TI 2017-002/IV
Tinbergen Institute Discussion Paper

Implied Volatility Sentiment: A Tale of Two Tails

Luiz Félix[1]
Roman Kräussl[2]
Philip Stork[3]

[1] VU University Amsterdam, APG Asset Management
[2] Luxembourg School of Finance, University of Luxembourg
[3] VU University Amsterdam, FEWEB, School of Finance
and Risk Management and Tinbergen Institute Amsterdam

Implied Volatility Sentiment: A Tale of Two Tails*

Luiz Félix
VU University Amsterdam
APG Asset Management

Roman Kräussl
Luxembourg School of Finance,
University of Luxembourg

Philip Stork†
VU University Amsterdam
Tinbergen Institute Amsterdam

January 2017

ABSTRACT

Low probability events are overweighted in the pricing of out-of-the-money index puts and single stock calls. This behavioral bias is strongly time-varying, and is linked to equity market sentiment and higher moments of the risk-neutral density. We find that our implied volatility (IV) sentiment measure, jointly derived from index and single stock options, explains investors' overweight of tail events well. When employed within a trading strategy, our IV-sentiment measure delivers economically significant results, which are more consistent than the ones produced by the market sentiment factor. Out-of-sample tests on reversal prediction show that our IV-sentiment measure adds value over and above traditional factors in the equity risk premium literature.

Keywords: Sentiment, implied volatility skew, equity-risk premium, reversals, predictability.

JEL classification: G12, G14, G17.

---

*We thank seminar participants at APG Asset Management Quant Roundtable for their helpful comments. We thank APG Asset Management for financial support and for making available part of the data.

†Philip Stork (corresponding author) is Professor at VU University Amsterdam, FEWEB, School of Finance and Risk Management, de Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Tel.: + 31 20 59 83651. E-mail address: p.a.stork@vu.nl.

1 Introduction

End-users of out-of-sample (OTM) options overweight small probability events, i.e., tail events. This behavioral bias, suggested by Tversky's and Kahneman's (1992) cumulative prospect theory (CPT), is claimed to be present in the pricing of OTM index puts and in OTM single stock calls (Barberis and Huang, 2008; Polkovnichenko and Zhao, 2013)[1]. Within the index option market, the typical end-users of OTM puts are institutional investors, who use them to protect their large equity portfolios. Because institutional investors have large portfolios and hold a substantial part of the total market capitalization, OTM index puts are frequently in high demand and, as a result, overvalued. The reason for such richness of OTM puts goes back to the 1987 financial market crash. Bates (1991) and Jackwerth and Rubinstein (1996) argue that the implied distribution of equity market expected returns from index options changed considerably since the 1987 market crash. Their findings demonstrate that, since the crash, a large shift in market participants' demand for such instruments took place, evidenced by the probabilities implied by options prices. Before the crash, the probability of large negative stock returns was close to the one suggested by a normal distribution. In contrast, just prior to the 1987 crash, the probability of large negative returns implied by option prices rose considerably. Such increased demand for hedging against tail risk events suggested a change in beliefs and attitude towards risk. Investors feared another crash and became more willing to give up upside potential in equities to hedge against the risk of drawdowns via put options. Bates (2003) suggest that even models adjusted for stochastic volatility, stochastic interest rates, and random jumps do not fully explain the high level of OTM puts' implied volatilities (IV). Accordingly, Garleanu et al. (2009) argue that excessive IV from OTM puts cannot either be explained by option-pricing models that take such institutional investors' demand pressure into account.

The literature also claims that OTM calls on single stocks are systematically expensive (Barberis and Huang, 2008; Boyer and Vorkink, 2014). The typical end-users of OTM single stock calls are individual investors. Bollen and Whaley (2004) state that changes in the IV structure of single stock options across moneyness are driven by the net purchase of calls by individual investors. The literature provides several explanations for such strong buying pressure of calls by retail investors. For example, Mitton and Vorkink (2007) and Barberis and Huang (2008) propose models in which investors have a clear preference for positive return skewness, or "lottery ticket" type of assets. In consequence of this preference, retail investors overpay for these leveraged securities, making OTM calls expensive and causing them to yield low forward returns. Cornell (2009) presents a behavioral explanation for the overpricing of

[1] We acknowledge that it is yet unclear whether the overweighting of small probabilities is a behavioral bias (i.e., a bias in beliefs) or caused solely by preferences. Barberis (2013) eloquently discusses how both phenomena are distinctly different and how both (individually or jointly) may potentially explain the existence of overpriced OTM options, as well as many other puzzling facts in financial markets. In this paper we take a myopic view and use only the first explanation, the existence of a behavioral bias, for ease of exposition.

single stock calls: because investors are overconfident in their stock-picking skills, they buy calls to get the most "bang for the buck". A connected explanation for the structural overpricing of single stock calls is leverage aversion or leverage constraint: because investors are averse to borrowing (levering) or constrained to do so, they buy instruments with implicit leverage to achieve their return targets.

Beyond this literature that supports the link between institutional and individual investor trading activity and the structural overvaluation of OTM options, we argue that short-term trading dynamics also influence the pricing of OTM options. For instance, Han (2008) provides evidence that the index options IV smirk is steeper when professional investors are bearish. He concludes that the steepness of the IV structure across moneyness relates to investors' sentiment. In the same line, Amin et al. (2004) argue that investors bid up the prices of put options after increases in stock market volatility and rising risk aversion, whereas such buying pressure wanes following positive momentum in equity markets. Mahani and Poteshman (2008) argue that trading in single stock call options around earnings announcements is speculative in nature and dominated by unsophisticated retail investors. Lakonishok et al. (2007) show evidence that long call prices increased substantially during bubble times (1990 and 2000) and that most of the single stock options' market activity consists of speculative directional call positions. Lemmon and Ni (2011) discuss that the demand for single stock options (dominated by speculative individual investors' trades) positively relates to sentiment. Lastly, Polkovnichenko and Zhao (2013) suggest that time-variation in overweight of small probabilities derived from index put options might depend on sentiment, whereas Felix et al. (2016) provide evidence that the time-varying overweight of small probabilities from single stock options largely links to sentiment.

The above studies suggest that OTM index puts and single stock calls are systematically overpriced and that the valuation misalignments fluctuate considerably over time, caused by changes in investor sentiment. In this paper, we delve deeper into it and investigate how overweight of small probabilities links to sentiment and forward returns.

The first contribution of this paper is to evaluate the information content of overweighted small probabilities from index puts and single stock calls, as a measure of sentiment. We assess the ability of this measure to predict forward equity returns and, more specifically, equity market reversals, defined as abrupt changes in the market direction[2]. Because we find overweight small probabilities to be strongly linked to IV skews, we hypothesize that reversals may follow not only periods of excessive overweight of tails but also periods of extreme IV skews[3].

[2]Reversals in the context of this paper are not to be confused with the, so-called, reversal (cross-sectional) strategy, i.e., a strategy that buys (sells) stocks with low (high) total returns over the past month, as first documented by Lehmann (1990). We focus on the overall equity market, rather than investigating single stocks.

[3]The literature on IV skew has largely explored the level of volatility skew across stocks and their cross-section of returns. However, insights on the link between the skew and the overall stock market are still incipient. The study by Doran et al. (2007) is one of the few that has tested the power of implied volatility skews as a predictor of aggregate market returns. However, they only analyze the relation between skews and one-day ahead returns (found to be weakly negatively related), and ignore any longer and perhaps more persistent effects. Similarly, several studies have already attempted to recognize the conditionality of forward equity market returns to other volatility-type of measures: Ang and Liu (2007) for realized variance, in Bliss and Panigirtzoglou (2004) for risk-aversion implied by risk-neutral probability distribution function embedded in cross-sections of options, Bollerslev et al. (2009) for variance risk premium, Driessen et al. (2013) for option-

One characteristic of the literature that analyzes the informational content of IV skews is that it evaluates index puts' IV skews and single stock calls' IV skews completely separated from each other. Our second contribution is that we are, to the best of our knowledge, the first in the literature to use IV skews jointly extracted from both the index and single stock option market as an indicator for investors' sentiment. Our sentiment measure, the so-called *IV-sentiment*, is calculated as the IV of OTM index puts minus the IV of OTM single stock calls. We conjecture that our *IV-sentiment* measure is an advance on understanding investors' sentiment because it captures the very distinct nature of these markets' two main categories of end-users: 1) IV from OTM puts captures institutional investors' willingness to pay for leverage to hedge their downside risk (portfolio insurance), as a measure of bearishness, whereas 2) IV from OTM single stock calls captures levering by individual investors for speculation on the upside ("lottery tickets" buying), as a measure of bullishness. Thus, a high level of *IV-sentiment* indicates bearish sentiment, as IV from index puts outpace the ones from single stock calls. In contrast, low levels of *IV-sentiment* indicate bullishness sentiment, as IV from single stock calls become high relative to the ones from index puts.

We find that our *IV-sentiment* measure predicts equity market reversals better than overweight of small probabilities itself and that it delivers positive risk-adjusted returns more consistently than the common Baker and Wurgler (2007) sentiment factor when evaluated via two trading strategies, a high-frequency and a low-frequency one. In univariate and multivariate predictive regression settings, our *IV-sentiment* measure improves the out-of-sample forecast ability of traditional equity risk-premium models. This result is partially due to the low correlation between the payoff of our *IV-sentiment* measure and the payoffs of traditional factors. Thus, the third contribution of our paper is to complement the literature on out-of-sample forecasting of the equity risk-premium and equity-market timing (Welch and Goyal, 2008; Campbell and Thompson, 2008; Rapach et al., 2010).

A final contribution of our work is to investigate the ability of our *IV-sentiment* measure to improve time-series momentum and cross-sectional momentum strategies. Our sentiment measure is uncorrelated to these strategies, especially at the tails, i.e., when momentum crashes. Consequently, we document an increase in the informational content of such momentum strategies when combined with the *IV-sentiment* strategy.

The remainder of this paper is organized as follows. Section 2 describes the data and the main methods employed in our empirical study. In Section 3, within three sub-sections, we focus on estimating overweight of small probabilities parameters from the index and single stock option markets as well as linking it to the Baker and Wurgler (2007) sentiment factor and other proxies for sentiment. In Section 4 we test how our sentiment proxy based on overweight of small probabilities relates to forward equity returns. Section 5 concludes.

implied correlations, Pollet and Wilson (2008) for historical correlations and Vilkov and Xiao (2013) for the risk-neutral tail loss measure. Most of these studies document a short-term negative relation between risk measures and equity market movements.

2 Data and Methodology

We use S&P 500 index options' IV data and single stock weighted average IV data from the largest 100 stocks of the S&P 500 index within our risk-neutral density (RND) estimations. The IV data comes from closing mid-option prices from January 2, 1998 to March 19, 2013 for fixed maturities for five moneyness levels, i.e., 80, 90, 100, 110, and 120, at the three-, six- and twelve-month maturity both for index and single stock options. Eq. (A.1k) in appendix A shows how weighted average single stock IV are computed. Weights applied are the S&P 500 index weights normalized by the sum of weights of stocks for which IVs across all moneyness levels are available. Following the S&P 500 index methodology and the unavailability of IV information for every stock in all days in our sample, stocks weights in this basket change on a daily basis. The sum of weights is, on average, 58 percent of the total S&P 500 index capitalization and it fluctuates from 46 to 65 percent. Continuously compounded stock market returns are calculated throughout our analysis from the basket of stocks weighted with the same daily-varying loadings used for aggregating the IV data[4]. For index options, we use the S&P 500 index prices to calculate continuously compounded stock market returns. Realized index returns and single stock returns are downloaded via Bloomberg.

Overweight of small probabilities is embedded in the cumulative prospect theory (CPT) model by means of the weighting function of the probability of prospects. Within the CPT model, overweight of small probabilities is measured by the probability weighting function parameters δ and γ for the left (losses) and right (gains) side of the return distribution, respectively. δ and $\gamma < 1$ imply overweight of small probabilities, whereas δ and $\gamma > 1$ imply underweight of small probabilities, and δ and γ equal to 1 means neutral weighting of prospects (see Tversky and Kahneman, 1992).

Our methodology builds on the assumption that investors' subjective density estimates should correspond, on average[5], to the distribution of realizations (Bliss and Panigirtzoglou, 2004). Thus, estimating CPT probability weighting function parameters δ and γ is only feasible if two basic inputs are available: the CPT subjective density function and the distribution of realizations, i.e. the empirical density function (EDF). The methodology applied by us to estimate these two parameters comprises of: 1) estimating the returns' risk-neutral density from option prices using a modified Figlewski (2010) method; 2) estimating the partial CPT density function using the CPT marginal utility function; 3) "undoing" the effect of the probability weighting function (w) to obtain the CPT subjective density function; 4) simulating time-varying empirical return distributions using the Rosenberg and Engle (2002) approach; and 5)

[4]We thank Barclays Capital for providing the implied volatility data. Barclays Capital disclosure: "*Any analysis that utilizes any data of Barclays, including all opinions and/or hypotheses therein, is solely the opinion of the author and not of Barclays. Barclays has not sponsored, approved or otherwise been involved in the making or preparation of this Report, nor in any analysis or conclusions presented herein. Any use of any data of Barclays used herein is pursuant to a license.*"

[5]This assumption implies that investors are somewhat rational, which is not inconsistent with the CPT-assumption that the representative agent is less than fully rational. The CPT suggests that investors are biased, not that decision makers are utterly irrational to the point that their subjective density forecast should not correspond, on average, to the realized return distribution.

minimizing the squared difference of the tail probabilities of the CPT and EDF to obtain daily optimal δ's and γ's.

Our starting point for obtaining the CPT probability weighting function parameters δ and γ is the estimation of RND from IV data. In order to estimate the RND, we first apply the Black-Scholes model to our IV data to obtain options prices (C) for the S&P 500 index. Once our data is normalized, so strikes are expressed in terms of percentage moneyness, the instantaneous price level of the S&P 500 index (S_0) equals 100 for every period for which we would like to obtain implied returns. Contemporaneous dividend yields for the S&P 500 index are used for the calculation of P as well as the risk-free rate from three-, six-, and twelve-month T-bills. Because we have IV data for five levels of moneyness, we implement a modified Figlewski (2010) method for extracting the RND structure. The main advantage of the Figlewski (2010) method over other techniques is that it extracts the body and tails of the distribution separately, thereby allowing for fat tails.

Once the RND is estimated, we must change the measure to translate it into the subjective density function, a *real-world* probability distribution. This operation is possible via the pricing kernel as follows:

$$\frac{f_P(S_T)}{f_Q(S_T)} = \Lambda \frac{U'(S_T)}{U'(S_t)} \equiv \varsigma(S_T), \tag{1}$$

where, $f_Q(S_T)$ is the RND, $f_P(S_T)$ is the *real-world* probability distribution, S_T is wealth or consumption, $\varsigma(S_T)$ is the pricing kernel, Λ is the subjective discount factor (the time-preference constant) and $U(\cdot)$ is the representative investor utility function.

Since CPT-biased investors price options as if the data-generating process has a cumulative distribution $F_{\widetilde{P}}(S_T) = w(F_P(S_T))$, where w is the weigthing function, its density function becomes $f_{\widetilde{P}}(S_T) = w'(F_P(S_T)) \cdot f_P(S_T)$ (Dierkes, 2009; Polkovnichenko and Zhao, 2013) and Eq. (1) collapses into Eq. (2):

$$\frac{w'(F_P(S_T)) \cdot f_P(S_T)}{f_Q(S_T)} = \varsigma(S_T) \tag{2}$$

which, re-arranged into Eq. (4) via Eqs. (3a) and (3b), demonstrates that for the CPT to hold, the subjective density function should be consistent with the probability weighted EDF:

$$\underbrace{f_Q(S_T)}_{\text{RND}} = \underbrace{w'(F_P(S_T))}_{\text{probability weighing}} \cdot \underbrace{f_P(S_T)}_{\text{EDF}} \cdot \underbrace{\varsigma(S_T)}_{\text{pricing kernel}} \tag{3a}$$

$$\underbrace{f_Q(S_T)}_{\text{RND}} = \underbrace{f_{\widetilde{P}}(S_T)}_{\text{probability weighted EDF}} \cdot \underbrace{\varsigma(S_T)}_{\text{pricing kernel}} \tag{3b}$$

$$\underbrace{\frac{f_Q(S_T)}{\lambda \frac{U'(S_T)}{U'(S_t)}} = \frac{f_Q(S_T)}{\varsigma(S_T)}}_{\text{Subjective density}} = \underbrace{f_{\widetilde{P}}(S_T)}_{\text{probability weighted EDF}} \tag{4}$$

Following Bliss and Panigirtzoglou (2004), Eq. (4) can be manipulated so that the time-preference constant Λ of the pricing kernel vanishes, producing Eq. (5), which directly relates the probability weighted EDF, the RND, and the marginal utility, $U'(S_T)$:

$$\underbrace{f_{\widetilde{P}}(S_T)}_{\text{probability weighted EDF}} = \underbrace{\frac{\lambda\frac{U'(S_T)}{U'(S_t)}Q(S_T)}{\int\frac{U'(S_t)}{U'(x)}Q(x)dx} = \frac{\frac{f_Q(S_T)}{U'(S_T)}}{\int\frac{f_Q(x)}{U'(x)}dx}}_{\text{Generic subjective density function}} \tag{5}$$

where $\int\frac{Q(x)}{U'(x)}dx$ normalizes the resulting subjective density function to integrate to one. Once the utility function is estimated, Eq. (5) allows us to convert RND into the probability weighted EDF. As the CPT marginal utility function is $U'(S_T) = \upsilon'(S_T)$, and, thus, $\upsilon'(S_T) = \alpha S_T^{\alpha-1}$ for $S_T >= 0$, and $\upsilon'(S_T) = -\lambda\beta(-S_T)^{\beta-1}$ for $S_T < 0$, we obtain Eq. (6) and (7):

$$f_{\widetilde{P}}(S_T) = \frac{\frac{f_Q(S_T)}{\alpha S_T^{\alpha-1}}}{\int\frac{f_Q(x)}{\alpha x^{\alpha-1}}dx} \quad for \quad S_T \geq 0, \quad and \tag{6}$$

$$\underbrace{f_{\widetilde{P}}(S_T)}_{\text{probability weighted EDF}} = \underbrace{\frac{\frac{f_Q(S_T)}{-\lambda\beta(-S_T)^{\beta-1}}}{\int\frac{f_Q(x)}{-\lambda\beta(-x)^{\beta-1}}dx}}_{\text{Partial CPT density function}} \quad for \quad S_T < 0, \quad and \tag{7}$$

Eqs. (6) and (7) relates on the LHS of the EDF, where probabilities are weighted according to the CPT probability distortion functions, to the subjective density function on the RHS, which is derived from the CPT value function, separately for gains and losses, what we call the partial CPT density function (PCPT). As the function $w(F_P(S_T))$ is strictly increasing over the domain [0,1], there is a one-to-one relationship between $w(F_P(S_T))$ and a unique inverse $w^{-1}(F_P(S_T))$. So, result $f_{\widetilde{P}}(S_T) = w'(F_P(S_T))f_P(S_T)$ also implies $f_{\widetilde{P}}(S_T).(w^{-1})'(F_P(S_T)) = f_P(S_T)$. This outcome allows us to directly relate the original EDF to the CPT subjective density function, by "undoing" the effect of the CPT probability distortion functions within the PCPT density function:

$$\underbrace{f_P(S_T)}_{\text{EDF}} = \underbrace{\frac{\frac{f_Q(S_T)}{\nu'(S_T)}}{\int\frac{f_Q(x)}{\nu'(x)}dx}(w^{-1})'(F_P(S_T))}_{\text{CPT density function}} \tag{8}$$

Thus, once the relation between the probability weighting function of EDF and the PCPT density is established, as in Eqs. (6) and (7), one can eliminate the weighting scheme affecting returns by applying the inverse of such weightings to the subjective density function without endangering such equalities, as in Eq. (8).

As the RND is converted into the subjective density function, we must also estimate daily empirical density functions (EDF). We built such time-varying EDFs from an invariant component, the *standardized innovation density*, and a time-varying part, the lagged conditional variance ($\sigma^2_{t|t-1}$) produced by an EGARCH model (Nelson, 1991). We first define the *standardized innovation*, being the ratio of empirical returns and their conditional standard deviation ($ln(S_t/S_{t-1})/\sigma_{t|t-1}$) produced by the EGARCH model. From the set of standardized innovations produced, we can then estimate a density shape, i.e., the *standardized innovation density.* The advantage of such a density shape versus a parametric one is that it may include the typically observed fat tails and negative skewness, which are not incorporated in simple parametric models, e.g., the normal distribution. This density shape is invariant and it is turned time-varying by multiplication of each standardized innovation by the EGARCH conditional standard deviation at time t, which is specified as follows:

$$ln(S_t/S_{t-1}) = \mu + \epsilon_t, \epsilon \sim f(0, \sigma^2_{t|t-1}) \tag{9a}$$

and

$$\sigma^2_{t|t-1} = \omega_1 + \alpha\epsilon^2_{t-1} + \beta\sigma^2_{t-1|t-2} + \vartheta Max[0, -\epsilon_{t-1}]^2, \tag{9b}$$

where α captures the sensitivity of conditional variance to lagged squared innovations (${\epsilon^2}_{t-1}$), β captures the sensitivity of conditional variance to the conditional variance (${\sigma^2}_{t-1|t-2}$), and ϑ allows for the asymmetric impact of lagged returns ($\vartheta Max[0, -\epsilon_{t-1}]^2$). The model is estimated using maximum log-likelihood where innovations are assumed to be normally distributed.

Up to now, we produced a one-day horizon EDF for every day in our sample but we still lack time-varying EDFs for the three-, six-, and twelve-month horizons. Thus, we use bootstrapping to draw 1,000 paths towards these desired horizons by randomly selecting single innovations (ϵ_{t+1}) from the one-day horizon EDFs available for each day in our sample. We note that once the first return is drawn, the conditional variance is updated ($\sigma^2_{t-1|t-2}$) affecting the subsequent innovation drawings of a path. This sequential exercise continues through time until the desired horizon is reached. To account for drift in the simulated paths, we add the daily drift estimated from the long-term EDF to drawn innovations, so that the one-period simulated returns equal $\epsilon_{t+1} + \mu$. The density functions produced by the collection of returns implied by the terminal values of every path and their starting points are our three-, six-, and twelve-month EDFs. These simulated paths contain, respectively, 63, 126, and 252 daily returns. We note that by drawing returns from stylized distributions with fat-tails and excess skewness, our EDFs for the three relevant horizons also embed such features. This estimation method for time-varying EDF is based on Rosenberg and Engle (2002).

Finally, once these three time-varying EDFs are estimated for all days in our sample, we estimate δ and γ for each of these days using Eq. (10) and (11).

$$w^+(\gamma, \delta = \gamma) = Min \sum_{b=1}^{B} W_b(EDF^b_{prob} - CPT^b_{prob})^2, \tag{10}$$

$$w^-(\delta, \delta = \gamma) = Min \sum_{b=1}^{B} W_b(EDF^b_{prob} - CPT^b_{prob})^2, \tag{11}$$

where, EDF^b_{prob} and CPT^b_{prob} are the probability within bin b in the empirical and CPT density functions and W_b are weights given by $\frac{1}{\frac{1}{\sqrt{2}}} \int_{0.5}^{\infty} e^{\frac{-x^2}{2}} dx = 1$, the reciprocal of the normalized normal probability distribution (above its median), split in the same total number of bins (B) used for the EDF and CPT. Parameters δ and γ are constrained by an upper bound of 1.75 and a lower bound of -0.25. The weights applied in these optimizations are due to the higher importance of matching probability tails in our analysis than the body of the distributions.

3 Overweight of tails: dynamics and dependencies

3.1 Time-varying CPT parameters

In this section, we evaluate the dynamics of the overweighting of tails within the single stock and index option markets. Descriptive statistics of the CPT's estimated δ and γ parameters via the methodology presented in Section 2 are provided in Table 1.

[Please insert Table 1 about here]

We report summary statistics of the estimated γ for three-, six- and twelve-month options in Panel A for the right tail from single stock options. The median and mean time-varying γ estimates for three-month options are 0.89 and 0.91, which considerably exceed the parameter value of 0.61 that is suggested by Tversky and Kahneman (1992). This finding suggests that overweight of small probabilities is present within the pricing of short-term single stock call options, but to a much lesser extent than suggested by the theory. Panel A also suggests that γ is highly time-varying and strongly sample dependent. Overweight of small probabilities in the single stock option market is very pronounced from 1998 to 2003 (present at 97 percent of all times), but infrequent from 2003 to 2008 (present at only 35 percent of all times). Our γ-estimates from three-month options range from 0 to 1.75 and the standard deviation of estimates is 0.23. In Panel B, we report summary statistics of the estimated δ from index options for the left tail. For δ estimated from three-month options, the median and mean estimates are both 0.68, implying a probability weighting that roughly matches the one in the CPT, which calibrates δ at 0.69. The δ-estimates are also time-varying, however, their standard deviation (0.08) is more than three times lower than for the γ-estimates. The range of δ-estimates is also much narrower than for γ, as it is between 0.29 and 1.01. In contrast to the γ-estimates, our δ-estimates reflect a consistent overweight of small probabilities across all sub-samples.

At the six-month maturity, overweight of small probabilities for γ seems even less acute than suggested by the theory and by the three-month options findings. The median and mean γ estimates for this maturity are 0.99 and 0.96, respectively. The distribution of γ is somewhat skewed to the right, i.e., towards a less pronounced overweight of small probabilities, as the median is higher than the mean. The 75^{th} quantile of γ (1.14) suggests an underweighting of probabilities already. For index options with six-month maturity, the estimated δ suggests an even more pronounced overweight of small probabilities (both the mean and median δ equal 0.60) than for three-month options. Overweight of small probabilities is again documented across all samples for δ but not for γ, in which overweight of small probabilities is more frequent than underweight of small probabilities only in the 1998-2003 sample.

The γ estimates for the twelve-month maturity tend even more towards probability underweighting than the six-month ones. The median γ is 1.03, whereas the mean γ is 1.01. Overweight of small probabilities appears in only 41 percent of all times in the overall sample and is roughly nonexistent in the 2003-2008 sample. Differently, the mean and median for the δ estimates from index options are 0.47 and 0.40, respectively, indicating an even stronger overweight of small probabilities than for single stock options and other maturities. We argue that such a pattern could be caused by institutional investors buying long-term protection, as twelve-month OTM index options are less liquid than short-term ones.

OTM index puts seem to be structurally expensive from the perspective of overweight of small probabilities, despite the fact that the degree of overvaluation varies in time. Concurrently, OTM single stock options are only occasionally expensive. Our γ estimates indicate an infrequent occurrence of overweight of small probabilities in single stock options, clustered within specific parts of our sample, e.g. the 1998-2003 period. Our results fit nicely within the seminal literature, for instance with Dierkes (2009), Kliger and Levy (2009), and Polkovnichenko and Zhao (2013), regarding the index option market, and with Felix et al. (2016) regarding the single stock option market.

3.2 Overweight of tails and sentiment

In order to evaluate how time-variation in overweight of small probabilities relates to sentiment, we run regressions between our proxies for overweight of tails (the explained variable), the Baker and Wurgler (2007) sentiment measure and other explanatory control variables. Since we aim to combine overweight of small probabilities parameters from both index options (bearish sentiment) and single stock options (bullish sentiment), we use the *Delta minus Gamma spread*, δ - γ, as the explained variable. The *Delta minus Gamma spread* captures the overweighting of small probabilities from both index options and single stock, because δ is the CPT tail overweight parameter estimated from the single stock market and γ is the equivalent parameter estimated from the index option market. The explanatory variables in these regressions are (1) the Baker and Wurgler (2007) sentiment measure[6], (2) the percentage of bullish investors minus the percentage of bearish investors given by the survey of the American Association of Individual Investors (AAII), (3) a proxy for individual investors' sentiment (see Han, 2008), and (4) a set of control variables among the ones tested by Welch and Goyal (2008)[7] as potential forecasters of the equity market. The data frequency used is monthly, as this is the highest frequency in which the Baker and Wurgler (2007) sentiment factor and the Welch and Goyal (2008) data set are available. Our regression samples start in January 1998 and ends in December 2011[8]. The OLS regression model applied is given as:

[6]Available at http://people.stern.nyu.edu/jwurgler/.

[7]The complete set of variables provided by Welch and Goyal (2008) that is used by us is discussed in Appendix B. In order to avoid multicollinearity in our regression analysis (some variables correlate 80 percent with each other), we exclude all variables that correlate more than 40 percent with others.

[8]This sample is only possible because Welch and Goyal (2008) and Baker and Wurgler (2007) have updated and made available their datasets after publication.

$$
\begin{aligned}
DGspread[\tau]_t = c + SENT_t + IISENT_t + E12_t + B/M_t + NTIS_t + \\
TBL_t + INFL_t + CORPR_t + SVAR_t + CSP_t + \epsilon_t,
\end{aligned}
\tag{12}
$$

where τ is the option horizon, $DGspread$ is the *Delta minus Gamma spread*, $SENT$ is the Baker and Wurgler (2007) sentiment measure, $IISENT$ is the AAII individual investor sentiment measure, $E12$ is the twelve-month moving sum of earnings of the S&P 5000 index, B/M is the book-to-market ratio, $NTIS$ is the net equity expansion, TBL is the risk-free rate, $INFL$ is the annual INFLation rate, $CORPR$ is the corporate spread, $SVAR$ is the stock variance, and CSP is the cross-sectional premium. We also run univariate models for each explanatory factor separately to understand their individual relation with the *Delta minus Gamma spread*:

$$
DGspread[\tau]_t = \alpha_i + \beta_i x_{i,t} + \epsilon_t, \tag{13}
$$

where x replaces the n explanatory variables specified in Eq. (12), given $i = 1...n$.

Table 2 Panel A reports the results of Eq. (12), estimated across our three maturities for the *Delta minus Gamma spread*. The explanatory power of the multivariate regression is very high, ranging from 36 to 57 percent. As expected, $SENT$ is positively linked to *Delta minus Gamma spread* and statistically significant across the three- and six-month maturities. This suggests that high sentiment exacerbates overweight of small probabilities measured as *Delta minus Gamma spread*. However, such a relation is negative and not significant at the twelve-month maturity. The univariate regressions of $SENT$ confirm the positive link between sentiment and *Delta minus Gamma spread* at shorter maturities. Once again, this relation is not present at the twelve-month horizon. The explanatory power of $SENT$ in the univariate setting is also high for the three- and six-month horizons, with 17 and 32 percent, respectively. This result strengthens our hypothesis that overweight of small probabilities increases at higher levels of sentiment and that sentiment seems to have a strong link to probability weighting by investors as priced by index puts and single stock call options. This conclusion, however, applies to the three- and six-month horizons only since the twelve-month univariate regression has a R^2 of zero.

[Please insert Table 2 about here]

$IISENT$ is also positively connected to *Delta minus Gamma spread* in the multivariate regression at the three- and six-month horizons but negatively at the twelve-month horizon. These results are confirmed by the univariate regressions, as $IISent$ is positively linked to *Delta minus Gamma spread* at the three- and six-month horizons. Explanatory power of these regressions is at 6 percent for both the three- and six-month maturities, which is relatively high. For the twelve-month maturity in the univariate regression, $IISENT$ is negatively linked to *Delta minus Gamma spread* and statistically significant.

Moving to the analysis of the other control variables in our regression, we observe that the results are less stable than for the sentiment proxies. Table 2 indicates that some signs of control variables change in both the multivariate and univariate regressions. TBL is the only control variable that remains statistically significant and keeps its sign across the multivariate and univariate models. The explanatory power of TBL is 21 percent in the univariate setting, whereas the other independent variable with high explanatory power is book-to-market with 27 percent. B/M is only statistically significant in the three-month maturity of the multivariate regressions. $NTIS$ is negatively and significantly linked to *Delta minus Gamma spread* in the univariate setting as well as in the multivariate regression in the twelve-month maturity. $SVAR$ is negatively and significantly linked to *Delta minus Gamma spread* in the univariate regression but in the multivariate regression this result is not observed. Overall, these empirical findings suggest that fundamentals have a relatively unstable link to the *Delta minus Gamma spread*.

We note that the stability of the relation between the sentiment factors and the *Delta minus Gamma spread* within the multivariate regressions evidences that sentiment and overweight of small probabilities are strongly connected.

3.3 Overweight of tails, IV skews and higher moments of the RND

In a next step, we assess the relationship between *Delta minus Gamma spread* and higher moments (skewness and kurtosis) of the RND implied by options and IV skew measures. We undertake this analysis for two reasons: 1) to understand to which extent *Delta minus Gamma spread* is connected to other metrics seemingly derived from IV, and 2) to approximate *Delta minus Gamma spread* by an easier-to-obtain measure, given the comprehensive estimation procedures required to compute γ and δ.

We expect the existence of a positive link between the estimated *Delta minus Gamma spread* and IV skew measures, because the presence of fat tails in the RND is a pre-condition for overweight of tail probabilities and a corollary of OTM's IVs to be rich versus at-the-money (ATM) IVs. Similarly, we observe negative skewness and fat-tails in RNDs only if OTM options are expensive versus ATM options and vice-versa[9]. Consequently, γ and δ are likely to be smaller than one (overweight of small probabilities), and *Delta minus Gamma spread* differs from zero if OTM options are expensive versus ATM options, which supports the use of IV skew as another proxy for overweight of tails.

The IV skew measures used at the beginning are the standard measures: 1) IV 90 percent (moneyness) minus ATM, 2) IV 80 percent minus ATM from index options (which captures bearish sentiment) and 3) IV 110 percent minus ATM and 4) IV 120 percent minus ATM from single stock calls (which captures bullish sentiment). However, as overweight of small probabilities is observed from the tails of the two markets jointly via *Delta minus Gamma spread*, and standard IV skew measures only capture information from one market at a time, we suggest a new IV-based measure. Our proposed IV skew sentiment metric, so-called *IV-sentiment*, is a combined measure of the index and single stock options markets. Our *IV-*

[9] While these relations are widely acknowledged, Jarrow and Rudd (1982) and Longstaff (1995) provide a formal theorem for the link between IV skew and risk-neutral moments, whereas Bakshi et al. (2003) offer a comprehensive empirical test of this proposition for index options.

sentiment measure is specified as follows:

$$IVsentiment = OTMindexputIV_{\tau p} - OTMsinglestockcallIV_{\tau c}, \tag{14}$$

where, the subscript $\tau = 1...3$ indexes the different option-maturities used, p specifies the moneyness levels 80 and 90 percent from index put options, and c specifies the moneyness levels 110 and 120 percent from single stock call options. Thus, our sentiment measure is calculated as permutations of IVs from the three-, six- and twelve-month maturities, and four points in the moneyness (80, 90, 110, and 120 percent) level grid, where the absolute distance from the two moneyness levels used per sentiment measure and the ATM level (100 percent moneyness) is kept constant. In other words, the *IV-sentiment* metric produced is restricted to the 80 minus 120 percent and the 90 minus 110 percent measures, hereafter called the *IV-sentiment 90-110* and *IV-sentiment 80-120* measures. From the granular data set across different moneyness levels and maturities, we create six distinct skew-based measures of *IV-sentiment*. Using such a construction, our *IV-sentiment* measure jointly incorporates bearishness sentiment from institutional investors and bullishness sentiment from retails investors, similarly to *Delta minus Gamma spread*.

We assess the isolated relationship between *Delta minus Gamma spread* and higher moments of the RND, (standard) IV skews, and our *IV-sentiment* measures using the univariate models presented by Eqs. (15) to (18). These models are estimated using OLS, where Newey-West standard errors are used for statistical inference. Our regression samples start in January 2, 1998 and end in March 19, 2013.

$$DGspread[\tau] = \alpha_t \left[\frac{K}{S}\right] + IVSent_t \left[\frac{K}{S}; \tau\right], \tag{15}$$

$$DGspread[\tau] = \alpha_t + KURT_t^m(\tau), \tag{16}$$

$$DGspread[\tau] = \alpha_t + SKEW_t^m(\tau), \tag{17}$$

$$DGspread[\tau] = \alpha_t \left[\frac{K}{S}\right] + IVSKEW_t \left[\frac{K}{S}; \tau\right], \tag{18}$$

where $\frac{K}{S}$ is the moneyness level of the option, τ is the option horizon, *DGspread* is the *Delta minus Gamma spread*, $IVSent$ is our *IV-sentiment* measure, $SKEW$ is the RND return skewness implied by options, $KURT$ is the RND return kurtosis implied by options, and $IVSKEW$ is the single market IV skew measure, for both index option and single stock option markets. We note that the superscript m for the variables $KURT$ and $SKEW$ aims to distinguish RND kurtosis and skewness obtained from either RND implied by index options ($m = io$) or single stock options ($m = sso$).

We estimate multivariate models of *Delta minus Gamma spread* regressed on RND skewness, kurtosis, IV skews and *IV-sentiment* to better understand the relation between these measures jointly and overweight of small probabilities:

$$DGspread[\tau] = \alpha_t \left[\frac{K}{S}\right] + SKEW_t^m(\tau) + KURT_t^m(\tau) + IVSent_t \left[\frac{K}{S};\tau\right], \tag{19}$$

Table 3 Panel A reports the estimates of Eqs. (15) to (18), when the *DGspread* is regressed on RND moments, IV skews and *IV-sentiment 90-110* in a univariate setting. The empirical findings suggest that *IV-sentiment* is the variable that explains *DGspread* the most across all maturities. The explanatory power of *IV-sentiment* is not only the highest but it is also the most consistent factor, as its R^2 ranges from 30 to 46 percent. *IV-sentiment* is negatively related to *DGspread*. Such a negative sign of the *IV-sentiment* regressor was expected because the *DGspread* rises with higher bullish sentiment, whereas higher *IV-sentiment* suggests a more pronounced bearish sentiment. Risk-neutral skewness and kurtosis also strongly explains *DGspread* (by roughly 30 percent), though only within the three-month maturity. Skewness and kurtosis explain *DGspread* by roughly 10 percent for six-month options, and 7 percent for twelve-month ones. The coefficient signs are in line with our expectations since high levels of RND skewness are associated with high *DGspread* (a bullish sentiment signal), while low levels of RND kurtosis (less pronounced fat-tails) are associated with high *DGspread*[10]. In contrast, standard IV skews explain very little of *DGspread* within the three-month maturity, only between 0 and 4 percent. At longer maturities, the IV skews are able to better explain *DGspread*, however, mostly when the skew measure comes from the single stock options market (between 17 and 21 percent). As a robustness check, we note that the regression results are virtually unchanged by the usage of either *IV-sentiment 90-110* or *80-120* measures. As a first impression, these results imply that *IV-sentiment* is strongly connected to *DGspread* and to overweight of small probabilities.

Panel B shows that when we evaluate the multivariate regressions, we find that *IV-sentiment* is the most stable regressor with respect to coefficient signs, being negatively linked to *DGspread* across all regressions, and is always statistically significant. These regressions have high explanatory power (ranging from 41 to 61 percent), especially when considering that the data frequency is daily, thus, potentially containing more noise than lower frequency data. In the multivariate regression we use the *IV-sentiment* 90-110, while the (unreported) results using *IV-sentiment* 80-120 are qualitatively the same. Due to likely multicollinearity in this multivariate model, we believe that our univariate models are more insightful than the former.

[Please insert Table 3 about here]

These findings strongly suggest that *DGspread* co-moves with our *IV-sentiment* measure within the three-, six-, and twelve-month maturities. Hence, we feel comfortable to use *IV-sentiment* to approximate the overweighting of small probabilities, similarly to *DGspread*.

[10]The regression results reported here use RND kurtosis and skewness from index options ($m = io$). The results when RND is extracted from single stock options ($m = sso$) are unreported but qualitatively the same as the coefficient signs are equal to the reported ones, and regressions' explanatory power are roughly in the same range.

4 Predicting with overweight of tails

4.1 Predicting returns with *DGspread* and *IV-sentiment*

Section 3.1 has documented that the overweighting of small probabilities is strongly time-varying and we hypothesize that it is linked to equity markets reversals. We hereby employ regression analysis to test if overweight of small probabilities (proxied by *Delta minus Gamma spread*) can predict equity market returns. Given the results of Section 3.3, in which our *IV-sentiment* measure strongly links to the *Delta minus Gamma spread*, we also run such predictive regressions by using *IV-sentiment* as the explanatory variable.

In order to test the predictability of these two metrics, we regress values of *Delta minus Gamma spread* and of our *IV-sentiment* measure on rolling forward returns with eight different investment horizons: 42, 84, 126, 252, 315, 525, 735, and 945 days, as specified by the Eqs. (20) and (21):

$$\frac{p_{t+h+1}}{p_{t+1}} = \alpha_h + \beta_h DGspread[\tau]_t + \epsilon_t, \tag{20}$$

$$\frac{p_{t+h+1}}{p_{t+1}} = \alpha_h + \beta_h IVSent[\tau]_t + \epsilon_t, \tag{21}$$

where p is the equity market price level, h is the forward return horizon, τ is the option maturity, α is the unconditional expected mean of forward returns, and β is the sensitivity of forward returns to *DGspread* and to *IV-sentiment*. We estimate Eqs. (20) and (21) via OLS with Newey-West adjustment to the standard deviation of regressors' coefficients due to the presence of serial correlation in forwards returns. Our regression samples start in January 2, 1998 and end in March 19, 2013.

Table 4 presents the empirical findings of forward returns regressed on *Delta minus Gamma spread*. The explanatory power of these regressions tends to have single-digit values, and it rarely exceeds ten percent. For the three-month horizon, the explanatory power of such predictive regressions rises steadily up to the two-year horizon (to nine percent), and drops then to four percent for forward returns at the 945-days horizon. We note that *Delta minus Gamma spread* tends to have low explanatory power and is not significant for short-horizons (42- to 126-days) and for higher maturities (twelve-month options). The coefficients of *Delta minus Gamma spread* are always negative for the three- and six-month maturities. This result was expected as it implies that a high (low) *Delta minus Gamma spread*, i.e., a bullish (bearish) sentiment predicts negative (positive) forward returns, i.e., reversals. For the twelve-month maturity, the coefficient signs are unstable, being negative (and statistically significant) for the 252-days horizon, while sometimes positive and insignificant for shorter horizons.

[Please insert Table 4 about here]

Panel B reports the regression results of Eq. (21), i.e., the outcomes of forward returns regressed on our *IV-sentiment 90-110* measure for three-, six-, and twelve-month maturities[11].

[11] The regression results for our *IV-sentiment 80-120* measure are qualitatively indifferent from the ones we

The pattern of R^2 across the different horizons tested is similar across the three option-maturities and analogous to the one observed for *Delta minus Gamma spread* for the same three-month horizon: R^2 rises from four percent to 28 percent when the horizon increases from 42 days (two months) to 525 days (two years), while after the two years horizon, explanatory power falls slightly for the 735 days (roughly three years) and collapses for the 925 days (3.7 years) horizon. We observe that the explanatory power for the six- and twelve-month option maturity is just slightly lower than for the three-month maturity. Statistical significance of the estimators is often high, across option maturities and return horizons. The coefficients for the *IV-sentiment 90-110* measure are always positive. This is as expected as it means that high (low) IV-sentiment, i.e., bearish (bullish) sentiment, predicts positive (negative) forward returns. The explanatory power, the stability of the coefficient signs, and the statistical significance of the regressors using our *IV-sentiment 90-110* measure clearly dominate the regression results that use *Delta minus Gamma spread*. These results strengthen our earlier findings that our *IV-sentiment* measure is a good representation of sentiment, especially concerning the prediction of equity market reversals.

4.2 *IV-sentiment* pair trading strategy

Our previous results suggest that *IV-sentiment* is more strongly connected to forward returns than *Delta minus Gamma spread* itself. We construct a trading strategy to further test the predictability power of *IV-sentiment*. The strategy consists of a high frequency (daily) trading rule that aims to predict equity reversals. Our hypothesis is that when the *IV-sentiment* measure is significantly higher (lower) than its normal level, overweight of small probabilities is then extreme and likely to mean-revert in the subsequent periods in tandem with the underlying market. The trading strategy, thus, buys (sells) equities when there is excessive bearishness/panic (excessive bullishness/complacency) indicated by the high (low) level of *IV-sentiment*.

The strategy is tested via a pair-trading rule among long and short positions in the S&P 500 index and a USD cash return index. For simplicity, such a strategy is implemented as a purely directional strategy where positions are constant in size and *IV-sentiment* is normalized via a Z-score. The trading rule enters a five percent long equities position when the *IV-sentiment* is higher than a pre-specified threshold, for example, its historical two standard deviation. The trading rule closes such a position, by entering into a full cash position, when such normalized *IV-sentiment* measure converges back to its average. Conversely, the rule enters a short equities position when the *IV-sentiment* is lower than its historical negative two standard deviation and buys back a full cash position when it converges to its average. Five basis points trading cost is charged over the five percent position traded in equities. In order to avoid strategy overfitting, we 1) compute the Z-score using multiple look-back periods, and 2) use multiple threshold levels to configure excessive sentiment[12]. We evaluate these contrarian strategies on

present for *IV-sentiment 90-110*.

[12]We also tested a percentile normalization and found results that are qualitatively similar to the use of Z-scores.

a volatility-adjusted basis using standard performance analytics such as the information ratio, downside risk characteristics, and higher moments of returns. We compare these strategies to 1) other contrarian strategies that make use of IV volatilities, such as an IV skew-based strategy, a volatility risk premia (VRP) strategy, and an implied-correlation-based (IC) strategy, 2) the equity market Beta, i.e., the S&P 500 index, and 3) alternative Beta strategies, such as writing put options, a 110-90 collar strategy, the G10 FX carry, equity cross-sectional momentum, and a time-series momentum strategy. We further evaluate such strategies by estimating the paired correlation coefficient between them, as well as tail and (distribution) higher-moment dependency statistics such as conditional co-crash (CCC) probabilities (see Appendix A.2) and co-skewness. Our back-test samples start in January 2, 1998 and end in December 4, 2015[13].

The boxplots of information ratios obtained by our *IV-sentiment* strategies and other IV-based strategies are provided in Figure 1. We see that the *IV-sentiment 90-110* strategy seems to perform better than the *IV-sentiment 80-120* strategy, as the information ratio means and dispersion of the former strategy dominate the ones for the latter. The average information ratio for the *IV-sentiment 90-110* strategy is positive for the three- and six-month option-maturities but negative for the twelve-month. For the three- and six-month strategies, all one-standard deviation boxes for the information ratio lay in positive territory, suggesting that the *IV-sentiment 90-110* strategy is robust to changes in look-back and outer-threshold parameters. Further, the *IV-sentiment 90-110* is superior to single-market IV skew-based strategies for the three- and six-month maturities, but not for the twelve-month maturity. At the three-month maturity, the average information ratio and dispersion for the *IV-sentiment 90-110* strategy are similar to the ones for the VRP strategy. However, for the six- and twelve-month maturities, the VRP strategies dominate the *IV-sentiment 90-110* based on the average information ratio, despite larger dispersion for the six-month maturity strategy.

[Please insert Figure 1 about here]

Figure 1 shows that the IC strategies seem to deliver relatively high and consistent information ratios, especially when calculated using the 80 and 90 percent moneyness levels. At the three- and six-month maturities, the performance of IC strategies match the performance of the *IV-sentiment 90-110* and VRP strategies. At the twelve-month horizon, the 80 and 90 percent IC strategies are superior to the *IV-sentiment 90-110* measure. Overall, the boxplots in Figure 1 suggest that the *IV-sentiment 90-110* strategy is robust to changes in parameters but also that its performance is matched by other IV-based strategies. Table 5 Panel A provides the performance analytics for the *IV-sentiment 90-110* strategy, as well as for alternative strategies.

[Please insert Table 5 about here]

We observe that the *IV-sentiment 90-110* strategy (using three-month option maturity) delivers returns (19 bps) and risk-adjusted returns (0.27) that are superior to many of the other

[13]As our *IV-sentiment* measure requires much less (cross-sectional) IV data than the *Delta minus Gamma spread* to be calculated, we were able to extend our full sample, which originally ended on March 19, 2013, until December 4, 2015

strategies compared, such as the S&P 500, the IV skew, the VRP, the IC, the G10 FX carry, and the equity momentum. Thus, the only strategies that deliver equal or higher risk-adjusted returns than our *IV-sentiment 90-110* strategy are the time-series momentum, the put writing, and a 90-110 collar. The return skewness for our *IV-sentiment* strategy is positive (0.10) and above the average of the other strategies. A strategy that has surprisingly high skewed returns is the IC (0.45). The drawdown characteristics such as the maximum drawdown, the average recovery time, and the maximum daily drawdown of our *IV-sentiment* strategy are very similar to the other IV-based strategies, except for the IC strategy, which is exposed to less downside risk.

In the following, we combine our *IV-sentiment* strategy with a simple buy and hold S&P 500, cross-sectional equity momentum, and time-series momentum strategies, on a standalone basis. These combinations are done by weighting returns in a 50/50 percent proportion. Statistics for the strategies are presented in columns (11) and (13) of Panel A of Table 5. We note that the combined strategies improve the information ratios of these three strategies. The information ratio for the S&P 500 rises from 0.20 to 0.25, for the cross-sectional momentum strategy from 0.16 to 0.24, and for the time-series momentum from 0.77 to 0.82. The drawdown characteristics are also marginally improved. We conjecture that these improvements in the information ratio should occur, at least partially, due to the low correlation and low higher moments-/tail-dependencies of our *IV-sentiment* strategy with these alternative strategies.

When we evaluate the correlation between the different strategies, the results in Table 5 Panel B indicate that our *IV-sentiment* strategy is, on average, positively related to other strategies. The highest correlation observed for the *IV-sentiment* strategy is with the IC strategy (0.70), which is an intuitive result given that these are the only two strategies driven jointly by the index option market and the single stock option market. The correlations of our *IV-sentiment* strategy with other IV-based strategies are also relatively high: 0.17 with the VRP and 0.40 with the IV skew 90 percent. The correlation of the *IV-sentiment* with the S&P 500 index is with 0.10, very low. The correlation of our *IV-sentiment* contrarian strategy with other strategies that perform poorly in "bad times" is also low, at 0.04 with the put writing, and 0.07 with the G10 FX carry, and 0.13 with the 90-110 collar strategy. The correlations between our contrarian strategy and equity momentum and time-series momentum are negative, at -0.16 and -0.11, respectively. We also note that other strategies can be highly correlated with each other, e.g., with 0.89 between the S&P 500 and the put-writing, whereas negative correlations are mostly observed for momentum strategies. Our findings on correlations among strategies are mostly reiterated by the estimated tail-dependence between them using co-skewness and CCC probabilities reported in Panel C of Table 5.

As a robustness check, we analyze whether our *IV-sentiment* high-frequency trading strategy performs well due to both its legs or whether its merit is concentrated in either the long- or the short-leg. We separate the performance of the two legs of the strategy as if they were two different strategies and we compute individual performance statistics. In order to visualize the results, we produce information ratios' (IRs) boxplots separately for the three option maturities, which are shown in Figure 2.

[Please insert Figure 2 about here]

The distribution of IRs for the long positions are shown in the plots at the upper part, while the distribution of IRs for the shorts are shown at the bottom. We note that the dispersion of IRs from the short-leg is much higher than from the long-leg; outliers are much more frequent in the short-leg. We find that the median IRs of long-legs are substantially higher than for short-legs. The IR distributions of the short positions seem slightly skewed to the negative side, whereas for the long positions they seem skewed to the positive side. These results indicate that the merit of our *IV-sentiment* strategy is concentrated in its buy-signal rather than in its sell-signal.

Figure 2 suggests that other IV-based strategies also seem to have their long-legs performing much better than their short-legs. This finding suggests that extreme bearish sentiment signals may be more reliable than extreme bullish sentiment signals. One explanation for this finding is the fact that the IV may be more reactive on the downside, due to the *leverage effect*[14]. In contrast, on the upside, a higher IV led by the bidding of call options might be offset by an overall lower IV. Our results are partially in line with the literature on cross-sectional returns and skew measures. Barberis and Huang (2008) suggest that stocks that have a high skew tend to have high subsequent returns, whereas for a call with a high skew this relation is inverse. However, other studies, such as Cremers and Weinbaum (2010), suggest that the relation between returns and volatility skews has the opposite direction. Assuming that there are systematic reasons for OTM implied volatilities across stocks to move in tandem, e.g., market risk, as suggested by Dennis and Mayhew (2002) and Duan and Wei (2009), then the logical consequence from the cross-sectional relation between the implied skew and returns would be that the overall equity market should reverse following times of extremely high skews.

Our results, thus, offer additional findings to the literature that explores the link between variance-measures and forward returns (Ang and Liu, 2007; Bliss and Panigirtzoglou, 2004; Doran et al., 2007; Pollet and Wilson, 2008). Most of these studies recognize a negative and short-term relation between risk measures and returns, where a high variance links to subsequent negative to low returns. In contrast, our findings suggest that a high level of IV skew relates to subsequent positive and high returns. Our finding is mostly in line with Bollerslev et al. (2009), who document that equity market reversals are predicted by the variance risk-premium.

Further, we aimed to compare the trading performance of the Baker and Wurgler (2007) sentiment measure to our high-frequency strategy but this was not possible as the former factor is only available on a monthly or quarterly frequency and was only published until 2010. Thus, in a next step, we compare how trading strategies using our suggested *IV-sentiment* measure compare with one that uses the sentiment factor of Baker and Wurgler (2007). We do this by implementing a low-frequency pair trade strategy using both predictors. This pair-trade

[14] The *leverage effect* refers to the typically observed negative correlation between equity returns and its changes of volatility and was first noted by Black (1976).

strategy is identical to the one applied above with the only difference being the rebalancing frequency and the number of observations in the look-back window. We use the following look-backs for the calculation of Z-scores: 1, 3, 6, 9, 12, 18, and 24 months. The *IV-sentiment* measures used are the *IV-sentiment 80-120* and *90-110* factors, available in our three different option maturities. Other back-test features (e.g., trading costs, strategy exit) are the same as for the high-frequency pair-trade strategy. Figure 3 provides our results by a series of boxplots. The empirical findings are displayed in columns for the different option maturities and in rows for the different statistics evaluated: 1) information ratio, 2) return skewness and 3) horizon, proxied by the average drawdown length (in months) observed per strategy.

[Please insert Figure 3 about here]

Our findings suggest that the IRs of the *IV-sentiment* strategies are much less dispersed than the ones for the sentiment factor by Baker and Wurgler (2007). The median IR for the *IV-sentiment 90-110* factor is also higher than for the other two strategies. The *IV-sentiment 90-110* factor is the only strategy in which almost all backtests deliver positive IRs, with the exception of a few outliers. This is not the case for the other strategies, as a substantial amount of backtests deliver negative IRs. In line with our earlier results, the *IV-sentiment 90-110* factor seems to dominate the *IV-sentiment 80-120* factor. The return skewness for the *IV-sentiment 90-110* strategy also dominates the ones for the other two strategies, as all boxplot features (median, one standard deviation, high and low percentile, and outliers) are superior. The *IV-sentiment 90-110* factor delivers the lowest median horizon of all strategies. The average horizons estimated for the *IV-sentiment 90-110* factor are 12, 13, and 19 months, respectively, for the strategies based on the three-, six- and twelve-month options. The dispersion of strategies' horizon is, however, higher for the *IV-sentiment 90-110* factor than for the Baker and Wurgler (2007) sentiment factor. We can conclude that our *IV-sentiment* measure seems to outperform a trading strategy based on the sentiment factor by Baker and Wurgler (2007) on several key aspects: IR, return skewness, and trade horizon.

4.3 Out-of-sample equity returns predictive tests

Following our hypothesis that extreme bearishness and bullishness sentiment might be followed by reversals in equity markets, we test in this subsection whether our *IV-sentiment* measure has out-of-sample predictive power in forecasting the equity risk premium (ERP), in line with the analysis introduced by Welch and Goyal (2008). We follow the methodology used by Campbell and Thompson (2008) and Rapach et al. (2010), who build on Welch and Goyal (2008). Hence, similarly to these three studies, our predictive OLS regressions are formulated as:

$$r_{t+1} = \alpha_i + \beta_i x_{i,t} + \epsilon_{t+1}, \tag{22}$$

where r_{t+1} is the monthly excess return of the S&P 500 index over the risk-free interest rate, x_t is an explanatory variable hypothesized to have predictive power, and ϵ_{t+1} is the error term.

Our predictive regressions also use the monthly data set provided by Welch and Goyal (2008)[15], but the scope of 14 explanatory variables used closely follows Rapach et al. (2010)[16].

From the predictive regressions in Eq. (22), we generate out-of-sample forecasts for the next quarter $(t+1)$ by using an expanding window. Following Rapach et al. (2010), the first parameters are estimated using data from 1947:1 until 1964:12, and forecasts are produced from 1965:1 until 2014:12. The estimating window for B/M starts slightly later than 1947:1, while the number of observations available allows forecasting B/M to start also at 1965:1. For the *IV-sentiment*-based regression, the data used for the first parameter estimation starts at 1998:1 and ends at 1999:12 so that out-of-sample forecasting is performed from 2000:1 to 2014:12 only.

Following Campbell and Thompson (2008) and Rapach et al. (2010), restrictions on the regression model specified by Eq. (22) are applied. The first restriction entails a sign restriction on the slope coefficients of Eq. (22) for the 14 Welch and Goyal (2008) variables used by us. The second restriction comprises setting negative forecasts of the ERP to zero. We specify an additional model containing both coefficient and forecast sign restrictions. The original Eq. (22) with no restrictions applied is called the *unrestricted model*, whereas the model with the two restrictions is called the *restricted model*. Once individual forecasts for r_{t+1} are obtained using the restricted and unrestricted models for every variable, weighted measures of central tendency (mean and median) of the N forecasts are generated by Eq. (23):

$$\hat{r}_{c,t+1} = \sum_{i=1}^{N} \omega_{i,t}\hat{r}_{i,t+1}, \tag{23}$$

where $(\omega_{i,t})_{i=1}^{N}$ are the combining weights available at time t. Our forecast combination method is a more simple and agnostic approach than the one used by Rapach et al. (2010)[17]. The mean and median combination methods are simply the equal weighed $(\omega_{i,t} = 1/N)$ average and median of the forecasts. Our benchmark forecasting model is the historical average model with the use of an expanding window.

We use the out-of-sample R^2 statistic method (R^2_{OS}) introduced by Campbell and Thompson (2008) and followed by Rapach et al. (2010) for forecast evaluation. This method compares the performance of a return forecast $\hat{r}_{t+1}$ and a benchmark or naïve return forecast $\bar{r}_{t+1}$ with the actual realized return (r_{t+1}). We note that this method can be applied either to the single factor-based forecast models as well as to the combined or multifactor forecast models, both

[15] Welch and Goyal (2008) monthly data was updated until December 2014 and is available at http://www.hec.unil.ch/agoyal/.

[16] These variables are: the dividend price ratio, the dividend yield, the earnings-price ratio, the dividend-payout ratio, the book-to-market ratio, the net equity issuance, the treasury bill rate, the long-term yield, the long-term return, the term spread, the default yield spread, the default return spread, the inflation rate and the stock variance.

[17] Rapach et al. (2010) classify their combination methods in two classes: the first class uses a mean, median, and trimmed mean approach for forecast combination, and the second class uses a discounted mean square prediction error (DMSPE) methodology. The DMSPE method aims to set combining weights as a function of the historical forecasting performance of the individual models over the out-of-sample period. This method weights more recent forecasts heavier than older ones by the use of one additional parameter. Despite the desirable features of such a second class combination method, we stick to the first class methods only because they are more transparent and do not require the choice of an additional parameter.

described in the previous section. The R^2_{OS} statistic is given as:

$$R^2_{OS} = 1 - \frac{\sum_{k=q_0+1}^{q}(r_{m+k} - \hat{r}_{m+l})^2}{\sum_{k=q_0+1}^{q}(r_{m+k} - \bar{r}_{m+l})^2}, \tag{24}$$

which evaluates the return forecasts from a predictive model (in the numerator) and the return forecasts from a benchmark or naïve model (in the denominator) by comparing the mean squared prediction errors (MSPE) for both methods. Because 1 is subtracted by the ratio of MSPEs in the R^2_{OS} statistic, its interpretation becomes: if $R^2_{OS} > 0$, then MSPE of $\hat{r}_{t+1}$ is smaller than for $\bar{r}_{t+1}$), indicating that the forecasting model outperforms the naïve (benchmark) model, and vice-versa.

The results from our out-of-sample equity returns predictive tests are reported in Table 6. Panel A reports the findings for the out-of-sample forecasting period between 1965:1 and 2014:12 for all individual variables except our *IV-sentiment* factor ($IVSent$), for which forecasts are only available from 2004:1-2014:12, and for the combined forecasts. For individual models, R^2_{OS} comes from the restricted model, whereas for the aggregated models, the results are reported for both the restricted and the unrestricted models. The results of the aggregate models are reported in means and medians, reflecting the aggregation method used.

[Please insert Table 6 about here]

Panel A suggests that performance is not consistent across factors within the longer history of the out-of-sample test. Some factors outperform others by a large amount. Concurrently, the performance of most single factors is quite inconsistent through time, as Figure 4 depicts: the slope and levels of R^2_{OS} constantly change from negative to positive and vice-versa for almost all factors. For some of them, R^2_{OS} even flips signal at times within the sample. In contrast, the aggregated models deliver better performance across restricted and unrestricted models using either averages or medians for aggregation method. Moreover, the performance of the weakest aggregate model (0.60) is superior to the best individual factor ($INFL$ at 0.56) within the full sample.

[Please insert Figure 4 about here]

Once we evaluate the 2004:1-2014:12 period, when $IVSent$ is used, we observe that the performance across factors remains inconsistent. The performance across individual factors looks less disperse in this sample than in the full sample, but overall worse. The $IVSent$ factor performs well (0.56), despite being strongly outperformed by the $SVAR$ factor, while other factors perform extremely poor ($NTIS$ at 0.61, $INFL$ at 0.56). The combined models that do not include $IVSent$ in their median versions (restricted and unrestricted) underperform the naïve forecasting benchmark as their R^2_{OS} is negative in the period. Interestingly, when our $IVSent$ factor is added to these models, the performance improves substantially, outperforming the benchmark. We observe the same for models based on the mean: the mean-unconstrained and the mean-constrained models ex-$IVSent$ show a R^2_{OS} of 0.06 and 0.09, respectively, within the period analyzed. When the $IVSent$ factor is added to them, R^2_{OS} improves to 0.15 and 0.17, respectively. Therefore, it appears that our $IVSent$ factor seems to impact the combined model in a very distinct way when compared to other factors.

We also find that our $IVSent$ is quite uncorrelated to other factors. The correlation coefficient of the $IVSent$ factor that uses three-month options with other individual factors is most of the times negative or close to zero, and only exceeds 0.5 when evaluated against LTY[18]. Such correlation is higher for the $IVSent$ factor computed using six- and twelve-month option maturities. These results suggest that the improvements made by our $IVSent$ factor to the combined models stem partially from diversification benefits rather than from forecast performance (R^2_{OS}) alone.

Observing the evolution of R^2_{OS} for the median-based (restricted and unrestricted) combined models in Plot A of Figure 5, we notice that both lines have slopes that are predominantly positive or flat. Positive slopes of the R^2_{OS} curve indicate that the combined model outperforms the benchmark out-of-sample. These R^2_{OS} lines match very closely the ones presented by Rapach et al. (2010) up to 2004, when their sample ends. The evolution of R^2_{OS} for our individual factors in Figure 4 is also very similar to Rapach et al. (2010): some R^2_{OS} curves are positively sloped during certain periods, but often all factors display negatively sloped curves. The R^2_{OS} curves for the $IVSent$ factor is mostly positively sloped but relatively flat from 2004 to 2007, as the last plot in Figure 4 indicates. These results reiterate the primary conclusion of Welch and Goyal (2008), Campbell and Thompson (2008) and Rapach et al. (2010): individual predictors that reliably outperform the historical average in forecasting the equity risk premium are rare but, once these models are sensibly restricted and aggregated in a multi-factor model, their out-of-sample predicting power improves considerably. This conclusion applies also to the inclusion of our $IVSent$ factor within the multi-factor model. Plot B of Figure 5 shows that the R^2_{OS} curves for the model that includes the $IVSent$ factor are visibly steeper than the ones that do not include it. Further, the findings in Figure 5 indicate that restricted models seem to be superior to unrestricted ones by having either higher or less volatile R^2_{OS}.

[Please insert Figure 5 about here]

However, even if the combined factor models perform much better than the individual predictors do, the red and black lines in Plots A and B of Figure 5 are not always positively sloped, which is in line with Rapach et al. (2010). The R^2_{OS} curve is strongly positively sloped from 1965 to 1975, more moderately positively sloped from 1975 to 1992, negatively sloped from 1992 to 2000, and then slightly positive to flat until 2008, when it sharply drops amid the global financial crisis up to December 2014. The addition of our $IVSent$ factor in the combined model produces the blue and green lines in Plot B of Figure 5. These new curves have an equally flat slope during the 2004 to 2008 period, while both experience a sharp rise since the beginning of 2008. These curves' profiles suggest that our $IVSent$ factor has considerably improved the out-of-sample performance of the combined model especially in times when the other factors broke down or did not provide an edge versus the historical average predictor. Thus, the inclusion of our $IVSent$ factor seems to revive the conclusion reached by the previous literature, where

[18] A full correlation matrix among the individual predictive factors tested by Rapach et al. (2010) and *IV-sentiment* factors can be provided upon request.

combined factor models are able to improve compared to individual factor models. At the same time, the recent poor performance of the combined models ex-$IVSent$ underscores that factor identification is still a major challenge for the specification of combined models. Overall, our empirical findings suggest that IV-based factors provide a relevant explanatory variable for the time-variation of equity returns.

4.4 Behavioral versus risk-sharing phenomena

Another perspective of equity market dynamics provided by IV-based factors that are jointly extracted from single stock and index options, is the implied correlation ($\bar{\rho}$). It is approximated by Eq. (25), which is derived in Appendix A.1:

$$\bar{\rho} \approx \frac{\sigma_I^2}{(\sum_{i=1}^{n} \sigma_i)^2}, \tag{25}$$

where σ_I^2 is the variance of index options, σ_i is the volatility of $i = 1...n$ stocks in the index, and w_i is the stocks' weight in the index. The implied correlation measures the level of the average correlation between stocks that are constituents of an index. The IV of index options, i.e., (σ_I^2), can be matched by the one of single stock options, weighted by its constituents' loadings in the index, i.e., $(\sum_{i=1}^{n} \sigma_i)^2$. Thus, if IV can be used as a measure of absolute expensiveness of an option, the implied correlation provides a relative valuation measure between the index and single stock options: a high (low) level of implied correlation means that index options are expensive (cheap) relative to single stock options.

Table 7 Panel A presents descriptive statistics of the implied correlations between the index and single stock options' IV. The means and medians suggest that the implied correlation monotonically decreases with an increase in the moneyness level. The implied correlation means range from 0.30 to 0.65, a somewhat wide range given that these are averaged measures. Such a relative high dispersion of implied correlations is confirmed by their standard deviations, which are around 0.14. The distributions of the implied correlation are mostly negative skewed, as medians are most of the times higher than their means. The most striking result is given by the maximum and minimum implied correlations: the maximum implied correlation observed across all maturities and moneyness levels reported reaches 135 percent. Implied correlations above 100 percent are observed for many options, mostly for puts at the 80 and 90 percent moneyness levels. This finding implies that in order to match the weighted IV of puts on single stocks that are part of the S&P 500 index to the IV of a put on the index (with same levels of moneyness), an average correlation above 100 percent between the single stock put options is required. However, as correlation coefficients are bounded between -100 and $+100$ percent, these levels of implied correlation are indicative of irrational behavior by investors, who bid up index puts to levels that contradict market completeness.

[Please insert Table 7 about here]

We also find that trading in the opposite direction of such evident irrational investor behavior has been very profitable, as implied correlations higher than 100 percent were very effective as an entry point for contrarian strategies. Across the maturities and moneyness levels where we can observe such biased behavior, a sentiment strategy that buys the equity market when the implied correlation is above 100 percent and sells it when the implied correlation falls back to 50 percent, yields an average net information ratio of 0.35, with information ratios ranging from 0.27 to 0.52.

The implied correlation means and medians provided by Panel A are far higher than the same measures from realized average pair-correlations between the 50 largest constituents of the S&P 500 index as of February 14, 2014, as provided in Panel B. Such average pair-correlations range from 0.25 to 0.36 when look-back periods of 30, 60, 90, 180, and 720 days are evaluated, which is substantially lower than most average implied correlations posted for the different option maturity and moneyness levels reported in Panel A. In fact, the average realized correlations are often below the 10^{th} percentile of the implied correlation for some options' maturity and moneyness levels. The 90^{th} percentile of realized correlations often match the average implied correlations reported. The maximum realized correlations are at most 84 percent, using an extremely short look-back of 30 days, much lower than the 135 percent observed for implied correlations. These empirical findings strongly suggest that implied correlations substantially overshoot realized ones. Similarly, the implied correlation reaches sometimes values as low as three percent for some options, especially on the call side (above ATM moneyness). This finding is also low when compared to put options. The minimum historical correlations from OTM puts is 0.18, whereas for call options it is 0.03. The fact that those extremely low values of the implied correlation from calls largely undershoot implied correlations from put options may also suggest less than fully rational pricing on the call side. It indicates that single stock options are expensive relative to index calls, which matches our postulation that individual investors use single stock calls to speculate on the upside.

Despite the strong evidence of irrational behavioral by investors provided by the extreme levels of implied correlation, which indirectly links to the IV skew being at extreme levels at times, we conjecture that such phenomena may also have a risk-bearing explanation. Reversal strategies such as the ones designed by us earn attractive long-term risk-adjusted returns, but are highly dependent on equity markets at the tail (see Table 5, Panel C). Additionally, *IV-sentiment*-based reversal strategies experience the largest daily drawdowns among all strategies evaluated (see Table 5, Panel A). Thus, their attractive risk-adjusted returns are, partially, compensation for downside risk. Therefore, the risk borne by investors that bet on reversals in equity markets is the risk of poor timing of losses (Campbell and Cochrane, 1999; Harvey and Siddique, 2000) and downside risk (Ang et al., 2006). In brief, betting on equity market reversals is a risky activity.

We note that this rational explanation for excesses in sentiment is also linked to *limits-to-arbitrage*. The *limits-to-arbitrage* literature defends that, as investors have finite access to capital (Brunnermeier and Pedersen, 2009) and feedback trading can keep markets irrational for a long period of time (De Long et al., 1990), contrarian strategies aiming to exploit the effect of irrational trading are not without risk. For example, once bearishness sentiment seems excessive, the risk of betting on a reversal may be tolerable only to a few investors, because 1) higher volatility drags investors risk budget usage closer to its limits, and 2) access to funding is limited. Thus, the ability to “catch a knife falling” in the equity markets is not suitable for all investors, as it involves high risk. Contrarian strategies are, then, mainly accessible to investors that have enough capital or funding liquidity. Similar considerations are career risk (Chan et al., 2002), negative skewness of returns (Harvey and Siddique, 2000), poor timing of losses (Campbell and Cochrane, 1999; Harvey and Siddique, 2000), and risk aversion of market makers (Garleanu et al., 2009). One final element in the characterization of reversals as a compensation for risk is the presence of correlation risk priced in index options (see Driessen et al., 2009, 2013; Krishnam and Ritchken, 2008; Jackwerth and Vilkov, 2015), which is present in assets that perform well when market-wide correlations are higher than expected.

5 Conclusion

End-users of OTM options overweight small probability events, i.e., tail events. This bias is strongly time-varying and present in both OTM index puts and OTM single stock calls, due to individual and institutional investors trading activity, respectively. Individual investors typically buy OTM single stock calls (“lottery tickets”) to speculate on the upside of equities (indicating bullish sentiment), whereas institutional investors typically buy OTM index puts (portfolio insurance) to protect their large equity holdings (indicating bearish sentiment). Hence, overweight of small probabilities derived from equity option prices should capture investors’ sentiment and, thus, potentially predict equity returns.

The parameters that directly capture overweight of small probabilities from option prices such as the Delta (δ) and Gamma (γ) CPT parameters or the *Delta minus Gamma spread* (as designed by us) are difficult to estimate. Due to the fact that *Delta minus Gamma spread* is found to be strongly linked to risk-neutral moments and IV skews, we circumvent such estimation challenges by proposing a simplified but still informative sentiment proxy: *IV-sentiment*. The uniqueness of our *IV-sentiment* measure is that it is jointly calculated from the IV of OTM index puts and OTM single stock call options. It aims to capture both bullish and bearish sentiment, respectively, from individual investors and institutional investors’ trading in options.

Our results confirm that the *IV-sentiment* carries substantial information. The first supporting evidence of this conclusion is that our *IV-sentiment* predicts mean-reversion better than the overweighting of small probabilities parameter *Delta minus Gamma spread*. We also test the predictive power of the *IV-sentiment* measure in the context of multi-factor predictive regressions and of two trading strategies, one high-frequency pair-trade and a low frequency strategy, which we compare to the Baker and Wurgler (2007) sentiment factor. In the high-frequency context, contrarian-trading strategies using our sentiment measure produce economically significant risk-adjusted returns. The joint use of information from the single stock and index option markets seems to be the reason for the superior forecast ability of our sentiment measure, because factors that use implied volatility skews from a single market achieve significantly inferior results. The performance of our *IV-sentiment* measure seems also more consistent in delivering a positive information ratio than the Baker and Wurgler (2007) sentiment factor. Moreover, it is more positively skewed, has a shorter horizon than the standard factor and allows for a daily strategy rebalancing. This is an interesting finding given the popularity of the Baker and Wurgler (2007) factor within the sentiment literature. Furthermore, our *IV-sentiment* factor seems to forecast returns as well as other well-known predictors of equity returns. Because our sentiment factor is uncorrelated to other predictors of the equity risk-premium, it significantly improves the quality of multi-factor predictive regressions, especially when such models are constrained, as in the terms of Campbell and Thompson (2008).

The prediction of reversals seems to be further enhanced when the volatility skews priced by OTM index puts and ATM single stock calls are clearly irrational, e.g., when implied correlations are higher than 100 percent. Timing market reversals using our *IV-sentiment* measure is, however, not without risk. Reversal strategies, as ours, are exposed to large drawdowns, which likely happen during 'bad times'. Nevertheless, we find that combining our sentiment strategy with other strategies, such as a long-only exposure to the S&P 500 index, time-series and cross-sectional equity momentum can strongly improve risk-adjusted returns. For momentum strategies, this finding is backed by the fact that our contrarian-sentiment strategy is negatively correlated to these strategies, with low dependence at the tails. These findings indicate a promising avenue for future research on (the prevention of) momentum crashes. Since sell-sentiment signals from our *IV-sentiment* measure may be not as reliable as buy signals, further refinement of our findings is warranted, as well as the search for factors that capture bullish sentiment better.

A Appendix: Methodology

A.1 Weighted average single stock IV and implied correlation approximations

Starting from the portfolio variance formula, Eq. (A.1a), we derive in the following the weighted average single stock IV, Eq. (A.1k), and the implied correlation approximation, Eq. (A.1i), as given in Eq. (25) in the main text:

$$\sigma_I^2 = \sum_{i,j=1}^{n} w_i w_j \rho_{ij} \sigma_i \sigma_j, \tag{A.1a}$$

where,

$$\rho_{ij}(x) = \begin{cases} \bar{\rho}, & if \quad i \neq j \\ 1, & if \quad i = j \end{cases} \tag{A.1b}$$

and where σ_I^2 is the equity index option implied variance and i and j are indexes for the constituents of such equity index, then:

$$\sigma_I^2 = \bar{\rho} \sum_{i \neq j}^{n} w_i w_j \rho_{ij} \sigma_i \sigma_j + \sum_{i=1}^{n} w_i^2 \sigma_i^2, \tag{A.1c}$$

$$= \bar{\rho} \sum_{i,j=1}^{n} w_i w_j \rho_{ij} \sigma_i \sigma_j + (1 - \bar{\rho}) \sum_{i=1}^{n} w_i^2 \sigma_i^2, \tag{A.1d}$$

$$= \bar{\rho} \left(\sum_{i=1}^{n} w_i \sigma_i \right)^2 + (1 - \bar{\rho}) \sum_{i=1}^{n} w_i^2 \sigma_i^2, \tag{A.1e}$$

$$= \bar{\rho} \left(\sum_{i=1}^{n} w_i \sigma_i \right)^2 + \sum_{i=1}^{n} w_i^2 \sigma_i^2 - \bar{\rho} \sum_{i=1}^{n} w_i^2 \sigma_i^2, \tag{A.1f}$$

$$= \bar{\rho} \left(\left(\sum_{i=1}^{n} w_i \sigma_i \right)^2 - \sum_{i=1}^{n} w_i^2 \sigma_i^2 \right) + \sum_{i=1}^{n} w_i^2 \sigma_i^2, \tag{A.1g}$$

$$\bar{\rho} \approx \frac{\sigma_I^2 - \sum_{i,j=1}^{n} w_i^2 \sigma_i^2}{\left(\sum_{i=1}^{n} w_i \sigma_i\right)^2 - \sum_{i,j=1}^{n} w_i^2 \sigma_i^2}. \tag{A.1h}$$

As $\sum_{i,j=1}^{n} w_i^2 \sigma_i^2$ is relatively small, we can simplify Eq. (A.1h) into (A.1i), the implied correlation:

$$\bar{\rho} \approx \frac{\sigma_I^2}{\left(\sum_{i=1}^{n} w_i \sigma_i\right)^2} \tag{A.1i}$$

In order to obtain the weighted average single stock implied volatility, Eq. (A.1k), we then square root both sides of the approximation and re-arrange its terms:

$$\sqrt{\bar{\rho}} \approx \frac{\sigma_I}{(\sum_{i=1}^{n} w_i \sigma_i)} \tag{A.1j}$$

with

$$\sum_{i=1}^{n} w_i \sigma_i \approx \frac{\sigma_I}{\sqrt{\bar{\rho}}}. \tag{A.1k}$$

A.2 Conditional co-crash probabilities

We use a bivariate Extreme Value Theory (EVT) method to calculate commonality on historical tail returns for the strategies highlighted in Section 4.2. EVT is well suited to measure contagion risk because it does not assume any specific return distribution. Our approach estimates how likely it is that one stock will experience a crash beyond a specific extreme negative return threshold conditional on another stock crash beyond an equally probable threshold. We refer to Hartmann et al. (2004) who use the conditional co-crash (CCC) probability estimator, which is applied to each pair of stocks in our sample, as follows:

$$\widehat{CCC}_{ij} = 2 - \frac{1}{k} \sum_{t=1}^{N} I[V_{it} > x_{i,N-k} \quad or \quad V_{jt} > x_{j,N-k}], \tag{A.2}$$

where the function I is the crash indicator function, in which $I = 1$ in case of a crash, and $I = 0$ otherwise, V_{it} and V_{jt} are returns for stocks i and j at time t; $x_{i,N-k}$, and $x_{j,N-k}$ are extreme crash thresholds. The estimation of the CCC-probabilities requires setting k as the number of observations used in Eq. (A.2).

B Appendix: Equity market control variables and predictors

The complete set and summarized descriptions of variables provided by Welch and Goyal (2008)[19] that are used in our study is given as:

1. **Dividendprice ratio (log), D/P**: Difference between the log of dividends paid on the S&P 500 index and the log of stock prices (S&P 500 index).
2. **Dividend yield (log), D/Y**: Difference between the log of dividends and the log of lagged stock prices.
3. **Earnings, E12**: 12-month moving sum of earnings on teh S&P500 index.
4. **Earnings-price ratio (log), E/P**: Difference between the log of earnings on the S&P 500 index and the log of stock prices.
5. **Dividend-payout ratio (log), D/E**: Difference between the log of dividends and the log of earnings.
6. **Stock variance, SVAR**: Sum of squared daily returns on the S&P 500 index.
7. **Book-to-market ratio, B/M**: Ratio of book value to market value for the Dow Jones Industrial Average.
8. **Net equity expansion, NTIS**: Ratio of twelve-month moving sums of net issues by NYSE-listed stocks to total end-of-year market capitalization of NYSE stocks.
9. **Treasury bill rate, TBL**: Interest rate on a three-month Treasury bill.
10. **Long-term yield, LTY**: Long-term government bond yield.
11. **Long-term return, LTR**: Return on long-term government bonds.
12. **Term spread, TMS**: Difference between the long-term yield and the Treasury bill rate.
13. **Default yield spread, DFY**: Difference between BAA- and AAA-rated corporate bond yields.
14. **Default return spread, DFR**: Difference between returns of long-term corporate and government bonds.
15. **Cross-sectional premium, CSP**: measures the relative valuation of high- and low-beta stocks.
16. **Inflation, INFL**: Calculated from the CPI (all urban consumers) using $t-1$ information due to the publication lag of inflation numbers.

[19] Available at http://www.hec.unil.ch/agoyal/.

References

Amin, K., Coval, J., Seyhun, H., 2004. Index option prices and stock market momentum. *Journal of Business* 77 (4), 835–874.

Ang, A., Chen, J., Xing, Y., 2006. Risk, return and dividends. *Review of Financial Studies* 19 (4), 1191–1239.

Ang, A., Liu, J., 2007. Risk, return and dividends. *Journal of Financial Economics* 85 (1), 1–38.

Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21 (Spring), 129–157.

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

Barberis, N., 2013. The psychology of tail events: Progress and challenges. *American Economic Review* 103 (3), 611–616.

Barberis, N., Huang, M., 2008. Stocks as lotteries: The implication of probability weighting for security prices. *American Economic Review* 98 (5), 2066–2100.

Bates, D., 1991. The crash of 87: Was it expected? the evidence from options markets. *Journal of Finance* 46 (3), 1009–1044.

Bates, D., 2003. Empirical option pricing: A retrospection. *Journal of Econometrics* 116, 387–404.

Black, F., 1976. Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association*, 171–181.

Bliss, R., Panigirtzoglou, N., 2004. Option-implied risk aversion estimates. *Journal of Finance* 59 (1), 407–446.

Bollen, N., Whaley, R., 2004. Does net buying pressure affect the shape of implied volatility function? *Journal of Finance* 59 (2), 711–754.

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

Boyer, B., Vorkink, K., 2014. Stock options as lotteries. *Journal of Finance* 69 (4), 1485–1527.

Brunnermeier, M., Pedersen, L., 2009. Market liquidity and funding liquidity. *Review of Financial Studies* 22 (6), 2201–2238.

Campbell, J., Cochrane, J., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107 (2), 205–251.

Campbell, J., Thompson, S., 2008. Predicting the equity premium out of sample: Can anything beat the historical average? *Review of Financial Studies* 21 (4), 1509–1531.

Chan, L., Chen, H.-L., Lakonishok, J., 2002. On mutual fund investment styles. *Review of Financial Studies* 15 (5), 1407–1437.

Cornell, B., 2009. The pricing of volatility and skewness: A new interpretation. *Journal of Investing* 18, 27–30.

Cremers, M., Weinbaum, D., 2010. Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis* 45, 335–367.

De Long, J., Shleifer, A., Summers, L., Waldmann, R., 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98 (4), 703–738.

Dennis, P., Mayhew, S., 2002. Risk-neutral skewness: evidence from stock options. *Journal of Financial and Quantitative Analysis* 37 (3), 471–493.

Dierkes, M., 2009. Option-implied risk attitude under rank-dependent utility. *Unpublished working paper. University of Münster, Münster*, Germany.

Doran, J., Peterson, D., Tarrant, B., 2007. Is there information in the volatility skew? *Journal of Future Markets* 27 (10), 921–959.

Driessen, J., Maenhout, P., Vilkov, G., 2009. The price of correlation risk: Evidence from equity options. *Journal of Finance* 64 (3), 1377–1406.

Driessen, J., Maenhout, P., Vilkov, G., 2013. Option-implied correlations and the price of correlation risk. *SSRN working paper no.* 2166829, 1–46.

Duan, J.-C., Wei, J., 2009. Systematic risk and the price structure of individual equity options. *Review of Financial Studies* 22 (5), 1981–2006.

Felix, L., Kraussl, R., Stork, P., 2016. Single stock call options as lottery tickets. *SSRN working paper* 2736900.

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* - Oxford University Press.

Garleanu, N., Pedersen, L. H., Poteshman, A. M., 2009. Demand-based option pricing. *Review of Financial Studies* 22 (10), 4259–4299.

Han, B., 2008. Investor sentiment and option prices. *Review of Financial Studies* 21 (1), 387–414.

Hartmann, P., Straetmans, S., de Vries, C., 2004. Asset market linkages in crisis periods. *The Review of Economics and Statistics* 86 (1), 313–326.

Harvey, C., Siddique, A., 2000. Conditional skewness in asset pricing tests. *Journal of Finance* 60 (3), 1263–1296.

Jackwerth, J., Rubinstein, M., 1996. Recovering probability distributions from options prices. *Journal of Finance* 51 (5), 1611–1631.

Jackwerth, J., Vilkov, G., 2015. , asymmetric volatility risk: Evidence from option markets. *SSRN working paper* 2325380.

Jarrow, R., Rudd, A., 1982. Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics* 10 (5), 347–369.

Kliger, D., Levy, O., 2009. Theories of choice under risk: Insights from financial markets. *Journal of Economic Behavior & Organization* 71 (2), 330–346.

Krishnam, C., P. R., Ritchken, P., 2008. Correlation risk. *SSRN Working Paper* 1027479, 1–31.

Lakonishok, J., Lee, I., Pearson, N., Poteshman, A., 2007. Option market activity. *Review of Financial Studies* 20 (3), 813–857.

Lehmann, B., 1990. Fads, martingales, and market efficiency. *Quarterly Journal of Economics* 105, 1–28.

Lemmon, M., Ni, S., 2011. The effects of investor sentiment on speculative trading and prices of stock and index options. *SSRN working paper no.* 1572427.

Longstaff, F., 1995. Option pricing and the martingale restriction. *Review of financial studies* 8 (4), 1091–1124.

Mahani, R., Poteshman, A., 2008. Overreaction to stock market news and misevaluation of stock prices by unsophisticated investors: evidence from the options market. *Journal of Empirical Finance* 15 (4), 635–655.

Mitton, T., Vorkink, K., 2007. Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies* 20 (4), 1255–1288.

Nelson, D., 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59 (2), 347–370.

Polkovnichenko, V., Zhao, F., 2013. Probability weighting functions implied in option prices. *Journal of Financial Economics* 107 (3), 580–609.

Pollet, J., Wilson, M., 2008. Average correlation and stock market returns. *Journal of Financial Economics* 96 (3), 364–380.

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

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

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5 (4), 297–323.

Vilkov, G., Xiao, Y., 2013. Option-implied information and predictability of extreme returns. *SAFE (Goethe University Frankfurt) Working Paper Series* 5, 1–36.

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

Table 1: Descriptive statistics

This table reports the summary statistics of the estimated cumulative prospect theory (CPT) parameters gamma (γ) from the single stock options market and delta (δ) from the index option market for each day in our sample as well as the optimizations' residual sum of squares (RSS). The parameters γ and δ define the curvature of the weighting function for gains and losses, respectively, which leads the probability distortion functions to have inverse S-shapes. The γ and δ parameters close to unity lead to weighting functions that are close to unweighted (neutral) probabilities, whereas parameters close to zero indicates large overweight of small probabilities. Panel A reports the summary statistics of gamma (γ) when we assume a parameter of risk aversion (λ) equal to 2.25 (the standard CPT parametrization). Panel B reports the summary statistics of delta (δ) under the same risk aversion assumption. Column headings % $\gamma < 1$ and % $\delta < 1$ report the percentage of observations in which parameters γ and δ are smaller than one, i.e., the proportion of the sample in which overweight of small probabilities is observed. We report this metric for the full sample as well as for three equal-sized splits of our full samples, namely: 98-03, from 1998-01-05 to 2003-01-30; 03-08, from 2003-01-31 to 2008-02-21; and 08-13, from 2008-02-22 to 2013-03-19.

Panel A - Gamma

Maturity	Min	25% Q	Median	Mean	75% Q	Max	StDev	% $\gamma < 1$	% $\gamma < 1$ (98-03)	% $\gamma < 1$ (03-08)	% $\gamma < 1$ (08-13)	RSS
3 months	-	0.74	0.91	0.89	1.04	1.75	0.23	64%	97%	35%	59%	0.0209
6 months	-	0.81	0.99	0.96	1.14	1.75	0.28	52%	92%	18%	46%	0.0170
12 months	0.04	0.91	1.03	1.01	1.14	1.75	0.22	41%	83%	11%	29%	0.0225

Panel B - Delta

Maturity	Min	25% Q	Median	Mean	75% Q	Max	StDev	% $\gamma < 1$	% $\gamma < 1$ (98-03)	% $\gamma < 1$ (03-08)	% $\gamma < 1$ (08-13)	RSS
3 months	0.29	0.64	0.68	0.68	0.72	1.01	0.08	100%	100%	100%	100%	0.0579
6 months	0.30	0.54	0.60	0.60	0.65	1.75	0.10	100%	100%	100%	100%	0.0198
12 months	-	0.40	0.45	0.47	0.52	1.75	0.10	100%	100%	100%	100%	0.0169

Table 2: Regression results: ***Delta minus Gamma spread***

Panel A reports the regression results for Eq. (12) in a multivariate setting. The dependent variable is *Delta minus Gamma spread* (δ - γ),while as explanatory variables we specify: 1) the Baker and Wurgler (2007) sentiment measure (*SENT*), 2) the individual investor sentiment (*IISENT*), and 3) the explanatory variables used by Welch and Goyal (2008), while excluding factors that correlate to each other in excess of 40 percent (see Appendix B for the full list of variables). Panel B reports the regression results for (13), in an univariate setting, in which *Delta minus Gamma spread* is regressed on the same set of explanatory variables. We report Newey-West adjusted standard errors in brackets. Asterisks ***, **, and * indicate significance at the one, five, and ten percent level, respectively.

Panel A - Multivariate				Panel B - Univariate													
Maturity	3m	6m	12m	3m	6m	12m	3m	6m	12m	6m	6m	6m	6m	6m	6m	6m	6m
Intercept	0.003	-0.491***	-0.490***	-0.063***	-0.369***	-0.520***	-0.064***	-0.365***	-0.508***	-0.048	0.131***	-0.055***	-0.121***	-0.055***	-0.053***	-0.039***	-0.052***
	(0.056)	(0.037)	(0.058)	(0.010)	(0.008)	(0.013)	(0.011)	(0.010)	(0.012)	(0.031)	(0.031)	(0.011)	(0.015)	(0.013)	(0.011)	(0.012)	(0.011)
SENT	0.030^*	0.064***	-0.024	0.071***	0.097***	-0.003											
	(0.017)	(0.013)	(0.019)	(0.014)	(0.016)	(0.016)											
IISENT	0.041	0.096**	-0.106**				0.123***	0.125**	-0.124***								
	(0.047)	(0.038)	(0.048)				(0.044)	(0.050)	(0.043)								
E12	0.000	-0.003	-0.028***							-0.001							
	(0.006)	(0.004)	(0.007)							(0.006)							
B/M	-0.364^*	0.125	0.163								-0.737***						
	(0.217)	(0.132)	(0.211)								(0.130)						
NTIS	0.560	0.259	-0.814									1.075**					
	(0.391)	(0.285)	(0.523)									(0.440)					
TBL	0.013	0.036***	0.029***										0.030***				
	(0.008)	(0.006)	(0.009)										(0.006)				
INFL	0.453	1.843	2.311											1.784			
	(2.507)	(1.885)	(2.176)											(3.350)			
CORPR	0.225	0.233	0.044												0.128		
	(0.285)	(0.202)	(0.273)												(0.472)		
SVAR	-1.426	3.519***	3.470^*													-3.376**	
	(1.331)	(1.153)	(1.982)													(1.307)	
CSP	-0.125	0.198	0.261														0.029
	(0.136)	(0.122)	(0.235)														(0.197)
R^2	36%	57%	30%	17%	32%	0%	6%	6%	5%	0%	27%	4%	21%	0%	0%	4%	0%
F-stats	8.2	19.5	6.4	32.5	72.9	0.0	9.1	9.4	8.0	0.0	58.0	7.1	40.7	0.7	0.2	6.1	0.0
AIC	-308.1	-369.1	-273.2	-326.1	-186.0	34.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BIC	-274.4	-335.4	-239.6	-320.0	-179.9	40.3	0.0	0.0	0.0	0.0	0.1	0.4	0.0	2.2	0.3	1.4	0.2

Table 3: Regression results: *Delta minus Gamma spread* and risk-neutral measures

Panel A reports the regression results for Eqs. (15), (16), (17) and (18) in an univariate setting. The dependent variable for these regressions is *Delta minus Gamma spread* (δ - γ), a proxy for overweight of small probabilities. As explanatory variables we specify the risk-neutral skewness and kurtosis, IV 110-ATM skew (from single stock options), IV 90-ATM skew (from index options), and our *IV-sentiment* measure in two permutations per maturity: 1) *IV-sentiment 90-110*, and 2) *IV-sentiment 80-120*. Our *IV-sentiment* measure is an IV skew measure that combines information from the index option market and the single stock option market, see Eq. (14). For instance, the *IV-sentiment 90-110* measure combines the IV from the 90 percent moneyness level from the index option market and the 110 percent moneyness level from the single stock option market. Panel B reports the regression results for Eq. (19) in a multivariate setting, in which *Delta minus Gamma spread* is regressed on the same set of explanatory variables. We report Newey-West adjusted standard errors in brackets. Asterisks ***, **, and * indicate significance at the one, five, and ten percent level, respectively.

Panel A - Univariate regressions												
Maturity	3m	6m	12m	3m	6m	12m	3m	6m	12m	3m	6m	12m
Intercept	0.019**	-0.219***	-0.436***	-0.045***	-0.254***	-0.460***	-0.295***	-0.499***	-0.683***	-0.186***	-0.368***	-0.593***
	(0.007)	(0.010)	(0.009)	(0.006)	(0.008)	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Skewness	0.122***	0.073***	0.054***									
	(0.003)	(0.004)	(0.003)									
Kurtosis				-0.015***	-0.009***	-0.007***						
				(0.000)	(0.000)	(0.000)						
IV-sentiment 90-110							-1.998***	-2.774***	-2.359***			
							(0.046)	(0.064)	(0.062)			
IV-sentiment 80-120										-1.606***	-2.438***	-2.124***
										(0.042)	(0.058)	(0.056)
R^2	32%	9%	7%	30%	10%	7%	34%	46%	36%	30%	45%	35%
F-stats	1861.1	408.6	285.7	1714.6	423.4	315.6	2085.1	3423.7	2228.0	1707.2	3199.9	2121.2
AIC	-2001	-13	-944	-1900	-27	-972	-2151	-2093	-2437	-1895	-1971	-2368
BIC	-1989	-1	-931	-1888	-14	-959	-2139	-2081	-2424	-1883	-1959	-2355

Panel A - Univariate regressions (continuation)							Panel B - Multivariate regressions		
Maturity	3m	6m	12m	3m	6m	12m	3m	6m	12m
Intercept	-0.195***	-0.141***	-0.332***	0.029	-0.052	-0.407***	-0.273***	-0.465***	-0.495***
	(0.010)	(0.013)	(0.011)	(0.028)	(0.034)	(0.027)	(0.025)	(0.038)	(0.040)
Skewness							0.093***	0.000	-0.032***
							(0.008)	(0.009)	(0.009)
Kurtosis							-0.002**	-0.007***	-0.009***
							(0.001)	(0.001)	(0.001)
IV-sentiment 90-110							-1.989***	-2.462***	-1.677***
							(0.063)	(0.108)	(0.126)
IV 110-ATM skew	1.082**	13.681***	16.172***				0.511	5.525***	8.106***
	(0.435)	(0.717)	(0.711)				(0.371)	(0.672)	(1.004)
IV 90-ATM skew				-4.941***	-8.399***	-4.997***	3.876***	4.129***	0.000
				(0.557)	(0.903)	(0.993)	(0.391)	(0.732)	(0.933)
R^2	0%	17%	21%	4%	4%	1%	61%	53%	42%
F-stats	10.3	810.0	1065.3	148.4	177.3	49.4	1214.5	903.8	707.4
AIC	-485	-362	-1612	-621	202	-717	-4154	-2636	-3008
BIC	-472	-349	-1599	-608	215	-705	-4116	-2598	-2970

Table 4: Regression results: Delta minus gamma spread and *IV-sentiment*

Panel A reports the regression results for Eq. (20), which regresses the *Delta minus Gamma spread* on eight different horizons for forward equity returns. Panel B reports the regression results for Eq. (21), which regresses the *IV-sentiment 90-110* measure on the same forward equity returns used in Panel A. The explained variables are forward returns for the S&P 500 index measured over the following horizons: 42, 84, 126, 252, 315, 525, 735, and 945 days. We report Newey-West adjusted standard errors in brackets. The asterisks ***, **, and * indicate significance at the one, five, and ten percent level, respectively.

Panel A - Delta minus Gamma spread									Panel B - IV-sentiment 90-110							
Three-month options																
Horizon	42	84	126	252	315	525	735	945	42	84	126	252	315	525	735	945
Intercept	0.000	-0.004	-0.009**	-0.016***	-0.023***	-0.030***	-0.027***	0.003	0.01***	0.01***	0.02***	0.04***	0.04***	0.06***	0.08***	0.08***
	(0.003)	(0.003)	(0.004)	(0.006)	(0.006)	(0.008)	(0.008)	(0.009)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)	(0.007)
DGspread/ IV-Sent	-0.03***	-0.08***	-0.13***	-0.26***	-0.33***	-0.43***	-0.39***	-0.19***	0.13***	0.26***	0.38***	0.65***	0.70***	1.11***	1.59***	1.52***
	(0.007)	(0.009)	(0.010)	(0.015)	(0.018)	(0.032)	(0.034)	(0.036)	(0.013)	(0.017)	(0.020)	(0.021)	(0.021)	(0.031)	(0.033)	(0.048)
R^2	1%	2%	4%	7%	9%	9%	7%	1%	4%	8%	10%	15%	16%	23%	34%	26%
F-stats	31.4	81.9	150.1	274.5	347.3	351.0	228.5	42.4	150.7	317.6	418.3	678.2	709.1	1083.3	1737.3	1096.3
AIC	0.0017	0.0023	0.0030	0.0042	0.0047	0.0061	0.0066	0.0075	0.0011	0.0015	0.0019	0.0026	0.0027	0.0034	0.0039	0.0047
BIC	0.0061	0.0085	0.0108	0.0157	0.0175	0.0232	0.0255	0.0288	0.0106	0.0146	0.0184	0.0251	0.0261	0.0337	0.0382	0.0460
Six-month options																
Horizon	42	84	126	252	315	525	735	945	42	84	126	252	315	525	735	945
Intercept	-0.004	-0.017***	-0.023***	-0.050***	-0.066***	-0.123***	-0.138***	-0.088***	0.02***	0.04***	0.05***	0.10***	0.12***	0.18***	0.19***	0.14***
	(0.003)	(0.004)	(0.005)	(0.007)	(0.008)	(0.009)	(0.008)	(0.009)	(0.002)	(0.002)	(0.003)	(0.005)	(0.005)	(0.007)	(0.009)	(0.011)
DGspread/ IV-Sent	-0.03***	-0.08***	-0.12***	-0.25***	-0.32***	-0.53***	-0.56***	-0.39***	0.24***	0.45***	0.64***	1.12***	1.38***	2.18***	2.22***	1.48***
	(0.007)	(0.009)	(0.010)	(0.014)	(0.017)	(0.023)	(0.024)	(0.028)	(0.022)	(0.027)	(0.029)	(0.032)	(0.040)	(0.050)	(0.059)	(0.066)
R^2	1%	3%	4%	8%	10%	16%	17%	7%	5%	8%	10%	15%	18%	26%	23%	9%
F-stats	31.7	114.0	145.2	301.5	399.4	669.1	651.7	232.6	186.6	344.4	447.9	660.3	803.0	1204.0	959.3	288.5
AIC	0.0023	0.0031	0.0040	0.0057	0.0062	0.0079	0.0084	0.0097	0.0014	0.0019	0.0025	0.0036	0.0040	0.0054	0.0063	0.0079
BIC	0.0056	0.0077	0.0099	0.0142	0.0159	0.0204	0.0221	0.0257	0.0176	0.0241	0.0304	0.0437	0.0486	0.0629	0.0717	0.0873
Twelve-month options																
Horizon	42	84	126	252	315	525	735	945	42	84	126	252	315	525	735	945
Intercept	0.007	0.004	0.012	-0.037***	-0.100***	-0.177***	-0.228***	-0.164***	0.02***	0.04***	0.06***	0.11***	0.14***	0.21***	0.21***	0.13***
	(0.005)	(0.007)	(0.009)	(0.013)	(0.014)	(0.019)	(0.020)	(0.021)	(0.002)	(0.003)	(0.003)	(0.005)	(0.006)	(0.008)	(0.009)	(0.012)
DGspread/ IV-Sent	0.00	-0.02	-0.01	-0.14***	-0.27***	-0.44***	-0.53***	-0.40***	0.25***	0.46***	0.68***	1.23***	1.53***	2.30***	2.23***	1.26***
	(0.009)	(0.012)	(0.015)	(0.022)	(0.024)	(0.033)	(0.037)	(0.040)	(0.024)	(0.029)	(0.030)	(0.035)	(0.043)	(0.056)	(0.068)	(0.077)
R^2	0%	0%	0%	2%	5%	8%	11%	5%	4%	7%	10%	15%	18%	23%	19%	5%
F-stats	0.0	2.8	0.9	62.8	198.7	313.8	396.4	166.3	160.0	299.2	409.8	636.9	806.9	1048.0	753.0	162.8
AIC	0.0037	0.0052	0.0066	0.0096	0.0105	0.0136	0.0144	0.0164	0.0016	0.0022	0.0028	0.0041	0.0046	0.0063	0.0074	0.0093
BIC	0.0066	0.0093	0.0119	0.0172	0.0190	0.0249	0.0267	0.0310	0.0194	0.0267	0.0336	0.0486	0.0539	0.0709	0.0813	0.0987

Table 5: *IV-sentiment* based pair-trade strategy

Panel A reports the results of contrarian pair-trade strategies based on our *IV-sentiment 90-110* indicator and on other IV-based strategies such as the IV Skew, the volatility-risk premia (VRP), and other traditional and alternative beta strategies, i.e., buy & hold the S&P500 index, put writing, 90-110 collar, G10 FX carry, cross-sectional equity momentum, and time-series momentum. The IV-based strategies use 252 days as the look-back period and +/- two standard deviations as convergence thresholds. The columns (11) and (12) of Panel A report statistics for strategies that combine the three-month *IV-sentiment 90-110* strategy (column (1)) with the buy & hold the S&P 500 index (column (5)) and the time-series momentum strategy (column (10)). Panel B reports the correlation coefficients of daily returns estimated over the period between January 2, 1998 and December 4, 2015, for the same strategies reported in Panel A. Panel C reports the co-skewness and the conditional co-crash (CCC) probabilities of the three-month *IV-sentiment 90-110* with the other strategies, which indicate the degree of tail-dependence among them.

Panel A Back-test results	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	IV-sentiment 90-110	IV Skew 3m 90	VRP 3m 90	IC 3m 90	S&P500	Put writing	90-110 Collar	G10 FX carry	Equity Momentum	CTA	S&P500 +IVSent	Eq Mom + IVSent	CTA +IVSent
Average return	0.19%	0.14%	0.12%	0.16%	0.14%	0.36%	0.19%	0.15%	0.11%	0.53%	0.17%	0.17%	0.57%
Volatility	0.69%	0.69%	0.69%	0.69%	0.69%	0.69%	0.69%	0.69%	0.69%	0.69%	0.69%	0.69%	0.69%
Information ratio	0.279	0.197	0.167	0.234	0.202	0.526	0.275	0.212	0.157	0.767	0.248	0.239	0.824
Skewness	0.10	-0.07	-0.01	0.45	-0.20	-0.61	-0.01	-0.91	-0.48	-0.37	-0.30	-0.60	-0.37
Kurtosis	16.98	26.42	30.96	19.85	7.93	23.40	2.15	11.87	4.10	2.91	13.07	7.15	2.91
Max drawdown	-1.7%	-1.6%	-2.8%	-1.7%	-2.9%	-2.4%	-2.8%	-3.1%	-2.8%	-1.3%	-2.8%	-2.6%	-1.3%
Avg. recovery time	0.43	0.42	0.41	0.35	0.17	0.05	0.17	0.18	0.23	0.13	0.17	0.22	0.14
Max daily drawdown	-0.55%	-0.53%	-0.49%	-0.47%	-0.33%	-0.52%	-0.28%	-0.49%	-0.33%	-0.31%	-0.42%	-0.43%	-0.30%

Panel B Correlations	IV-sentiment 90-110	IV Skew 3m 90	VRP 3m 90	IC 3m 90	S&P500	Put writing	90-110 Collar	G10 FX carry	Equity Momentum	CTA
IV-sentiment	1	0.408	0.177	0.703	0.101	0.038	0.128	0.068	-0.160	-0.109
IV Skew	0.408	1	0.551	0.589	0.176	0.165	0.079	0.158	-0.046	-0.033
VRP	0.177	0.551	1	0.510	0.407	0.422	0.179	0.155	-0.134	-0.110
IC	0.703	0.589	0.510	1	0.343	0.308	0.238	0.139	-0.212	-0.131
S&P500	0.101	0.176	0.407	0.343	1	0.891	0.880	0.281	-0.052	-0.150
Put writing	0.038	0.165	0.422	0.308	0.891	1	0.682	0.260	-0.085	-0.164
90-110 Collar	0.128	0.079	0.179	0.238	0.880	0.682	1	0.234	0.133	-0.046
G10 FX carry	0.068	0.158	0.155	0.139	0.281	0.260	0.234	1	0.023	-0.050
Equity Momentum	-0.160	-0.046	-0.134	-0.212	-0.052	-0.085	0.133	0.023	1	0.303
CTA	-0.109	-0.033	-0.110	-0.131	-0.150	-0.164	-0.046	-0.050	0.303	1

Panel C Tail dependence with IV sentiment	Sentiment 90-110	IV Skew 3m 90	VRP 3m 90	IC 3m 90	S&P500	Put writing	90-110 Collar	G10 FX carry	Equity Momentum	CTA
Co-skewness	1.6E-12	-2.9E-13	1.0E-11	4.6E-12	-6.5E-10	-3.3E-09	9.6E-10	-9.8E-10	5.9E-10	-3.2E-09
1% cond. crash prob.	100.0%	51.1%	36.2%	70.2%	23.4%	21.3%	23.4%	6.4%	14.9%	23.4%
2% cond. crash prob.	100.0%	42.6%	46.8%	72.3%	31.9%	27.7%	28.7%	12.8%	8.5%	27.7%
5% cond. crash prob.	100.0%	39.7%	44.4%	82.9%	29.1%	32.5%	26.1%	15.0%	10.3%	32.5%

Table 6: Out-of-sample equity risk premium

This table reports the results from the predictive regressions of individual factor models and of combined-factor models relative to the historical average naïve (benchmark) model. R^2_{OS} is the Campbell and Thompson (2008) out-of-sample R^2 statistic. If $R^2_{OS} > 0$, then mean squared prediction errors (MSPE) of $\bar{r}_{t+1}$, i.e., the predictive regression forecast, is smaller than for $\bar{r}_{t+1}$, i.e., the naïve forecast, indicating that the forecasting model outperforms the latter (benchmark) model. Panel A reports the results for the full out-of-sample period available (1965:1-2014:12) for all variables tested by Rapach et al. (2010). Panel B reports the results for the latest period within the entire out-of-sample history (2004:1-2014:12) and includes the three-month *IV-sentiment 90-110* factor (*IVSent*) in addition to the variables tested by Rapach et al. (2010).

Individual predictive regression model forecast				Combination forecasts	
Predictor	R^2_{OS} (%)	Predictor	R^2_{OS} (%)	Combining method	R^2_{OS} (%)
(1)	(2)	(3)	(4)	(5)	(6)
Panel A. 1965:1-2014:12 out-of-sample period					
D/P	-0.34	LTY	-0.33	Mean-Unconstrained	1.16
D/Y	-0.13	TMS	-0.58	Median-Unconstrained	0.6
E/P	-0.47	LTR	0.26		
D/E	-0.88	DFY	-0.8	Mean-Constrained	1.21
B/M	-1.02	DFR	-0.64	Median-Constrained	0.63
NTIS	-0.96	TBL	-0.02		
INFL	0.56				
SVAR	0.02				
Panel B. 2004:1-2014:12 out-of-sample period					
D/P	-0.19	LTY	0.14	Mean-Unconstrained	0.06
D/Y	-0.12	TMS	-0.22	Median-Unconstrained	-0.08
E/P	-0.3	LTR	0	Mean-Unconstrained + IVSent	0.15
D/E	-0.49	DFY	-0.29	Median-Unconstrained+IVSent	0.06
B/M	-0.04	DFR	-0.15	Mean-Constrained	0.09
NTIS	-0.61	TBL	-0.01	Median-Constrained	-0.06
INFL	-0.59	IVSent	0.56	Mean-Constrained + IVSent	0.17
SVAR	0.96			Median-Constrained + IVSent	0.04

Table 7: Implied and realized correlations

Panel A reports the descriptive statistics for the implied correlations between index options and single stock options for three month options at the 80, 90, ATM (100), 110, and 120 percent moneyness levels, and for six- and twelve-month options at the 80 and 90 percent moneyness levels over the full sample, which extends from January 2, 1998 to March 19, 2013. The implied correlation ($\bar{p}$) is approximated by the Eq. (25): $\bar{p} \approx \frac{\sigma_I^2}{(\sum_{i=1}^{n} w_i \sigma_i)^2}$, where σ_I^2 is the implied volatility of an index option and $\sum_{i=1}^{n} w_i \sigma_i$ is the weighted average single stock implied volatility, as in Eq. (A.1k) of Appendix A. Panel B reports the descriptive statistics for the average pair-correlations for the 50 largest constituents of the S&P500 index calculated over the same sample, which extends from January 2, 1998 to March 19, 2013.

Panel A Implied correlations

Statistics \ Maturity, moneyness	3m 80%	3m 90%	3m ATM	3m 110%	3m 120%	6m 80%	6m 90%	12m 80%	12m 90%
Mean	0.65	0.56	0.45	0.35	0.3	0.64	0.56	0.6	0.54
Median	0.67	0.57	0.45	0.35	0.3	0.65	0.56	0.61	0.54
Minimum	0.24	0.18	0.12	0.07	0.03	0.26	0.21	0.26	0.22
Maximum	1.35	1.11	0.86	0.72	0.68	1.07	0.95	1.1	1.01
10th percentile	0.44	0.35	0.27	0.17	0.13	0.41	0.34	0.39	0.34
90th percentile	0.81	0.73	0.63	0.53	0.49	0.8	0.73	0.77	0.72
Standard deviation	0.15	0.14	0.14	0.13	0.14	0.14	0.14	0.14	0.13
Skew	-0.46	-0.39	0.1	0.29	0.29	-0.6	-0.38	-0.26	-0.2
Excess Kurtosis	0.6	-0.02	-0.37	-0.38	-0.66	0.09	-0.18	0.03	-0.22

Panel B - Realized correlations

Statistics \ Look-back period	30 Days	60 Days	90 Days	180 Days	720 Days
Mean	0.3	0.25	0.25	0.26	0.36
Median	0.27	0.22	0.24	0.25	0.31
Minimum	0	0.01	0	0.01	0.06
Maximum	0.84	0.69	0.67	0.61	0.74
10th percentile	0.1	0.05	0.04	0.07	0.08
90th percentile	0.54	0.47	0.48	0.52	0.71
Standard deviation	0.17	0.16	0.16	0.16	0.2
Skew	0.66	0.38	0.6	0.37	0.42
Excess Kurtosis	-0.02	-0.68	-0.21	-0.86	-0.88

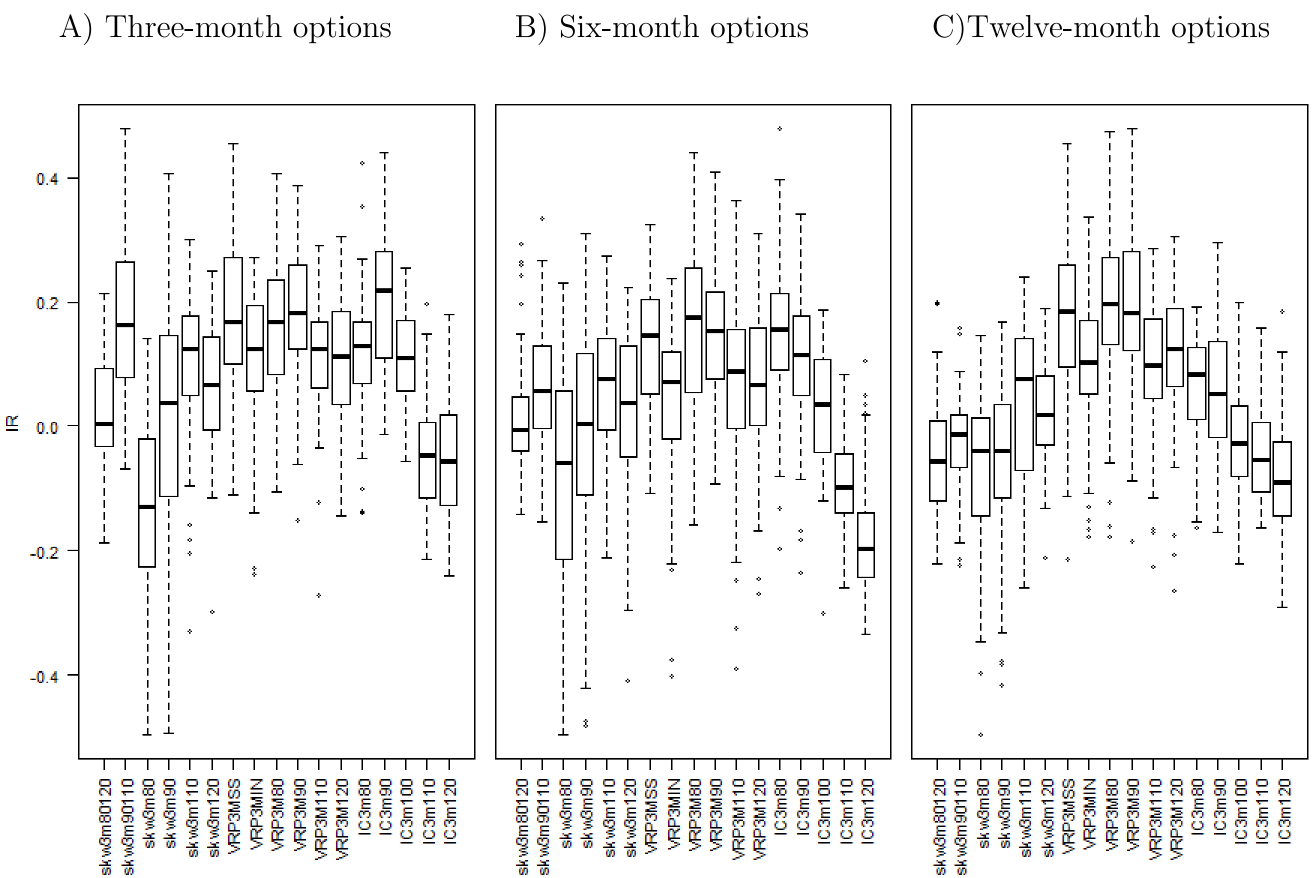


Figure 1: **Information ratio boxplots for daily IV-based strategies.** The boxplots depict the distribution of information ratios (IR) obtained by the IV-based strategies tested, when different look-back periods and outer-thresholds are used per factor-specific strategy. Boxplot A depicts the distribution of IRs when the IV factor used is obtained from three-month options. Panels B and C depict the same information while using the IV factors obtained from six- and twelve-month options, respectively.

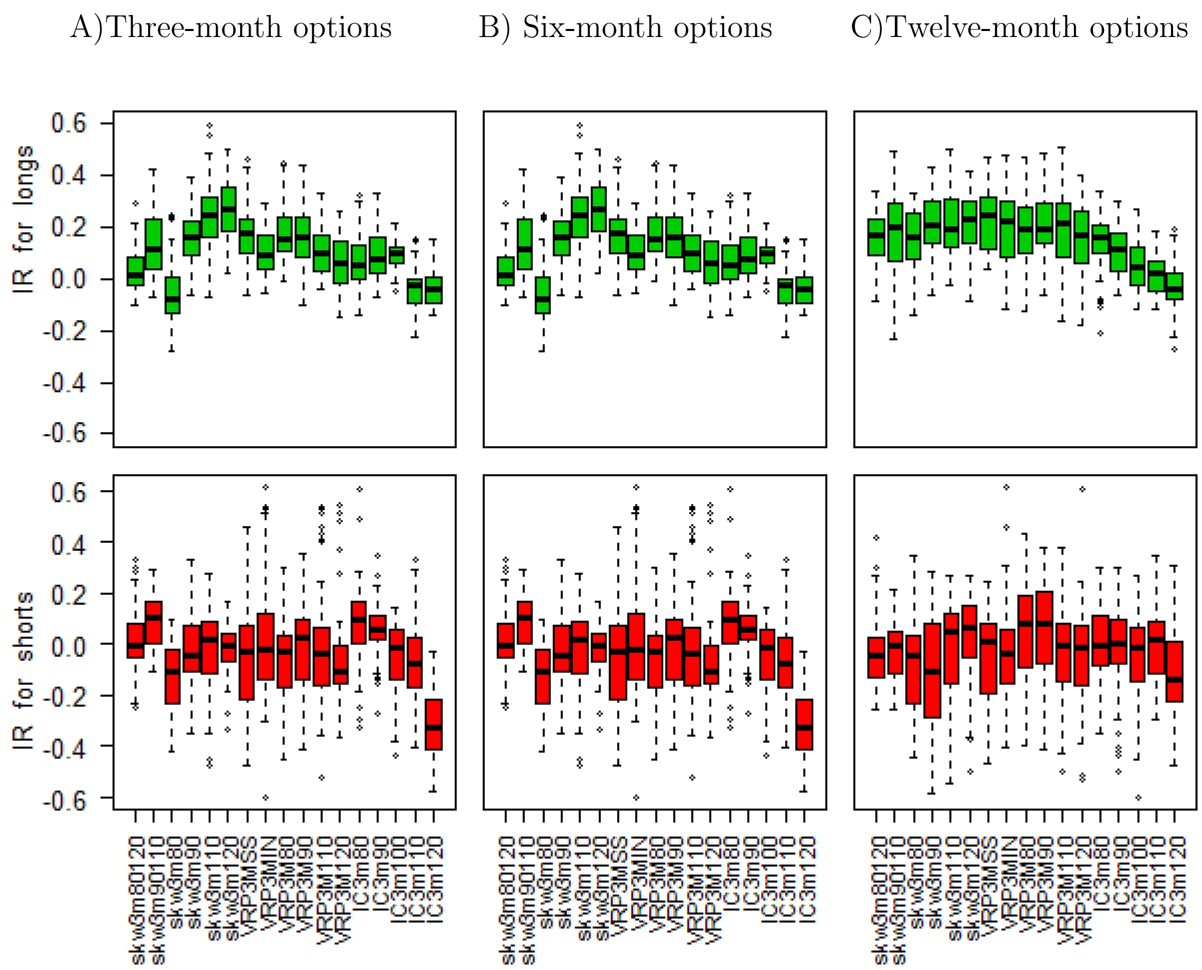


Figure 2: **Information ratio boxplot for long- and short-leg of IV-based strategies.** The boxplots depict the distribution of information ratios (IRs) obtained by the IV-based strategies tested, when different look-back periods and outer-threshold are used per factor-specific strategy. Boxplots on the top row (in green) refer to IRs produced by the long-leg of IV-based strategies, whereas the ones in the second row refer to the short-leg of the same strategies. Boxplot A depicts the distribution of IRs when the IV factor used is obtained from three-month options. Panels B and C depict the same information while using the IV factors obtained from six- and twelve-month options, respectively.

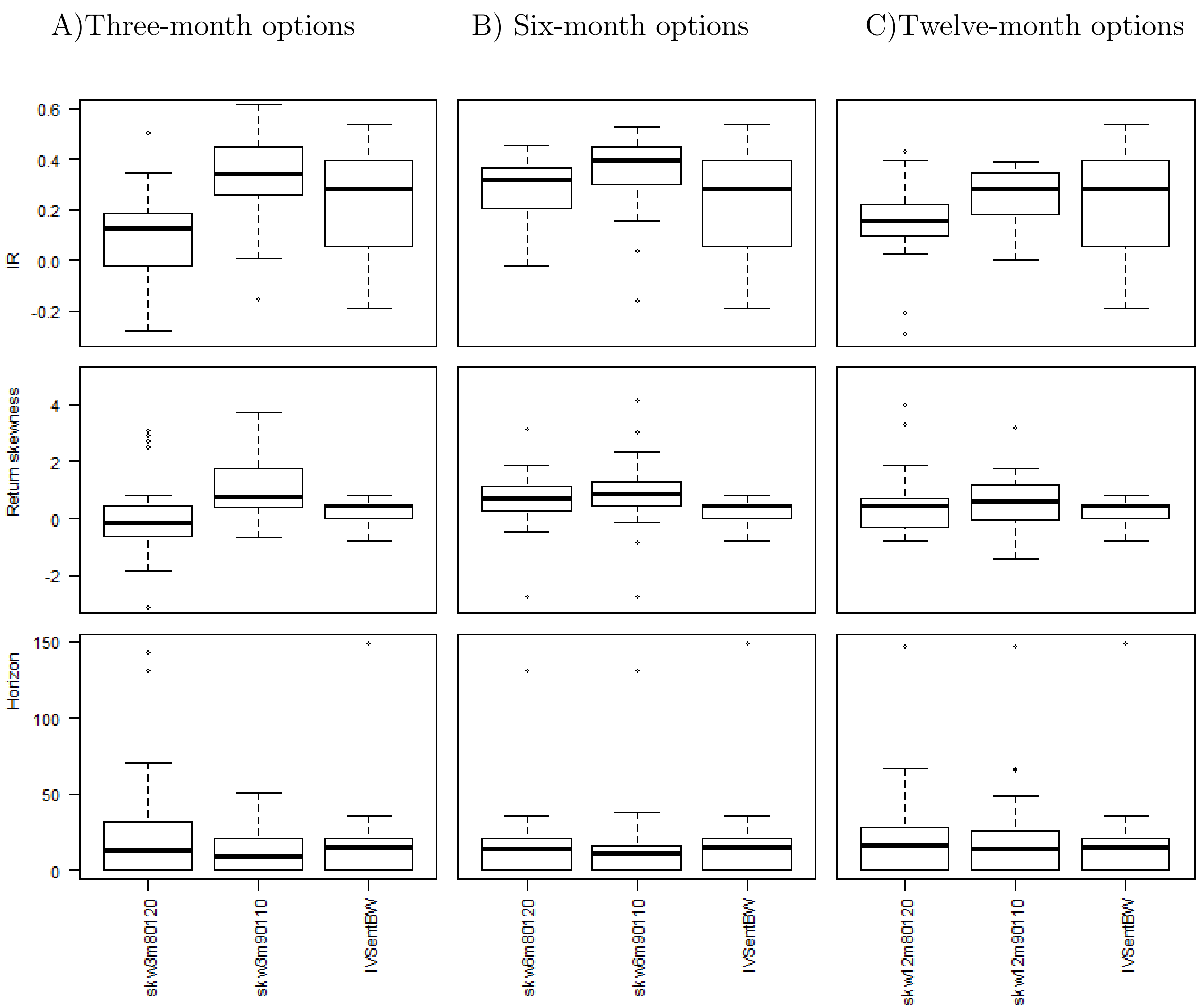


Figure 3: **Information ratio, skewness and horizon for monthly IV-based strategies.** The boxplots depict the distribution of information ratios (IRs), return skewness, and trade horizon (average drawdown) obtained by the *IV-sentiment* strategies tested, as well as the Baker and Wurgler (2007) sentiment factor when different look-back periods and outer-threshold are used per strategy. Boxplot A depicts the distribution of these statistics when the IV factor used is obtained from three-month options. Panel B and C depicts the same information but, respectively, when the IV factors used are obtained from six- and twelve-month options. Boxplots of IR, return skewness and trade horizon for the Baker and Wurgler (2007) factor are the same across option horizons but are shown for comparison with the *IV-sentiment* strategies.

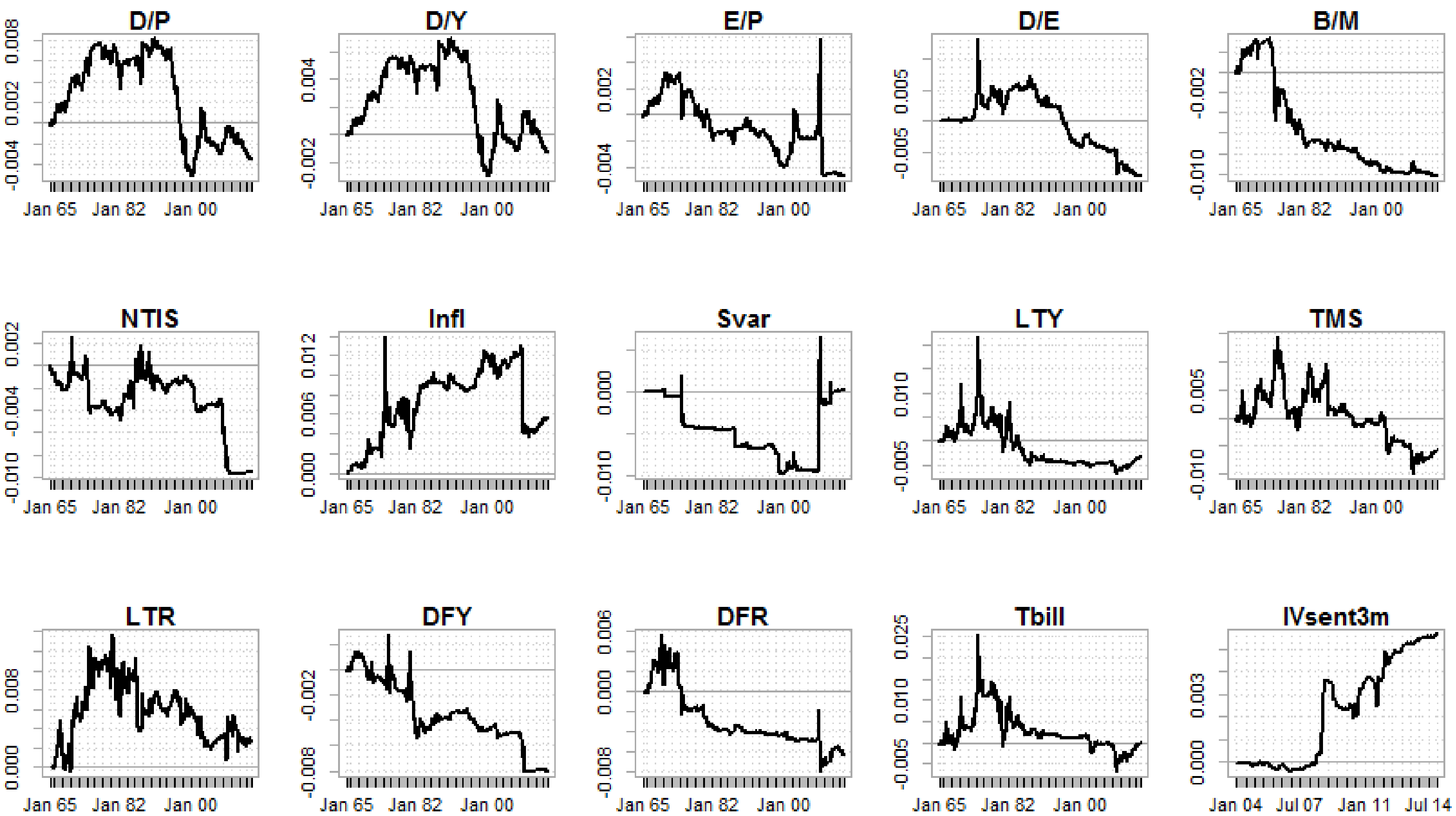
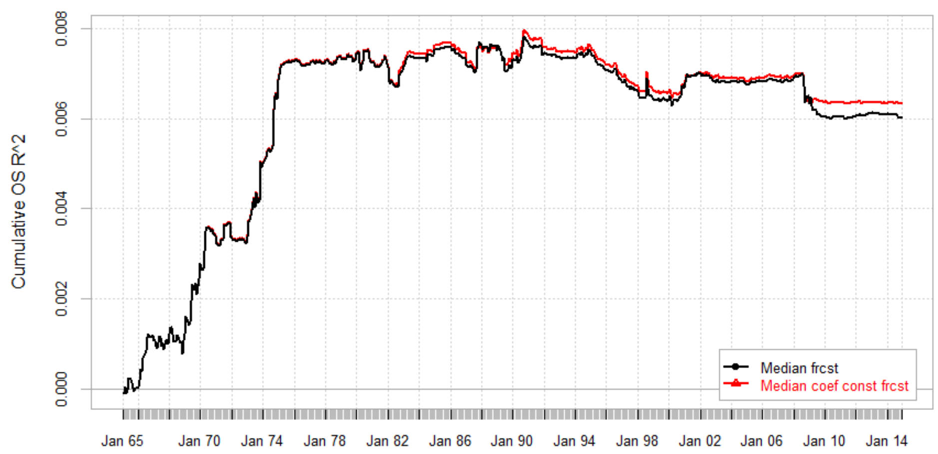
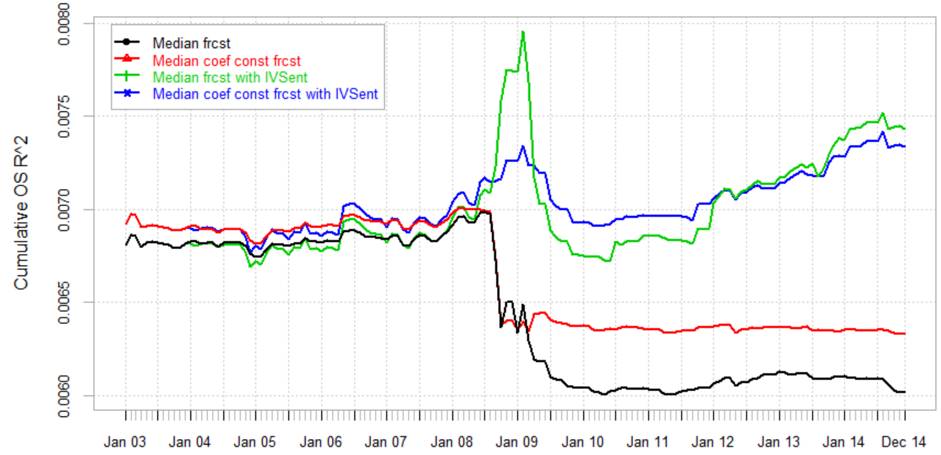


Figure 4: **Cumulative R^2_{OS} of single factor predictive regressions.** The lines in every plot depict the cumulative out-of-sample squared errors (R^2_{OS}) calculated by Eq. (24) for the historical average benchmark-forecasting model minus the cumulative squared prediction errors for the single-factor forecasting models constructed by using 14 out of all the explanatory variables suggested by Welch and Goyal (2008), as well as the *IV-sentiment 90-110* factor with a three-month maturity. Positive values of R^2_{OS} mean that single-factor forecasting models that employ the Welch and Goyal (2008) factors and *IVsent* outperform the historical average benchmark-forecasting model.



(a) Without *IV-sentiment*



(b) With *IV-sentiment*

Figure 5: **Cumulative R^2_{OS} of combined predictive regressions.** The black line in Plot A depicts the cumulative squared errors for the historical average benchmark-forecasting model minus the cumulative squared prediction errors for the aggregated predictive regression-forecasting model construct by using 14 Welch and Goyal (2008) explanatory variables in univariate unrestricted models. The green and red lines in Plot A depict the same forecast evaluation statistic, i.e., the cumulative R^2_{OS}, when such 14 univariate models are restricted as suggested by Campbell and Thompson (2008). The red line represents the cumulative R^2_{OS} when coefficients are constrained to have the same sign as the priors suggest. Plot B zooms in on the 2003:1-2014:12 period, where the black and red lines are the same as in Plot A, whereas the green and blue lines are the cumulative R^2_{OS} when our *IV-sentiment* factor is added to the multifactor forecasts model for the unrestricted and restricted model, respectively. The forecasting period is 1965:1-2014:12 for all variables except *IVSent*, for which forecasts are only available from 2004:1-2014:12. Forecast aggregation in both models is done by calculating the mean of the $t+1$ forecast from each individual predictive regression.