroll, total (SA, Thousands)	US BLS
House starts	Housing Starts, Total (AR, SA, Thousands)	US CB
Build perm	Building Permits, Total (AR, SA, Thousands)	US CB
M&T invent	Manufacturing and Trade Inventories (SA, Billions $, 2009=100)	TCB
Consumption	Personal Consumption Expenditure, Overall, Total (AR, SA, Billions $, 2009=100)	US BEA
Retail sales	Retail Sales, Total excluding food services (Current Prices, SA, Millions $)	US CB
	Money, Credit and Treasury Yields	
M1	Money supply M1 (Current Prices, SA, Billions $)	US FED
M2 (real)	Money supply M2 (Current Prices, SA, Billions $) / Price Index, Personal Consumption Expenditure, Overall, Total (SA, Index, 2009=100)	US FED US BEA
Reserves tot	Reserves Depository Institutions, Total reserves (Current Prices, SA, Millions $)	US FED
C&I loans	Commercial and Industrial Loans Outstanding (Current Prices, SA, Millions $)	TCB
CPI	Consumer Prices, All items (SA, Index, 1982-1984=100)	US BLS
3-m t-bill	Interest Rate: US Treasury Bills, Secondary Market, 3-Month (% Per Annum, NSA)	US FED
6-m t-bill	Interest Rate: US Treasury Bills, Secondary Market, 6-Month (% Per Annum, NSA)	US FED
1-yr t-bond	Interest Rate: US Treasury Constant Maturities, 1-Year (% Per Annum, NSA)	US FED
5-yr t-bond	Interest Rate: US Treasury Constant Maturities, 5-Year (% Per Annum, NSA)	US FED

This table lists the name, a brief description and the source of each macroeconomic variable. Panel A lists the real activity variables while Panel B the money, credit and treasury yield variables. The data sources are the US Bureau of Economic Analysis (US BEA), the US Federal Reserve (US FED), the US Bureau of Labor Statistics (US BLS), the US Census Bureau (US CB) and the Conference Board (TCB).

C.2 Detailed Results with Newey and West (1987) Covariance Matrix

Table C.2: Predicting real activity for 6-month horizon - Newey-West covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
Pers income	-0.852**	-2.346**	-0.214	0.182	0.940					0.276		
	(-2.592)	(-2.432)	(-0.260)	(0.236)	(1.596)						0.000	
	-0.680**	-2.128**	-0.491	0.294	0.863*	-0.398	0.074	0.639***	-0.753**	0.314		
	(-2.049)	(-2.503)	(-0.577)	(0.558)	(1.696)	(-1.046)	(0.214)	(2.649)	(-2.357)		0.000	0.054
Ind prod	0.404	-3.556***	0.317	-0.496	1.806**					0.155		
	(0.771)	(-3.279)	(0.353)	(-0.512)	(2.116)						0.000	
	0.744*	-2.853***	-0.382	-0.390	1.701**	-1.676***	0.567	0.978**	-0.833	0.219		
	(1.680)	(-2.926)	(-0.390)	(-0.521)	(2.269)	(-3.802)	(1.090)	(2.434)	(-1.350)		0.000	0.000
Cap util	1.579***	-2.463***	-0.400	0.160	1.276**					0.270		
	(3.935)	(-3.207)	(-0.548)	(0.242)	(2.179)						0.000	
	1.724***	-1.989***	-0.752	0.096	1.168*	-0.714*	0.114	0.210	-0.599	0.305		
	(4.547)	(-2.685)	(-0.902)	(0.196)	(1.951)	(-1.910)	(0.301)	(0.846)	(-1.371)		0.000	0.404
Unempl	-0.064	0.759**	0.028	0.169	-0.347					0.268		
	(-0.664)	(2.398)	(0.103)	(0.554)	(-1.133)						0.000	
	-0.134	0.576**	0.189	0.167	-0.306	0.240**	-0.060	-0.174*	0.344***	0.358		
	(-1.612)	(2.439)	(0.786)	(0.885)	(-1.357)	(2.232)	(-0.602)	(-1.744)	(3.305)		0.000	0.019
Payroll	-0.419**	-1.192**	0.023	-0.225	0.444					0.332		
	(-2.117)	(-2.517)	(0.046)	(-0.536)	(0.815)						0.000	
	-0.289*	-0.891**	-0.255	-0.205	0.382	-0.537***	0.147	0.344***	-0.482***	0.431		
	(-1.793)	(-2.218)	(-0.537)	(-0.684)	(0.893)	(-3.713)	(0.913)	(2.736)	(-3.043)		0.000	0.001
House starts	5.647*	-14.433	8.224	1.310	3.777					0.061		
	(1.917)	(-1.507)	(1.123)	(0.177)	(0.856)						0.077	
	6.542**	-15.792*	7.920	3.149	3.605	1.587	-0.612	5.930**	-4.680	0.089		
	(2.434)	(-1.714)	(0.995)	(0.425)	(0.852)	(0.634)	(-0.278)	(2.215)	(-1.322)		0.073	0.166
Build perm	5.586*	-13.050*	8.027	1.157	5.264					0.076		
	(1.845)	(-1.718)	(1.422)	(0.174)	(1.276)						0.027	
	6.811**	-13.349*	6.739	3.008	4.937	0.652	-0.054	6.825***	-6.809*	0.131		
	(2.460)	(-1.909)	(1.081)	(0.434)	(1.250)	(0.267)	(-0.020)	(2.895)	(-1.779)		0.026	0.022
M&T invent	-0.429	-1.417	-0.143	-0.867	0.690					0.236		
	(-1.034)	(-1.400)	(-0.123)	(-0.816)	(0.652)						0.000	
	-0.194	-0.821	-0.636	-0.881	0.543	-1.013***	0.171	0.517*	-0.899***	0.338		
	(-0.527)	(-0.880)	(-0.562)	(-1.152)	(0.661)	(-3.068)	(0.528)	(1.800)	(-3.088)		0.000	0.000
Consumption	-0.475*	-1.000*	0.474	0.050	0.194					0.069		
	(-1.796)	(-1.969)	(0.828)	(0.114)	(0.332)						0.082	
	-0.347	-0.858*	0.232	0.177	0.193	-0.426**	0.265	0.555***	-0.373	0.124		
	(-1.455)	(-1.767)	(0.401)	(0.528)	(0.357)	(-1.985)	(1.281)	(3.015)	(-1.446)		0.167	0.002
Retail sales	-0.123	-3.703*	0.232	2.247*	0.882					0.048		
	(-0.263)	(-1.867)	(0.182)	(1.690)	(0.625)						0.199	
	0.216	-3.056*	-0.417	2.361**	0.705	-1.027*	0.213	1.043**	-1.505*	0.121		
	(0.637)	(-1.755)	(-0.337)	(2.049)	(0.636)	(-1.878)	(0.514)	(2.211)	(-1.781)		0.158	0.206

This table reports the results of predictive regressions of 6-month growth (or change) in real activity variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table C.3: Predicting real activity for 12-month horizon - Newey-West covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
Pers income	-0.537	-2.043***	0.372	-0.064	0.702					0.193		
	(-1.579)	(-3.611)	(0.795)	(-0.119)	(1.361)						0.000	
	-0.351	-2.102***	0.244	0.140	0.637	-0.261	-0.235	0.766**	-0.554	0.244		
	(-1.087)	(-4.161)	(0.429)	(0.413)	(1.423)	(-0.763)	(-0.660)	(2.573)	(-1.589)		0.000	0.084
Ind prod	0.678	-2.269***	0.351	-0.530	1.630*					0.063		
	(1.190)	(-2.847)	(0.471)	(-0.577)	(1.932)						0.000	
	0.996*	-2.068***	-0.023	-0.266	1.523*	-1.158***	0.071	1.147**	-0.574	0.117		
	(1.949)	(-2.608)	(-0.028)	(-0.383)	(1.942)	(-3.464)	(0.117)	(2.013)	(-1.142)		0.000	0.001
Cap util	1.664***	-1.484***	-0.212	-0.027	1.223**					0.257		
	(4.045)	(-2.894)	(-0.402)	(-0.052)	(2.257)						0.000	
	1.800***	-1.548***	-0.236	0.057	1.136**	-0.224	-0.443	0.380	-0.337	0.279		
	(4.566)	(-3.000)	(-0.391)	(0.122)	(2.019)	(-0.855)	(-1.283)	(1.045)	(-0.985)		0.000	0.482
Unempl	-0.199*	0.634***	-0.046	0.147	-0.333					0.169		
	(-1.693)	(3.716)	(-0.268)	(0.665)	(-1.295)						0.000	
	-0.268***	0.531***	0.081	0.122	-0.310	0.203*	0.012	-0.193*	0.254**	0.243		
	(-2.727)	(3.780)	(0.392)	(0.853)	(-1.480)	(1.961)	(0.103)	(-1.703)	(2.233)		0.000	0.149
Payroll	-0.201	-1.091***	0.140	-0.176	0.427					0.181		
	(-0.794)	(-3.040)	(0.437)	(-0.463)	(0.912)						0.000	
	-0.061	-0.885***	-0.115	-0.107	0.385	-0.487***	0.060	0.427**	-0.428**	0.279		
	(-0.288)	(-2.727)	(-0.320)	(-0.486)	(1.032)	(-2.905)	(0.294)	(2.505)	(-2.394)		0.000	0.009
House starts	7.063**	-5.031	7.464	-2.934	2.665					0.128		
	(2.410)	(-1.190)	(1.588)	(-0.612)	(0.554)						0.115	
	8.005***	-6.428	6.920	-0.773	2.718	0.598	0.379	6.532**	-3.490*	0.192		
	(2.831)	(-1.552)	(1.302)	(-0.211)	(0.579)	(0.361)	(0.175)	(2.378)	(-1.892)		0.071	0.065
Build perm	7.307**	-3.725	7.958	-2.611	1.599					0.151		
	(2.327)	(-0.977)	(1.536)	(-0.507)	(0.286)						0.130	
	8.340***	-4.699	7.202	-0.588	1.568	-0.658	0.585	6.375**	-3.127	0.204		
	(2.751)	(-1.181)	(1.220)	(-0.127)	(0.278)	(-0.363)	(0.271)	(2.377)	(-1.631)		0.097	0.153
M&T invent	-0.030	-1.996***	0.281	-0.620	0.841					0.203		
	(-0.067)	(-2.654)	(0.404)	(-0.815)	(1.021)						0.000	
	0.216	-1.564**	-0.194	-0.546	0.756	-0.964***	0.095	0.649**	-0.730**	0.304		
	(0.562)	(-2.253)	(-0.247)	(-1.220)	(1.162)	(-2.821)	(0.238)	(2.326)	(-2.118)		0.000	0.020
Consumption	-0.303	-0.873**	0.687	-0.338	0.378					0.036		
	(-0.906)	(-2.119)	(1.436)	(-0.797)	(0.689)						0.000	
	-0.179	-0.816*	0.508	-0.158	0.372	-0.424**	0.270	0.641***	-0.228	0.118		
	(-0.600)	(-1.852)	(1.028)	(-0.529)	(0.740)	(-2.230)	(1.193)	(2.739)	(-1.058)		0.007	0.004
Retail sales	0.362	-2.245***	0.524	0.618	1.099					0.031		
	(0.681)	(-2.618)	(0.658)	(0.553)	(0.873)						0.007	
	0.657	-2.269***	0.292	0.909	0.991	-0.562	-0.315	1.144*	-0.808*	0.086		
	(1.389)	(-3.187)	(0.343)	(1.131)	(0.914)	(-1.224)	(-0.641)	(1.784)	(-1.827)		0.004	0.291

This table reports the results of predictive regressions of 12-month growth (or change) in real activity variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table C.4: Predicting money, credit and yield variables for 6-month horizon - Newey-West covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
M1	3.110***	0.370	-0.308	1.050	0.357					0.323		
	(3.615)	(0.364)	(-0.285)	(0.852)	(0.272)						0.044	
	3.028***	0.690	-0.339	0.758	0.256	0.266	-0.283	-0.739*	-0.326	0.331		
	(3.445)	(0.693)	(-0.282)	(0.708)	(0.211)	(0.313)	(-0.437)	(-1.759)	(-0.473)		0.028	0.171
M2 (real)	-0.174	2.233**	-0.050	-1.250***	-0.407					0.077		
	(-0.406)	(2.348)	(-0.058)	(-2.774)	(-0.747)						0.000	
	-0.175	2.357**	-0.121	-1.301***	-0.390	-0.319	0.164	-0.119	0.205	0.067		
	(-0.404)	(2.480)	(-0.137)	(-2.840)	(-0.709)	(-0.640)	(0.450)	(-0.616)	(0.580)		0.000	0.859
Reserves tot	12.192*	34.963	-24.483*	-13.001	12.454					0.018		
	(1.787)	(1.358)	(-1.792)	(-0.761)	(1.484)						0.135	
	8.559	27.782	-18.055	-13.600	14.964	10.938	-0.121	-9.856	16.644	0.045		
	(1.454)	(1.165)	(-1.199)	(-0.660)	(1.492)	(1.307)	(-0.020)	(-1.284)	(1.054)		0.077	0.683
C&I loans	-4.114**	-4.776	-2.043	0.760	1.357					0.243		
	(-1.993)	(-1.193)	(-0.534)	(0.242)	(0.264)						0.003	
	-3.558**	-2.344	-3.235	-0.127	0.499	-3.490***	-0.789	-0.556	-2.352***	0.413		
	(-2.132)	(-0.678)	(-0.916)	(-0.052)	(0.120)	(-2.672)	(-0.897)	(-0.560)	(-2.629)		0.000	0.000
CPI	-0.295*	-1.126*	-0.017	0.864**	-0.091					0.101		
	(-1.762)	(-1.773)	(-0.037)	(2.169)	(-0.357)						0.028	
	-0.261*	-1.016	-0.099	0.849**	-0.132	-0.021	-0.041	0.064	-0.310	0.113		
	(-1.655)	(-1.583)	(-0.203)	(2.136)	(-0.531)	(-0.111)	(-0.194)	(0.407)	(-1.602)		0.026	0.515
3-m t-bill	0.245	-0.689**	-0.156	0.262	0.376					0.061		
	(1.111)	(-2.059)	(-0.524)	(1.055)	(1.120)						0.109	
	0.293	-0.634**	-0.221	0.290	0.354	-0.175	0.019	0.164**	-0.132	0.060		
	(1.289)	(-2.007)	(-0.769)	(1.277)	(1.096)	(-1.099)	(0.128)	(2.066)	(-1.176)		0.081	0.280
6-m t-bill	0.268	-0.846***	0.009	0.250	0.321					0.083		
	(1.211)	(-2.603)	(0.030)	(1.028)	(0.972)						0.022	
	0.319	-0.792***	-0.051	0.276	0.292	-0.181	-0.002	0.164**	-0.138	0.087		
	(1.415)	(-2.660)	(-0.191)	(1.229)	(0.932)	(-1.173)	(-0.016)	(2.347)	(-1.258)		0.018	0.184
1-yr t-bond	0.256	-0.946***	0.179	0.280	0.244					0.079		
	(1.130)	(-2.778)	(0.648)	(1.221)	(0.797)						0.013	
	0.304	-0.897***	0.128	0.302	0.212	-0.156	-0.023	0.150**	-0.145	0.081		
	(1.332)	(-2.872)	(0.495)	(1.394)	(0.727)	(-1.021)	(-0.142)	(2.071)	(-1.300)		0.011	0.293
5-yr t-bond	-0.132	-1.000***	0.485*	0.738***	-0.056					0.056		
	(-0.614)	(-3.050)	(1.677)	(3.085)	(-0.292)						0.014	
	-0.113	-1.074***	0.525	0.780***	-0.077	0.106	-0.114	0.129	-0.122	0.051		
	(-0.537)	(-3.314)	(1.638)	(3.561)	(-0.375)	(0.678)	(-0.645)	(1.409)	(-1.221)		0.002	0.420

This table reports the results of predictive regressions of 6-month growth (or change) in money, credit and yield variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table C.5: Predicting money, credit and yield variables for 12-month horizon - Newey-West covariance matrix

	TERM	**FV$^{(1)}$**	**FV$^{(2)}$**	**FV$^{(3)}$**	**FV$^{(4)}$**	**FSC$^{(1)}$**	**FSC$^{(2)}$**	**FSC$^{(3)}$**	**FSC$^{(4)}$**	**Adj. R^2**	**Joint FV**	**Joint FSC**
M1	2.702***	-0.353	-0.212	2.182	-0.406					0.339		
	(3.162)	(-0.328)	(-0.218)	(1.557)	(-0.383)						0.140	
	2.581***	0.682	-0.870	1.701	-0.313	-0.156	0.765*	-1.055**	-0.588	0.369		
	(3.055)	(0.816)	(-0.750)	(1.515)	(-0.318)	(-0.257)	(1.720)	(-1.996)	(-1.026)		0.020	0.139
M2 (real)	-0.369	0.783	0.143	-0.320	-0.458					0.034		
	(-0.895)	(1.397)	(0.232)	(-0.666)	(-1.313)						0.186	
	-0.394	1.303**	-0.230	-0.442	-0.391	-0.471	0.783***	-0.198	-0.048	0.066		
	(-0.974)	(2.419)	(-0.337)	(-0.880)	(-1.179)	(-1.441)	(3.410)	(-0.934)	(-0.171)		0.000	0.012
Reserves tot	5.158	9.496	-6.415	16.364	-10.484					0.010		
	(0.919)	(0.753)	(-0.645)	(1.010)	(-0.821)						0.004	
	1.690	14.292	-7.062	11.773	-8.838	3.284	8.651	-15.066	7.187	0.035		
	(0.306)	(1.325)	(-0.647)	(0.821)	(-0.779)	(0.546)	(1.041)	(-1.409)	(0.932)		0.007	0.724
C&I loans	-2.995	-7.845**	-0.001	1.454	1.536					0.260		
	(-1.320)	(-2.090)	(0.000)	(0.455)	(0.385)						0.000	
	-2.195	-5.667*	-1.770	0.920	1.036	-3.485***	-1.068	0.370	-2.977***	0.482		
	(-1.318)	(-1.777)	(-0.623)	(0.469)	(0.344)	(-2.757)	(-0.895)	(0.331)	(-4.101)		0.000	0.000
CPI	-0.203	-0.556**	-0.207	0.517**	-0.010					0.095		
	(-1.335)	(-2.022)	(-0.559)	(2.411)	(-0.034)						0.000	
	-0.144	-0.657***	-0.163	0.547***	-0.063	0.077	-0.405***	0.135	-0.237**	0.222		
	(-1.194)	(-2.684)	(-0.459)	(2.911)	(-0.256)	(0.645)	(-3.154)	(1.142)	(-2.264)		0.000	0.004
3-m t-bill	0.449*	-0.675*	0.050	0.164	0.293					0.138		
	(1.707)	(-1.654)	(0.168)	(0.675)	(0.776)						0.365	
	0.499*	-0.694*	0.025	0.215	0.271	-0.095	-0.076	0.193*	-0.119	0.142		
	(1.869)	(-1.813)	(0.088)	(0.921)	(0.767)	(-0.536)	(-0.418)	(1.875)	(-1.141)		0.355	0.371
6-m t-bill	0.452*	-0.761*	0.149	0.167	0.246					0.154		
	(1.705)	(-1.908)	(0.506)	(0.696)	(0.666)						0.194	
	0.503*	-0.774**	0.118	0.215	0.225	-0.092	-0.080	0.190**	-0.134	0.161		
	(1.886)	(-2.092)	(0.417)	(0.917)	(0.653)	(-0.525)	(-0.456)	(1.976)	(-1.256)		0.206	0.313
1-yr t-bond	0.421	-0.817**	0.275	0.185	0.177					0.147		
	(1.576)	(-2.052)	(0.904)	(0.811)	(0.506)						0.139	
	0.466*	-0.831**	0.247	0.227	0.158	-0.058	-0.081	0.169*	-0.140	0.152		
	(1.750)	(-2.244)	(0.841)	(0.990)	(0.484)	(-0.330)	(-0.486)	(1.728)	(-1.274)		0.155	0.392
5-yr t-bond	0.007	-0.625**	0.486*	0.330*	-0.088					0.045		
	(0.034)	(-2.172)	(1.713)	(1.831)	(-0.343)						0.223	
	0.010	-0.686**	0.518*	0.353*	-0.090	0.097	-0.075	0.049	-0.041	0.033		
	(0.051)	(-2.553)	(1.840)	(1.900)	(-0.344)	(0.670)	(-0.705)	(0.558)	(-0.431)		0.134	0.824

This table reports the results of predictive regressions of 12-month growth (or change) in money, credit and yield variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Newey and West (1987) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

C.3 Detailed Results with Hodrick (1992) Covariance Matrix

Table C.6: Predicting real activity for 6-month horizon - Hodrick covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
Pers income	-0.852***	-2.346**	-0.214	0.182	0.940*					0.276		
	(-2.682)	(-2.316)	(-0.316)	(0.264)	(1.888)						0.033	
	-0.680**	-2.128**	-0.491	0.294	0.863*	-0.398	0.074	0.639***	-0.753**	0.314		
	(-2.162)	(-2.144)	(-0.698)	(0.423)	(1.772)	(-1.553)	(0.375)	(2.754)	(-2.588)		0.026	0.017
Ind prod	0.404	-3.556**	0.317	-0.496	1.806***					0.155		
	(0.882)	(-2.338)	(0.337)	(-0.626)	(2.619)						0.026	
	0.744	-2.853*	-0.382	-0.390	1.701**	-1.676***	0.567*	0.978***	-0.833*	0.219		
	(1.653)	(-1.931)	(-0.400)	(-0.497)	(2.491)	(-4.597)	(1.798)	(3.169)	(-1.739)		0.023	0.000
Cap util	1.579***	-2.463**	-0.400	0.160	1.276*					0.270		
	(4.017)	(-2.295)	(-0.533)	(0.271)	(1.951)						0.006	
	1.724***	-1.989*	-0.752	0.096	1.168*	-0.714**	0.114	0.210	-0.599*	0.305		
	(4.433)	(-1.905)	(-0.962)	(0.165)	(1.767)	(-2.319)	(0.463)	(0.900)	(-1.701)		0.006	0.115
Unempl	-0.064	0.759**	0.028	0.169	-0.347*					0.268		
	(-0.586)	(2.377)	(0.121)	(0.609)	(-1.834)						0.020	
	-0.134	0.576*	0.189	0.167	-0.306	0.240**	-0.060	-0.174**	0.344***	0.358		
	(-1.238)	(1.922)	(0.783)	(0.613)	(-1.624)	(2.547)	(-0.759)	(-1.985)	(3.335)		0.018	0.002
Payroll	-0.419***	-1.192***	0.023	-0.225	0.444*					0.332		
	(-3.867)	(-2.798)	(0.077)	(-0.823)	(1.799)						0.008	
	-0.289***	-0.891**	-0.255	-0.205	0.382	-0.537***	0.147*	0.344***	-0.482***	0.431		
	(-2.780)	(-2.156)	(-0.843)	(-0.748)	(1.572)	(-5.503)	(1.819)	(3.767)	(-4.605)		0.006	0.000
House starts	5.647	-14.433	8.224	1.310	3.777					0.061		
	(0.982)	(-1.287)	(0.861)	(0.144)	(0.443)						0.675	
	6.542	-15.792	7.920	3.149	3.605	1.587	-0.612	5.930*	-4.680	0.089		
	(1.121)	(-1.488)	(0.759)	(0.350)	(0.421)	(0.393)	(-0.167)	(1.866)	(-1.192)		0.633	0.402
Build perm	5.586	-13.050*	8.027	1.157	5.264					0.076		
	(1.531)	(-1.733)	(1.283)	(0.162)	(0.844)						0.191	
	6.811*	-13.349*	6.739	3.008	4.937	0.652	-0.054	6.825**	-6.809**	0.131		
	(1.874)	(-1.854)	(0.979)	(0.429)	(0.796)	(0.225)	(-0.018)	(2.456)	(-2.208)		0.277	0.059
M&T invent	-0.429	-1.417*	-0.143	-0.867	0.690					0.236		
	(-1.605)	(-1.892)	(-0.239)	(-1.283)	(1.307)						0.020	
	-0.194	-0.821	-0.636	-0.881	0.543	-1.013***	0.171	0.517**	-0.899***	0.338		
	(-0.740)	(-1.041)	(-0.992)	(-1.272)	(1.035)	(-4.164)	(0.920)	(2.369)	(-3.917)		0.019	0.000
Consumption	-0.475*	-1.000*	0.474	0.050	0.194					0.069		
	(-1.731)	(-1.778)	(1.007)	(0.134)	(0.318)						0.379	
	-0.347	-0.858	0.232	0.177	0.193	-0.426*	0.265	0.555***	-0.373**	0.124		
	(-1.298)	(-1.517)	(0.478)	(0.505)	(0.316)	(-1.719)	(1.640)	(3.562)	(-2.025)		0.540	0.001
Retail sales	-0.123	-3.703*	0.232	2.247*	0.882					0.048		
	(-0.157)	(-1.875)	(0.160)	(1.741)	(0.582)						0.351	
	0.216	-3.056	-0.417	2.361*	0.705	-1.027	0.213	1.043**	-1.505**	0.121		
	(0.284)	(-1.637)	(-0.274)	(1.878)	(0.462)	(-1.448)	(0.430)	(2.411)	(-2.357)		0.337	0.073

This table reports the results of predictive regressions of 6-month growth (or change) in real activity variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table C.7: Predicting real activity for 12-month horizon - Hodrick covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
Pers income	-0.537	-2.043***	0.372	-0.064	0.702*					0.193		
	(-1.380)	(-3.512)	(0.998)	(-0.139)	(1.859)						0.011	
	-0.351	-2.102***	0.244	0.140	0.637*	-0.261	-0.235	0.766***	-0.554**	0.244		
	(-0.895)	(-3.768)	(0.585)	(0.324)	(1.711)	(-1.118)	(-1.107)	(3.725)	(-2.260)		0.005	0.003
Ind prod	0.678	-2.269***	0.351	-0.530	1.630**					0.063		
	(1.587)	(-2.606)	(0.506)	(-0.964)	(2.559)						0.032	
	0.996**	-2.068**	-0.023	-0.266	1.523**	-1.158***	0.071	1.147***	-0.574*	0.117		
	(2.352)	(-2.369)	(-0.031)	(-0.540)	(2.425)	(-3.546)	(0.227)	(3.100)	(-1.953)		0.054	0.001
Cap util	1.664***	-1.484**	-0.212	-0.027	1.223**					0.257		
	(4.612)	(-2.371)	(-0.368)	(-0.064)	(2.109)						0.069	
	1.800***	-1.548**	-0.236	0.057	1.136*	-0.224	-0.443*	0.380	-0.337	0.279		
	(5.028)	(-2.499)	(-0.391)	(0.150)	(1.957)	(-0.835)	(-1.894)	(1.363)	(-1.584)		0.048	0.104
Unempl	-0.199*	0.634***	-0.046	0.147	-0.333**					0.169		
	(-1.768)	(3.579)	(-0.298)	(0.895)	(-2.144)						0.001	
	-0.268**	0.531***	0.081	0.122	-0.310**	0.203***	0.012	-0.193***	0.254***	0.243		
	(-2.377)	(3.250)	(0.463)	(0.802)	(-2.006)	(2.658)	(0.165)	(-2.635)	(3.644)		0.002	0.001
Payroll	-0.201*	-1.091***	0.140	-0.176	0.427**					0.181		
	(-1.951)	(-4.561)	(0.828)	(-1.154)	(2.394)						0.000	
	-0.061	-0.885***	-0.115	-0.107	0.385**	-0.487***	0.060	0.427***	-0.428***	0.279		
	(-0.610)	(-3.839)	(-0.661)	(-0.729)	(2.200)	(-6.117)	(0.784)	(4.929)	(-4.679)		0.000	0.000
House starts	7.063	-5.031	7.464	-2.934	2.665					0.128		
	(1.226)	(-0.786)	(1.049)	(-0.444)	(0.359)						0.575	
	8.005	-6.428	6.920	-0.773	2.718	0.598	0.379	6.532**	-3.490	0.192		
	(1.371)	(-1.141)	(0.852)	(-0.126)	(0.365)	(0.186)	(0.122)	(2.377)	(-1.202)		0.612	0.151
Build perm	7.307**	-3.725	7.958*	-2.611	1.599					0.151		
	(2.066)	(-0.778)	(1.761)	(-0.539)	(0.314)						0.219	
	8.340**	-4.699	7.202	-0.588	1.568	-0.658	0.585	6.375***	-3.127*	0.204		
	(2.349)	(-1.047)	(1.425)	(-0.131)	(0.307)	(-0.300)	(0.301)	(2.844)	(-1.816)		0.359	0.026
M&T invent	-0.030	-1.996***	0.281	-0.620	0.841**					0.203		
	(-0.115)	(-4.151)	(0.815)	(-1.558)	(2.323)						0.000	
	0.216	-1.564***	-0.194	-0.546	0.756**	-0.964***	0.095	0.649***	-0.730***	0.304		
	(0.822)	(-3.627)	(-0.512)	(-1.403)	(2.107)	(-4.676)	(0.625)	(4.004)	(-3.622)		0.001	0.000
Consumption	-0.303	-0.873*	0.687*	-0.338	0.378					0.036		
	(-1.135)	(-1.843)	(1.882)	(-0.993)	(0.990)						0.018	
	-0.179	-0.816*	0.508	-0.158	0.372	-0.424*	0.270*	0.641***	-0.228*	0.118		
	(-0.670)	(-1.900)	(1.330)	(-0.492)	(0.984)	(-1.842)	(1.870)	(4.646)	(-1.713)		0.075	0.000
Retail sales	0.362	-2.245*	0.524	0.618	1.099					0.031		
	(0.492)	(-1.664)	(0.510)	(0.624)	(1.014)						0.503	
	0.657	-2.269*	0.292	0.909	0.991	-0.562	-0.315	1.144**	-0.808**	0.086		
	(0.884)	(-1.805)	(0.270)	(0.960)	(0.924)	(-0.862)	(-0.742)	(2.435)	(-2.052)		0.488	0.141

This table reports the results of predictive regressions of 12-month growth (or change) in real activity variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table C.8: Predicting money, credit and yield variables for 6-month horizon - Hodrick covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
M1	3.110***	0.370	-0.308	1.050	0.357					0.323		
	(4.054)	(0.162)	(-0.218)	(0.887)	(0.322)						0.584	
	3.028***	0.690	-0.339	0.758	0.256	0.266	-0.283	-0.739*	-0.326	0.331		
	(3.789)	(0.314)	(-0.232)	(0.638)	(0.234)	(0.346)	(-0.589)	(-1.811)	(-0.549)		0.734	0.118
M2 (real)	-0.174	2.233**	-0.050	-1.250***	-0.407					0.077		
	(-0.450)	(2.100)	(-0.068)	(-2.612)	(-0.604)						0.032	
	-0.175	2.357**	-0.121	-1.301***	-0.390	-0.319	0.164	-0.119	0.205	0.067		
	(-0.444)	(2.319)	(-0.164)	(-2.692)	(-0.579)	(-0.834)	(0.740)	(-0.557)	(0.800)		0.027	0.646
Reserves tot	12.192***	34.963	-24.483	-13.001	12.454					0.018		
	(2.650)	(1.511)	(-1.505)	(-1.054)	(1.466)						0.212	
	8.559**	27.782	-18.055	-13.600	14.964*	10.938*	-0.121	-9.856*	16.644	0.045		
	(2.550)	(1.382)	(-1.261)	(-1.077)	(1.722)	(1.807)	(-0.044)	(-1.831)	(1.530)		0.175	0.165
C&I loans	-4.114***	-4.776**	-2.043	0.760	1.357					0.243		
	(-2.946)	(-2.232)	(-0.837)	(0.359)	(0.473)						0.002	
	-3.558**	-2.344	-3.235	-0.127	0.499	-3.490***	-0.789	-0.556	-2.352***	0.413		
	(-2.471)	(-1.053)	(-1.250)	(-0.060)	(0.172)	(-3.441)	(-0.978)	(-0.740)	(-3.048)		0.003	0.000
CPI	-0.295	-1.126*	-0.017	0.864**	-0.091					0.101		
	(-1.609)	(-1.836)	(-0.042)	(2.162)	(-0.373)						0.154	
	-0.261	-1.016*	-0.099	0.849**	-0.132	-0.021	-0.041	0.064	-0.310*	0.113		
	(-1.465)	(-1.768)	(-0.237)	(2.076)	(-0.549)	(-0.135)	(-0.313)	(0.420)	(-1.843)		0.161	0.238
3-m t-bill	0.245	-0.689***	-0.156	0.262	0.376*					0.061		
	(1.520)	(-2.968)	(-0.817)	(1.409)	(1.758)						0.001	
	0.293*	-0.634***	-0.221	0.290*	0.354	-0.175	0.019	0.164*	-0.132	0.060		
	(1.788)	(-2.837)	(-1.100)	(1.693)	(1.633)	(-1.535)	(0.171)	(1.850)	(-1.587)		0.001	0.225
6-m t-bill	0.268*	-0.846***	0.009	0.250	0.321					0.083		
	(1.705)	(-3.007)	(0.043)	(1.392)	(1.576)						0.002	
	0.319**	-0.792***	-0.051	0.276	0.292	-0.181	-0.002	0.164**	-0.138*	0.087		
	(2.008)	(-2.951)	(-0.247)	(1.635)	(1.412)	(-1.609)	(-0.025)	(2.081)	(-1.806)		0.002	0.148
1-yr t-bond	0.256	-0.946***	0.179	0.280	0.244					0.079		
	(1.488)	(-3.173)	(0.819)	(1.596)	(1.110)						0.009	
	0.304*	-0.897***	0.128	0.302*	0.212	-0.156	-0.023	0.150*	-0.145*	0.081		
	(1.751)	(-3.106)	(0.568)	(1.840)	(0.947)	(-1.280)	(-0.206)	(1.747)	(-1.712)		0.009	0.232
5-yr t-bond	-0.132	-1.000**	0.485	0.738**	-0.056					0.056		
	(-0.592)	(-2.386)	(1.402)	(2.166)	(-0.197)						0.127	
	-0.113	-1.074**	0.525	0.780**	-0.077	0.106	-0.114	0.129	-0.122	0.051		
	(-0.513)	(-2.584)	(1.357)	(2.313)	(-0.272)	(0.634)	(-0.831)	(1.241)	(-0.938)		0.092	0.452

This table reports the results of predictive regressions of 6-month growth (or change) in money, credit and yield variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Table C.9: Predicting money, credit and yield variables for 12-month horizon - Hodrick covariance matrix

	TERM	$FV^{(1)}$	$FV^{(2)}$	$FV^{(3)}$	$FV^{(4)}$	$FSC^{(1)}$	$FSC^{(2)}$	$FSC^{(3)}$	$FSC^{(4)}$	Adj. R^2	Joint FV	Joint FSC
M1	2.702***	-0.353	-0.212	2.182**	-0.406					0.339		
	(3.884)	(-0.272)	(-0.245)	(2.380)	(-0.468)						0.068	
	2.581***	0.682	-0.870	1.701**	-0.313	-0.156	0.765*	-1.055**	-0.588	0.369		
	(3.471)	(0.545)	(-0.823)	(2.044)	(-0.368)	(-0.288)	(1.907)	(-2.489)	(-1.252)		0.182	0.002
M2 (real)	-0.369	0.783	0.143	-0.320	-0.458					0.034		
	(-1.079)	(1.354)	(0.324)	(-0.873)	(-0.958)						0.684	
	-0.394	1.303**	-0.230	-0.442	-0.391	-0.471*	0.783***	-0.198	-0.048	0.066		
	(-1.109)	(2.224)	(-0.450)	(-1.337)	(-0.827)	(-1.760)	(4.025)	(-0.902)	(-0.242)		0.179	0.002
Reserves tot	5.158*	9.496	-6.415	16.364**	-10.484					0.010		
	(1.675)	(0.895)	(-0.867)	(2.203)	(-1.374)						0.101	
	1.690	14.292	-7.062	11.773*	-8.838	3.284	8.651**	-15.066**	7.187	0.035		
	(0.490)	(1.269)	(-0.958)	(1.750)	(-1.256)	(0.967)	(1.997)	(-2.032)	(1.522)		0.111	0.271
C&I loans	-2.995**	-7.845***	-0.001	1.454	1.536					0.260		
	(-2.280)	(-4.864)	(0.000)	(0.984)	(0.870)						0.000	
	-2.195	-5.667***	-1.770	0.920	1.036	-3.485***	-1.068*	0.370	-2.977***	0.482		
	(-1.629)	(-3.861)	(-0.988)	(0.640)	(0.582)	(-4.172)	(-1.750)	(0.590)	(-4.697)		0.000	0.000
CPI	-0.203	-0.556*	-0.207	0.517**	-0.010					0.095		
	(-1.170)	(-1.880)	(-0.783)	(2.577)	(-0.039)						0.007	
	-0.144	-0.657**	-0.163	0.547***	-0.063	0.077	-0.405***	0.135	-0.237**	0.222		
	(-0.823)	(-2.098)	(-0.605)	(2.671)	(-0.255)	(0.629)	(-3.312)	(0.851)	(-2.221)		0.006	0.005
3-m t-bill	0.449***	-0.675***	0.050	0.164	0.293*					0.138		
	(2.675)	(-3.879)	(0.313)	(1.302)	(1.817)						0.000	
	0.499***	-0.694***	0.025	0.215*	0.271	-0.095	-0.076	0.193**	-0.119***	0.142		
	(2.928)	(-4.207)	(0.165)	(1.933)	(1.640)	(-0.874)	(-0.776)	(2.287)	(-2.734)		0.000	0.036
6-m t-bill	0.452***	-0.761***	0.149	0.167	0.246					0.154		
	(2.778)	(-3.930)	(0.923)	(1.385)	(1.530)						0.000	
	0.503***	-0.774***	0.118	0.215**	0.225	-0.092	-0.080	0.190**	-0.134***	0.161		
	(3.075)	(-4.149)	(0.766)	(1.994)	(1.367)	(-0.860)	(-0.843)	(2.434)	(-3.125)		0.000	0.017
1-yr t-bond	0.421**	-0.817***	0.275	0.185	0.177					0.147		
	(2.442)	(-4.039)	(1.540)	(1.446)	(1.050)						0.000	
	0.466***	-0.831***	0.247	0.227**	0.158	-0.058	-0.081	0.169**	-0.140***	0.152		
	(2.705)	(-4.185)	(1.395)	(2.031)	(0.919)	(-0.498)	(-0.791)	(1.983)	(-2.998)		0.000	0.029
5-yr t-bond	0.007	-0.625**	0.486*	0.330	-0.088					0.045		
	(0.032)	(-2.424)	(1.761)	(1.433)	(-0.333)						0.024	
	0.010	-0.686***	0.518*	0.353*	-0.090	0.097	-0.075	0.049	-0.041	0.033		
	(0.048)	(-2.867)	(1.662)	(1.668)	(-0.341)	(0.675)	(-0.657)	(0.502)	(-0.457)		0.013	0.852

This table reports the results of predictive regressions of 12-month growth (or change) in money, credit and yield variables. Details about the variables can be found in Table C.1 of Appendix C. The sample period is 1996:01-2012:12. For each variable the first two rows present the results of a simple model containing only term spread and forward variances as predictors. The next rows present the results of a model augmented with forward skewness coefficients. A constant term is always included into the model but is omitted for brevity. Significance tests are based on the covariance matrix suggested by Hodrick (1992) with lag length equal to the forecasting horizon. Individual coefficient t-statistics can be found in parentheses. ***, ** and * denote significance at 1%, 5% and 10% level respectively. P-values of Wald tests of joint significance for forward variances and forward skewness coefficients are reported at the last two columns.

Appendix D

Appendix to Chapter 5

D.1 Additional Results

In this section we provide some additional results using the consumption-wealth ratio of Lettau and Ludvigson (2001) and the stock market illiquidity of Amihud (2002). Consumption-wealth ratio data are obtained from Sydney Ludvigson's website[1] and a monthly time-series is created from the most recent quarterly observations. Stock market illiquidity is created by averaging the illiquidity measures of all the NYSE/AMEX stocks within a given month. Both variables are insignificant across all horizons in univariate regressions and have very limited impact on the significance of our dispersion in expectations measures when added in the predictive model.

[1] `http://www.econ.nyu.edu/user/ludvigsons/`

Table D.1: 1-month horizon predictability - additional variables

Panel A: Univariate			**Panel B: Bivariate**					
		$\widetilde{\mathbf{R}}^2$ **(%)**	**DISP**	**Z**	$\widetilde{\mathbf{R}}^2$ **(%)**	**DISP***	**Z**	$\widetilde{\mathbf{R}}^2$ **(%)**
DISP	-9.68	2.23						
	(-2.36)**							
	[-2.44]**							
DISP*	-9.20	1.97						
	(-2.29)**							
	[-2.35]**							
VRP	16.09	7.04	-9.27	15.85	9.10	-8.14	15.53	8.51
	(4.66)***		(-2.47)**	(4.34)***		(-2.24)**	(4.28)***	
	[2.49]**		[-2.35]**	[2.46]**		[-2.11]**	[2.41]**	
CAY	3.07	-0.22	-9.70	-0.05	1.75	-9.38	-0.47	1.49
	(0.79)		(-2.23)**	(-0.01)		(-2.15)**	(-0.11)	
	[0.81]		[-2.33]**	[-0.01]		[-2.23]**	[-0.12]	
ILLIQ	2.52	-0.31	-9.54	0.75	1.76	-9.11	0.42	1.48
	(0.63)		(-2.27)**	(0.18)		(-2.18)**	(0.10)	
	[0.64]		[-2.34]**	[0.19]		[-2.24]**	[0.10]	
Panel C: Trivariate								
	DISP	**VRP**	**Z**	$\widetilde{\mathbf{R}}^2$ **(%)**	**DISP***	**VRP**	**Z**	$\widetilde{\mathbf{R}}^2$ **(%)**
CAY	-10.29	16.40	-3.21	8.90	-9.38	16.06	-3.37	8.33
	(-2.58)**	(4.32)***	(-0.80)		(-2.37)**	(4.28)***	(-0.81)	
	(-2.44)**	(2.46)**	(-0.75)		[-2.23]**	[2.42]**	[-0.77]	
ILLIQ	-9.64	16.18	-2.00	8.75	-8.61	15.86	-2.15	8.17
	(-2.50)**	(4.26)***	(-0.49)		(-2.29)**	(4.22)***	(-0.52)	
	[-2.36]**	[2.43]**	[-0.47]		[-2.13]**	[2.39]**	[-0.50]	

This table reports the results of 1-month ahead predictive regressions for the excess return on the CRSP value-weighted index. The sample period is 1996:01-2012:12. Panel A reports the results of univariate regressions, Panel B the results of bivariate regressions and Panel C the results of trivariate regressions. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), consumption-wealth ratio (CAY) and stock market illiquidity (ILLIQ). Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance in 1%, 5% and 10% level.

Table D.2: Univariate long-horizon predictability - additional variables

	h=3		**h=6**		**h=12**		**h=24**	
		$\widetilde{\mathbf{R}}^2$ **(%)**		$\widetilde{\mathbf{R}}^2$ **(%)**		$\widetilde{\mathbf{R}}^2$ **(%)**		$\widetilde{\mathbf{R}}^2$ **(%)**
DISP	-7.39	3.61	-5.47	3.54	-6.67	10.89	-5.19	12.72
	(-2.59)**		(-2.25)**		(-2.90)***		(-2.35)**	
	[-2.33]**		[-1.83]*		[-2.38]**		[-2.02]**	
DISP*	-7.10	3.30	-5.10	3.00	-6.65	10.80	-5.20	12.74
	(-2.65)***		(-2.24)**		(-3.03)***		(-2.49)**	
	[-2.31]**		[-1.81]*		[-2.40]**		[-2.11]**	
VRP	13.05	12.32	8.34	8.90	3.94	3.46	2.90	3.58
	(5.25)***		(3.96)***		(2.40)**		(2.29)**	
	[3.62]***		[3.33]***		[2.06]**		[1.77]*	
CAY	3.41	0.37	3.17	0.85	3.24	2.17	3.52	5.56
	(1.01)		(0.96)		(0.91)		(1.11)	
	[0.90]		[0.83]		[0.84]		[0.93]	
ILLIQ	2.57	-0.01	1.46	-0.22	0.95	-0.29	0.07	-0.56
	(0.70)		(0.41)		(0.25)		(0.02)	
	[0.65]		[0.37]		[0.24]		[0.02]	

This table reports the results of 3-, 6-, 12- and 24-month ahead univariate predictive regressions for the excess return on the CRSP value-weighted index. The sample period is 1996:01-2012:12. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), consumption-wealth ratio (CAY) and stock market illiquidity (ILLIQ). Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance in 1%, 5% and 10% level.

Table D.3: Trivariate long-horizon predictability - additional variables

	DISP	**VRP**	**Z**	$\widetilde{R}^2$ **(%)**	**DISP***	**VRP**	**Z**	$\widetilde{R}^2$ **(%)**
				Panel A: 3-month horizon				
CAY	-7.47	13.08	-1.33	15.33	-6.76	12.84	-1.43	14.50
	(-3.15)***	(4.43)***	(-0.41)		(-2.94)***	(4.43)***	(-0.42)	
	[-2.27]**	[3.60]***	[-0.34]		[-2.09]**	[3.55]***	[-0.35]	
ILLIQ	-7.21	13.01	-0.93	15.27	-6.46	12.77	-1.04	14.45
	(-3.15)***	(4.46)***	(-0.28)		(-2.95)***	(4.46)***	(-0.30)	
	[-2.24]**	[3.58]***	[-0.23]		[-2.06]**	[3.53]***	[-0.25]	
				Panel B: 6-month horizon				
CAY	-5.25	8.20	0.04	11.74	-4.53	8.03	0.05	10.78
	(-2.19)**	(3.24)***	(0.01)		(-1.88)*	(3.25)***	(0.02)	
	[-1.69]*	[3.32]***	[0.01]		[-1.54]	[3.28]***	[0.01]	
ILLIQ	-5.42	8.35	-0.89	11.84	-4.75	8.18	-0.95	10.89
	(-2.44)**	(3.27)***	(-0.27)		(-2.17)**	(3.26)***	(-0.27)	
	[-1.77]*	[3.41]***	[-0.22]		[-1.64]	[3.37]***	[-0.23]	
				Panel C: 12-month horizon				
CAY	-6.41	3.71	0.63	13.81	-6.30	3.48	0.33	13.09
	(-2.51)**	(1.96)*	(0.19)		(-2.46)**	(1.87)*	(0.09)	
	[-2.24]**	[2.09]**	[0.16]		[-2.23]**	[1.98]**	[0.08]	
ILLIQ	-6.68	3.92	-0.60	13.81	-6.55	3.65	-0.81	13.23
	(-2.90)***	(2.02)**	(-0.19)		(-2.90)***	(1.91)*	(-0.25)	
	[-2.35]**	[2.19]**	[-0.15]		[-2.34]**	[2.07]**	[-0.20]	
				Panel D: 24-month horizon				
CAY	-4.61	2.45	1.77	16.88	-4.48	2.26	1.65	16.07
	(-1.63)	(1.75)*	(0.53)		(-1.55)	(1.65)*	(0.47)	
	[-1.69]*	[2.02]**	[0.46]		[-1.66]*	[1.84]*	[0.42]	
ILLIQ	-5.19	2.89	-0.90	15.87	-5.11	2.67	-0.99	15.35
	(-2.24)**	(1.91)*	(-0.27)		(-2.23)**	(1.78)*	(-0.29)	
	[-1.98]**	[2.32]**	[-0.23]		[-2.02]**	[2.16]**	[-0.25]	

This table reports the results of 3- (Panel A), 6- (Panel B), 12- (Panel C) and 24-month (Panel D) ahead trivariate predictive regressions for the excess return on the CRSP value-weighted index. The sample period is 1996:01-2012:12. The forecasting variables are the two dispersion in options traders' expectations measures (DISP, DISP*), variance risk premium (VRP), consumption-wealth ratio (CAY) and stock market illiquidity (ILLIQ). Reported coefficients indicate the percentage annualized excess return resulting from a one standard deviation increase in each predictor variable. Newey and West (1987) and Hodrick (1992) t-statistics with lag length equal to the forecasting horizon are reported in parentheses and square brackets respectively. ***, ** and * denote significance in 1%, 5% and 10% level.