Market Maker Inventory, Bid-Ask Spreads, and the Computation of Option Implied Risk Measures

Bjørn Eraker Daniela Osterrieder

June 13, 2022

Abstract

We present empirical evidence showing that Option Implied Risk Measures (OIRMs) are substantially impacted by bid-ask spreads in underlying options. Asking prices are more sensitive to shocks than bids, leading to highly skewed distributions of spreads. We derive and estimate a model of market making that empirically matches these asymmetric responses as well as the time-series properties of bid-ask spreads. Using these estimates to obtain bias-corrected option quotes, we compute several popular OIRMs. We find that fear and risk premia associated with market events that affect the center of the return distribution or unpredictable return jumps are on average overstated when relying on option mid-quotes, whereas risk associated with return-tail events is larger once the bias has been corrected.

Keywords: asymmetric bid-ask spreads, biased option mid-quotes, market maker inventory, option implied risk measures

JEL codes: G12, G13, C58

Corresponding author: Bjørn Eraker, email: bjorn.eraker@wisc.edu, University of Wisconsin - Madison, 975 University Ave, 53706, Madison, WI.

We thank Mathieu Fournier, Nishad Kapadia, Dmitriy Muravyev, Andrea Vedolin, Ivan Shaliastovich, Mark Ready, and seminar participants at 11th International Conference on Computational and Financial Econometrics, Midwest Finance Association Meeting (2018), Norwegian Business School (BI), and OptionMetrics (2017).

1 Introduction

Since the pioneering work of Black and Scholes (1973) and Breeden and Litzenberger (1978), academics have understood that risk-neutral distributions of stock returns can be reverse engineered from option prices alone. This has led many researchers to compute option-implied measures of variance and higher order moments of stock returns. The most widely known option-implied risk measure (OIRM) is the volatility index, VIX, which is computed in near real time by the Chicago Board of Options Exchange (CBOE). Bakshi, Kapadia, and Madan (2003) derive a formula for computing option-implied risk-neutral skewness which has subsequently been applied by the CBOE to compute their skewness index (SKEW). Schneider and Trojani (2018) derive formulae for higher order moments while Bollerslev and Todorov (2011) and Bollerslev, Todorov, and Xu (2015) derive measures of jump risk based primarily on prices of deep out-of-the-money (OTM) index options.

Option prices are subject to microstructure noise. Bliss and Panigirtzoglou (2002) study option-implied density functions and find that they are heavily impacted by small perturbations in prices. Hentschel shows that implied volatility computations are sensitive to price errors, especially options with low strike prices. Dennis and Mayhew (2009) show that noise induces a bias in the estimates of risk-neutral moments, including implied volatility. Muravyev and Pearson (2020) point out that transactions usually take place inside of posted bid-ask quotes, effectively leading to a lower cost of trading than the usual half spread or effective spread estimates that are typically used to gauge trading costs. The authors further report that trades are mostly seller-initiated on average, as trade prices tend to be closer to bid quotes than ask quotes.

We contribute to this literature by studying the impact of market microstructure noise on the computation of OIRMs. First, we present reduced-form evidence on the determinants of option bid-ask spreads. Spreads are empirically negatively related to volatility changes (measured by the differenced realized variance), returns, and a month-end seasonal dummy. Spreads depend positively on concurrent volatility, on the expiration cycle of options, and on the interaction between inventory, vega (sensitivity to volatility changes), and local variance of the market makers' positions. Spreads also correlate positively with VIX levels. In particular, the spreads widen considerably during high-volatility periods such as the financial crisis 2008-2009.

To understand the relationship between bid-ask spreads, volatility, and inventories we derive a simple model of market-making behavior. The model features a market maker who will take a volatility position such as a delta-hedged option position that could, for instance, be a (synthetic) variance swap. Her total risk is then driven by inventory multiplied by the vega of the option position. The market maker posts bid and asking prices and faces an exogenous price-elastic demand. In line with the empirical findings in George and Longstaff (1993), the length between order arrivals depends on the size of bid and ask mark-ups, which are defined as the differences between the market maker's posted quote and an objective value of the option or option portfolio. In equilibrium, the market maker increases her asking price (bid price) in response to negative (positive) inventories. The size of the asymmetric adjustment is impacted by the volatility. Thus, the model delivers a (non-linear) relationship between market makers' inventory, volatility, and the bid-ask spread, similarly to Ho and Stoll (1981), although importantly it features time-varying spreads. We estimate the model using likelihood inference, which allows us to recover estimates of model parameters and the mark-ups on bid and ask prices charged by market makers. We find that objective, true values of options are likely to be closer to bids than asks. This finding is supported by Muravyev and Pearson (2020), presuming that trade prices are reflective of actual market value on average.

In order to investigate the impact of the market microstructure noise on the computation of OIRMs, we compute the VIX, the variance risk premium (VRP) of Bollerslev et al. (2009), the Fear Index (FI) of Bollerslev and Todorov (2011) and Bollerslev et al. (2015), the SKEW, and the associated skewness risk premium (SRP) of e.g. Sasaki (2016) using midpoints and bias-corrected option quotes that result from our model estimates, respectively. For the VIX and the VRP, the mean of the midpoint exceeds the bias-corrected computations by 3.7% and 25.3% in monthly variance units, respectively. In monthly volatility units, the latter difference is 0.299 or 11.9%. To put it into perspective, if an investor demands an extra 11.9% premium every month in order to hold the volatile security, she will collect an additional 286.2% volatility premium in expectation over a year. This in itself is a relatively large number. However, this number dwarfs in comparison to the asymmetry's relative impact on higher order measures. By computing our equivalent of the CBOE SKEW index, we show that when it is computed from bias-corrected prices, its standard deviation is more than ten times larger than when it is computed from midpoints. The FI differs substantially in both mean and standard deviation (as well as higher order moments) when computed from bias-corrected prices and mid-quotes. The differences in the SKEW and the FI computed from the two different quotes result from the fact that these indices are highly dependent on OTM option prices. Yet, OTM options are often quoted at minimum bids of 5 cents while the asking prices may be anything from 10 cents and up. This implies that OTM options, which are eligible to be included in the computation of the indices because they have a quoted bid, will mechanically have a minimum spread of 100%, and often are quoted with spreads of 10, 20, or even 30 cents.

The remainder of the paper is organized as follows. In the next section, we present reduced-form evidence pertaining to the asymmetry of bid and asking prices for options. Sections 3 and 4 present our theoretical model along with empirical estimates of model parameters. In Section 5 we examine the implications for OIRMs. Section 6 concludes.

2 Data and Asymmetric Spreads

We rely on daily data from OptionMetrics collected over the period 1998-01-02 to 2014-08-29. Since prices increase over time, raw bid and ask quotes are nonstationary. To facilitate the construction of a stationary time series of comparable bid and ask prices over time, we compute daily series

$$VIX^j = \sum_i w_i C_t(K_i)^j \tag{1}$$

for $j = \{a, b\}$ and where

$$w_i = \frac{1}{T}\frac{\triangle K_i}{K_i^2}e^{r(T-t)} \tag{2}$$

which is identical to the formula for computing the squared CBOE (squared) VIX index, with the exception of a small constant term[1]. Accordingly, we refer to VIX_t^a and VIX_t^b as measures of the average bid and ask prices. Note that the squared VIX index is, up to the dropped constant term, equal to $\frac{1}{2}(VIX_t^b + VIX_t^a)$. We let VIX_t^m denote the daily squared VIX variance series computed from mid-quotes. All three VIX measures are reported as a percentage variance figures scaled into monthly units.

[1] See the CBOE VIX white paper `https://cdn.cboe.com/resources/futures/vixwhite.pdf`.

We also use a second data set consisting of SPX options holdings data for firms and retail customers, distributed by the CBOE. This dataset has previously been used by Pan and Poteshman (2006), Lakonishok, Lee, Pearson, and Poteshman (2007), Chordia, Kurov, Muravyev, and Subrahmanyam (2018), and Chen, Joslin, and Ni (2018) among others.

The Option Clearing Corporation (OCC) groups each option transaction into one of three categories: public customer, firm proprietary trader, and market maker. The customer classification includes retail and institutional investors (e.g. hedge funds). Firms are broker-dealers that are not market makers and that trade for their own accounts, or for other firms. The CBOE data set contains only non-market maker records (Pan and Poteshman, 2006), and subclassifies the public customer transactions further into small, medium, and large customers. The data categorizes transactions on whether they comprise new positions, or close out existing positions. In particular, it distinguishes between OPEN BUY (an opening of a new long position), CLOSE SELL (a sale of an existing long position), OPEN SELL (a new short position), and CLOSE BUY (buy to cover an existing short position).

Our raw data set has two dimensions; trading days, t, and unique option contracts i (a unique contract is defined by the expiration date, the strike price, and whether the option is a put or a call). We track each unique contract i over its lifespan, that is from the first trading day in our sample (implicitly assuming that i did not exist before the first recorded trade) to the day before expiration.

Every day, t, within this period, we compute the market maker's holdings. More precisely, following Chen et al. (2018), we add the BUY positions and subtract the SELL positions from the previous day's holdings for every option contract and every investor type separately. Expiring contracts are nullified. As in Muravyev (2016), Chordia et al. (2018), and Gârleanu et al. (2009) the market maker takes the other side of the transaction, and hence her net inventory for contract i on day t, $n_{i,t}$, is the negative of the sum of firm and customer holdings. We also compute daily vega and squared vega for every contract within its lifespan.

[Figure 1 about here.]

The resulting series are shown in Figure 1. The plot shows the number of options held by market makers, n_t, defined as the sum of all $n_{i,t}$, and the total squared vega of their positions computed as

$n_t \times vega_t^2$. We use this definition of exposure, rather than $n_t \times vega_t$, as it corresponds to the relevant state variable in our model below. As seen, market makers are on average net sellers. This is in line with the results in Pan and Poteshman (2006), who confirm that between 1990 and 2001 index options we more actively bought than sold by non-market makers.

Figure 1 further shows that market makers' inventories gradually became more negative over the 1998 to 2007 period. In a perhaps prophetical and certainly fortunate rebalancing in early 2007, market makers reduced short positions dramatically. Chordia et al. (2018) note that the market makers' trading activity in index options positively predicts market returns. Since large financial institutions dominate market making, this is consistent with Hendershott, Livdan, and Schürhoff (2015) who provide empirical evidence suggesting that institutions, unlike retail investors, have an informational advantage in predicting stock returns.

We further collect several additional financial and economic series. Our risk-free rate series is the 3-month U.S. T-Bill rate, obtained from the FRED FED data base. We rely on the TAQ database to compute daily close-to-close SPY returns as well as realized variances (RV) from 5-minute intraday returns[2]. We are further interested in obtaining a measure for the conditional variance of the VIX. To obtain an estimate of $\mathrm{Var}_t(\triangle \mathrm{VIX}_{t+1})$ or equivalently $\mathrm{Var}_t(\mathrm{VIX}_{t+1})$, we estimate the following GARCH specification,

$$\mathrm{VIX}_{t+1} = a + b^{(d)}\mathrm{VIX}_t + b^{(w)}\frac{1}{5}\sum_{j=1}^{5}\mathrm{VIX}_{t-j+1} + b^{(m)}\frac{1}{22}\sum_{j=1}^{22}\mathrm{VIX}_{t-j+1} + \omega_t u_t \tag{3}$$

$$\omega_{t+1}^2 = \alpha_0 + \alpha_1 u_t^2 + \beta\omega_t^2 \tag{4}$$

The model for the dynamics of the squared volatility index in (3) is an additive cascade model inspired by the work of Corsi (2009). The parameters in this model are straightforwardly estimated using likelihood inference and we omit the details. Our estimate of $\mathrm{Var}_t(\mathrm{VIX}_{t+1})$ are the filtered ω_t^2s.

[Table 1 about here.]

The summary statistics of the data are in Table 1. Three stylized facts are confirmed by the data. First, asking prices are more volatile than the bids. Table 1 shows that the variance of the

[2] OptionMetrics data is available from 1996 onwards. We discard the first two years in compiling our data set, since the observations on the SPY in TAQ are not sufficiently liquid to compute measures (e.g. RV) that rely on high-frequency intraday returns.

indices computed from bid quotes is between 71.5% and 96.9% of the variance of the indices obtained from asks[3]. Second, market makers' inventories are negative on average; the sample mean of n_t in Table 1 is -30$\times 10^4$. Lastly, all three spread measures exhibit a strong positive skew. Returns are zero on average, very volatile, negatively skewed, and - as usual - leptokurtic. All variance series (VIX, RV, and GARCH-model forecasts) have a strong positive skew and a very high kurtosis. Finally, we confirm that the risk-free rate has been low throughout our sample period, with a minimum of zero and a maximum of 6.3% annually.

2.1 Empirical Evidence on Spread Asymmetry

In the following we present reduced form descriptive evidence on option bid-ask spreads.

[Figure 2 about here.]

Figure 2 shows our computed VIX^m and VIX^b indices on the first trading day of every month, as well as the VIX-spread $VIX^a - VIX^b$. It is clear that the two follow each other very closely in levels. There's also an obvious correlation between the level of the VIX-spread and the level of VIX.

[Figure 3 about here.]

Figure 3 shows the average spread in ATM option prices computed from daily data. The upper panel, which shows the raw time series, displays what appears to be rapid, periodic movements, along with longer term variation which is related, at least in part, to the level of volatility. The periodic variation is shown in the bottom graph by plotting the sub-period of 2003 along with end-of-month and option-maturity indicators. As seen from this plot, there is a clear tendency for the spread to contract on the last day of the month and to increase on the maturity dates.

[Table 2 about here.]

[Table 3 about here.]

[3]The lower variability of VIX^b relative to VIX^a cannot be attributed to the fact that on each date, options with bids of zero are discarded before the VIX for that day is computed. If instead, we also include bid prices ≤ 0 in the index computation, we find standard deviations of 41.04 (51.04) for VIX^b (VIX^a). These numbers are basically the same as in Table 1.

Table 2 shows the results from running regressions of the VIX spread, $VIX^a - VIX^b$, onto various explanatory variables. A contemporaneous negative S&P 500 return ("Return") positively impacts the spread and vice versa. Not surprisingly, higher levels of realized volatility ("RV level") are associated with higher VIX spreads. As already seen in Figure 3, expiration and month-end dummies correlate significantly with the spread. The regressions also include the variable $|\text{Inv}\times vega^2 \times RV|$, which measures the aggregate market maker vega position multiplied realized volatility. This variable is theoretically motivated from our model to be introduced below. As seen, this variable strongly and significantly impacts the VIX spread.

Table 3 presents the equivalent results of regression of VIX^a and VIX^b onto the same variables as in Table 2. The main takeaway here is that asking prices tend to move more than bids in response to shocks to the explanatory variables. For example, a negative (positive) contemporaneous S&P return will have a larger positive (negative) impact on asking prices, than bids. The same is true for the other explanatory variables. To check whether this difference in impacts is significant, we conduct a Wald test for the equality of the estimated effect of the variable of interest, $|\text{Inv}\times vega^2 \times RV|$, on VIX^a and VIX^b. In the first regression specification (column 1 of Table 3) we find a Wald statistic of 12.7739 and a corresponding p-value of 0.0004 from the $\chi^2(1)$-distribution. Similarly, in the fourth and fifth regression specification (columns 4 and 5 of Table 3) we find Wald statistics of 12.4262 and 12.2179, respectively, with corresponding p-values of 0.0004 and 0.0005. Thus, we reject the equality of the impact in each case.

To further interpret these estimates, consider the impact of a shock to market makers' inventory. If the variable $vega^2 \times RV$ is at its steady state average, and if the absolute value of inventory were to increase by an amount equal to its unconditional standard deviation, the increase in the VIX squared spread is 2.9 annualized variance units. Since the average spread is 7.5, this increase represents a 38% increase in spread. At the lower 10th percentile of $vega^2 \times RV$ the change in the spread is 0.38 in response to a one standard deviation increase in inventory. The comparable number is 6.21 when evaluated at the 90th percentile of $vega^2 \times RV$.

Another way to study this impact is to imagine that the inventory, in absolute value, were to increase by an amount equal to the standard deviation of its daily changes. This is a much smaller quantity than the unconditional standard deviation. Computations then show that the spread would increase by 0.51, or 6.8% on average. Both these numbers point to a large, economically significant impact of inventory on spreads.

3 Model

In the following we derive a simple model that expresses the mark-ups charged by market makers for buy and sell orders, respectively.

The mark-ups, ϵ_t^j for $j = \{a, b\}$ are defined as

$$C_t^{\text{ask}} \equiv C_t + \epsilon_t^a \tag{5}$$

$$C_t^{\text{bid}} \equiv C_t - \epsilon_t^b. \tag{6}$$

Thus, if C_t^{ask} is the observed market asking price, $C_t^{\text{ask}} - \epsilon_t^a$ is an estimate of the "true value" of the option. The true value of the option is unobserved, but by this assumption it is contained within the bid-ask spread. Since the markups are assumed positive, market makers are unwilling to sell (buy) at a price below (above) the true value. This is reasonable in options markets given the existence of hard arbitrage bounds on prices as well as a well-developed theory of option valuation.

Assume that a market maker has a position $\bar{n}_t = \{n_{t,i}\}$ for $i = 1, .., N_t$ in options. Let P_t denote the prices at the start of a trading period $[t, t+1]$, taken to be one day. The market value of the options is then $\bar{n}_t' P_t$ and the total change in the value of the positions is $\bar{n}_t'(P_{t+1} - P_t)$ over the course of a day.

Assume further that the option prices are functions of the underlying stock price level, S_t, and the VIX index, such that $P_t = P(t, S_t, VIX_t)$. This is equivalent to assuming that the option price depends on the stock price and its spot variance, say σ_t^2, under mild regularity conditions. The total delta of the market maker's position is then

$$\delta_{p,t} = \bar{n}_t' \frac{\partial P_t}{\partial S_t}, \tag{7}$$

where $\frac{\partial P_t}{\partial S_t}$ is a vector of partial derivatives (deltas). The vega of the portfolio is

$$vega_{p,t} = \bar{n}_t' \frac{\partial P_t}{\partial VIX_t}. \tag{8}$$

We assume that S_t and VIX_t are diffusion processes. Then the value of the option portfolio evolves according to

$$dV_t = \mu(\cdot)dt + \delta_{p,t}dS_t + vega_{p,t}dVIX_t, \tag{9}$$

where $\mu(\cdot)$ is a drift term, and dV_t, dS_t and $dVIX_t$ are the infinitesimal increments to the unhedged options portfolio, underlying stock (or stock index) and VIX index, respectively.

There is no reason for the market maker to take on delta risk, and most do not (see e.g. Jameson and Wilhelm, 1992, and Christoffersen, Goyenko, Jacobs, and Karoui, 2017). Following Stoikov and Saglam (2009) among many others, we therefore assume that the position is fully delta hedged. Hence, the market maker takes an offsetting position $-\delta_{p,t}$ in stock such that the dS_t-term in (9) cancels. The delta-hedged portfolio value, V_t^H, then evolves according to

$$dV_t^H = \mu(\cdot)dt + vega_{p,t}dVIX_t. \tag{10}$$

Instead of modeling bid-ask spreads across an entire spectrum of all possible options contracts, we simplify the problem and study the average bid-ask spreads on a portfolio of options. Let $C_t = \sum_i w_{i,t}P_{i,t}$ with weights $w_{i,t}$ denote the price of this option portfolio.

Below we derive a model for the bid-ask spread of C_t. If the weights $w_{i,t}$ match the composition of the VIX index, as in equation (2), we can interpret C_t as a synthetic variance swap (see for example Eraker and Wu (2017) and references therein).

A market maker trading variance swaps faces a discrete-time profit given approximately by

$$\pi_{t+1} = n_t\mu + n_t vega_{vs,t} \triangle VIX_{t+1}, \tag{11}$$

where n_t now denotes the number of variance swap contracts the market maker holds. The variance of the market maker's profit is thus

$$V(n_t) := \mathrm{Var}_t(\triangle\pi_{t+1}) = {n_t}^2 vega^2_{vs,t}\mathrm{Var}_t(\triangle VIX_{t+1}), \tag{12}$$

where $vega_{vs,t}$ represents the vega of a synthetic variance swap.

We assume that option market orders arrive with Poisson intensities that depend on the price, or the mark-up of the market makers. The greater the mark-ups, the lower the demand, and thus there are fewer arrivals of market orders. Specifically, we assume

$$\lambda(\epsilon_t^j) = A_j e^{-D_j \epsilon_t^j} \tag{13}$$

for bids and asks, $j = \{a, b\}$. The parameters A_j and D_j are assumed to be positive such that the arrival intensities are positive, and negatively dependent upon the size of the mark-ups. Positive D_j's create a standard downward sloping "demand function" where arrivals increase as ϵ_t decreases. The market maker controls her inventory probabilistically: For example, if she has a large negative inventory, she might increase ϵ_t^a so as to deter additional sell orders and simultaneously decrease ϵ_t^b, so as to attract buy orders. We assume that zero is a lower bound for ϵ_t. That is, the market maker will not post a limit buy (limit sell) above (below) the "true" price C_t. We can weaken this assumption by assuming that there exists a lower no-arbitrage bound on the value of the option, say $C_t^{lb} \leq C_t$. For example, the C_t^{lb} could be the intrinsic value, or the market equivalent value of a close substitute, for example options on another closely related index, or a VIX futures position. The $\epsilon_t > 0$ assumption is also made in Stoikov and Saglam (2009) and Chan and Chung (2012).

Consider a market maker who posts bids and asks for x variance swap contracts. If the market maker's bid is hit by a market order his inventory becomes $n_t + x$. The conditional variance is

$$\mathrm{Var}_t(\epsilon_t^a, \epsilon_t^b) = \lambda(\epsilon_t^b)V(n_t + x) + \left(1 - \lambda(\epsilon_t^b)\right)V(n_t) + \lambda(\epsilon_t^a)V(n_t - x) + (1 - \lambda(\epsilon_t^a))\,V(n_t). \tag{14}$$

The model features a monopolistic market maker, which is a realistic assumption when studying SPX options, since they "have only a single market maker who provides continuous quotes" (Chordia et al., 2018). We assume that the market maker is risk-averse and she determines her mark-ups, ϵ_t^a

and ϵ_t^b , so as to maximize a quadratic utility function

$$\max_{\epsilon_t^a,\epsilon_t^b} \lambda(\epsilon_t^b)\epsilon_t^b x + \lambda(\epsilon_t^a)\epsilon_t^a x - \gamma \mathrm{Var}_t(\epsilon_t^a, \epsilon_t^b), \tag{15}$$

where γ is a risk aversion coefficient. The terms $\lambda(\epsilon_t^j)\epsilon_t^j x$ represent the expected profit from a single purchase or sale. The market clears through market orders executed at the market maker's bid or ask, C_t^{bid} and C_t^{ask} in (5) -(6).

Proposition 1. *The optimal mark-ups are given by*

$$\begin{aligned}
\epsilon_t^a &= \max\left(\frac{1}{D_a} + \gamma\,(-2n_t + x)\, vega_{vs,t}^2\, Var_t(\triangle VIX_{t+1}), 0\right) && (16)\\
\epsilon_t^b &= \max\left(\frac{1}{D_b} + \gamma\,(2n_t + x)\, vega_{vs,t}^2\, Var_t(\triangle VIX_{t+1}), 0\right) && (17)
\end{aligned}$$

4 Estimation and Inference in the Structural Model

Our model is essentially one that applies to a single asset. In order to apply it to options we compute aggregate options data as described in Section 2. We compare these aggregate quantities across all strikes and maturities to aggregate bid-ask spreads as measured by the difference in VIX computed from asks and bids.

4.1 Identification

Our estimation method is straightforward. Let $Z_t^{\mathrm{theory}}(\Theta) = \epsilon_t^a + \epsilon_t^b$ denote the theoretical spread and let Z_t^{data} denote some measure of observed spread in the data. We assume that

$$Z_t^{\mathrm{data}} - Z_t^{\mathrm{theory}}(\Theta) \sim N(0, s^2). \tag{18}$$

Empirically, the distribution of spreads is asymmetric and has fat tails, which may translate to the residuals $Z_t^{\mathrm{data}} - Z_t^{\mathrm{theory}}(\Theta)$, as well. For robustness, we thus also consider the alternative assumption

$$Z_t^{\mathrm{data}} - Z_t^{\mathrm{theory}}(\Theta) \sim GH(0, s^2, \upsilon, \zeta, \rho), \tag{19}$$

where GH is the Generalized Hyperbolic distribution with the parameterization used in Bensaïda and Slim (2016). This parameterization has the advantage that the first two parameters of the distribution determine the mean and variance, while remaining three values determine the tails and skewness of the distribution. With the mean of the distribution set to zero, it can be used to specify a non-Gaussian likelihood function for the non-linear regression implied by (19). The parameter s^2 is interpretable as a residual variance, comparable to the specification with normal errors (18).

The distributional assumptions in (18) or (19) allow us to specify a likelihood function, $\mathcal{L}(\Theta)$. To facilitate numerical stability, we estimate the re-parameterized vectors $\Theta^{(N)} = \{D_a^{-1},\, D_b^{-1},\, \gamma,\, x,\, s\}$ and $\Theta^{(GH)} = \{D_a^{-1},\, D_b^{-1},\, \gamma,\, x,\, s,\, \upsilon,\, \zeta,\, \rho\}$. That is, we estimate the reciprocal of the demand elasticity parameters, D_a and D_b.

There are a number of other minor details to note in regards to the parameter estimation and identification. First, our estimator does not allow us to identify the average arrivals, as determined by the coefficients, A_a and A_b. On the other hand, the parameters $D_a^{-1}, D_b^{-1}, \gamma$ and x, are identified up to a constant. To see this, note that if $\hat{\Theta}^{(N)} = \left\{\hat{D_a^{-1}}, \hat{D_b^{-1}}, \hat{\gamma}, \hat{x}, \hat{s}\right\}$ is an estimator of $\Theta^{(N)}$ then rescaling the data $\tilde{n}_t = \kappa n_t$ by a positive constant κ, leads to the MLE $\tilde{\Theta}^{(N)} = \left\{\frac{1}{\kappa}\hat{D_a^{-1}}, \frac{1}{\kappa}\hat{D_b^{-1}}, \frac{1}{\kappa}\hat{\gamma}, \frac{1}{\kappa}\hat{x}, \hat{s}\right\}$. Note that $\hat{s}$ is unaffected by rescaling, implying that the $R^2 = 1 - \hat{s}^2/\mathrm{Var}(Z_t^{\mathrm{data}})$ is unaffected as well. This is, of course, standard for linear regression models. Unlike linear regressions, however, parameter estimates in our model are not invariant to the addition of a constant because the identification of D_a and D_b depends on the sign of n_t.

It should further be noted that the parameters D_a^{-1} and D_b^{-1} are identified from the changing sign of inventory[4]. If observed inventory were negative throughout, we would not be able to identify D_b^{-1}; conversely if $n_t > 0$ for all t, D_a^{-1} would not be identifiable. Fortunately, we observe periods with positive as well as negative inventory in our data, although for the majority of our sample period market maker inventories are negative.

Finally, rather than estimating the expected order size x from spread data, we set it equal to the average order size $\bar{x} = \frac{1}{T}\sum_t |n_t - n_{t-1}|$. With this, we are able to interpret the estimate as an actual average order size, rather than just a free parameter.

[4]See the breakpoints shown in Figure 9 in Appendix 2

4.2 Parameter Estimates

[Table 4 about here.]

Table 4 reports parameter estimates for the the equilibrium model. We report posterior means and standard deviations to summarize the Bayesian analysis, and maximum likelihood point estimates. The two methods give almost identical numerical results, indicating that the posterior distributions are relatively symmetrically distributed with modes close to the MLE. The parameter estimates for Equations (18) and (19) are also quite close. The bid-ask asymmetry is marginally more pronounced, the risk aversion is somewhat lower, and the error variance elevated when errors are not Gaussian. The in-sample fit, as measured by R^2, of the predicted spread resulting from Equation (19) is slightly lower than the corresponding fit from (18), which is why we opt to work with the latter estimates henceforth.

[Figure 4 about here.]

The most interesting aspect of the parameter estimates is the large difference in the elasticity parameters D_a^{-1} and D_b^{-1}. To see what these parameter estimates imply, Figure 4 plots the estimated ask and bid mark-ups. As seen, the market makers' mark-up is much greater on the asking side. In fact, the mark-ups for the bids are only non-zero during the height of the financial crisis, and zero otherwise. This implies that, for the most part, market makers are willing to buy at the "true" price with no mark-up. To sell, they require a premium that is time varying, especially during stress periods.

[Figure 5 about here.]

The implication that the average ask mark-up, ϵ_t^a, exceeds the average of ϵ_t^b suggests that that the midpoint, $\frac{1}{2}(C_t^{\text{ask}} + C_t^{\text{bid}})$, is upwardly biased for the true price, C_t. In fact, the model implies that the midpoint is never an unbiased estimate of the true value of the option, since $\epsilon_t^a \neq \epsilon_t^b$ at every point in the sample. Conditionally, the bias in the midpoints is substantial, as large as half the spread itself for most t. The results also imply that the second moment of ϵ_t^a is greater than the variance of ϵ_t^b, which means that there are economic shocks that asymmetrically impact asking prices.

Our model provides a reasonable, but not perfect, fit to the data and the R^2 is reported to be 39% in Table 4. Figure 5 plots the data vs. the theoretical spread based on the MLE estimates. The model captures the increases in the spread observed during stress periods, as for example during the financial crisis, reasonably well. There are, however, some periods during which the model fails to capture the low observed spreads. During the late nineties and early two-thousands, the model does well in capturing the average spreads. For a long stretch in the mid two-thousands and also after mid 2012, the model produces a spread that is too large. It is easy to see where this is coming from: In the absence of a large positive or negative inventory, the model produces a spread that is essentially equal to D_a^{-1}. The model hits this lower bound whenever market maker inventory falls within a range, as seen in Figure 9.

As is necessarily the case with any model, our theoretical model that describes spreads as a function of market maker's inventory, price sensitivity to changes in the volatility of the underlying asset, and the conditional variance of the spot variance, is an abstraction of reality. Other drivers of spread asymmetries that have been suggested in the extant literature are, for instance, the risk of trading with traders that have an informational advantage, or the compensation required by liquidity providers for facing possibly sharp changes in asset prices or tail risk (see Weller, 2019). Standard errors of model estimates account for potentially omitted factors, in the sense that they increase if the variability of the unexplained part of the model increases. Thus, Figures 4 and 5 report confidence intervals that are based on these standard errors[5], in an attempt to caution against uncertainty. Throughout the remainder of this paper, we will report standard errors for our model-implied measures that are based on the estimation in this section as a reminder to the reader that there is statistically uncertainty.

5 Impact on OIRMs

The model in the previous section suggests that option midpoints are biased estimates for the "true" price, which in turn seems to lie closer to the bid quote. We now examine how this bias impacts OIRMs such as VIX, VRP, FI, which is an option implied jump-risk measure of Bollerslev et al. (2015), and on the SKEW index from the Chicago Board of Options Exchange, as well as the corresponding SRP. These risk measures are of interest for finance academics and practitioners because they are believed to capture risk or risk aversion in financial markets.

[5] These standard errors, as well as the ones reported for bias-corrected OIRMs below, are computed by finite differences and the Delta Method.

Relying on our MLE model-estimation results for Equation (18), we can correct OIRMs for the asymmetry in spreads. Specifically, we compute the OIRMs from $\hat{C}_t$, i.e. our estimate for the "true value" or "theoretical price" of the option. We refer to these OIRMs as bias-corrected prices, denoted $(\cdot)^{BC}$. To obtain the bias-corrected prices, we merge the OptionMetrics database with the inventory database. Only contracts that are listed in both data sets are retained[6]. We then use the parameter estimates from Section 4 to calculate the bias-corrected option price as

$$\hat{C}_t = C_t^{\text{bid}} + \hat{\epsilon}_t^b. \tag{20}$$

C_t^{bid} is the observed bid quote and $\hat{\epsilon}_t^b$ is a function of the parameter estimates, the option's inventory and vega, and the conditional variance of VIX (see Eq. (16)-(17)). As an alternative, we could have computed bias-corrected prices as $\hat{C}_t = C_t^{\text{ask}} - \hat{\epsilon}_t^a$, which has the disadvantage that prices can become negative. The correlation between the two bias-corrected prices is 0.99, which suggests that it does not impact results largely, which alternative we chose.

5.1 VIX and Variance Risk Premium

We compute our own VIX series based on midpoints of the bid and the ask, VIX^m, following the methodology for computing the CBOE VIX index. The VIX, which is a proxy for the standard deviation of the S&P 500 return distribution, is thus subject to market events that affect the center of the distribution rather than the far tails. Table 5 presents summary statistics and shows that our series is basically identical to the index published by the CBOE. Then we apply the same methodology to obtain VIX^{BC}.

The average daily VIX computed from $\hat{C}_t$ is 43.243, and the average of VIX^m is 44.842, measured as a percentage variance scaled into monthly units. These numbers correspond to 22.937 and 23.358 in the more commonly reported annualized percentage standard deviation units. To test whether these two values are statistically different, we can rely on the standard error of the former estimate, which is obtained from the model estimation in Section 4. The average of VIX^m is not contained in the 90% and the 95% confidence interval of the average of VIX^{BC}, yet it lies within the 99% confidence interval.

[6]The process of merging the two data sets reduces the number of option contracts from 5,588,744 (in the original OptionMetrics database) to 4,383,052.

As second approach to testing the equality of the OIRM-moments, we re-estimate the first moment by nonlinear GMM, restricting the averages of the two series, VIX^m and VIX^{BC} to be identical[7]. The resulting GMM J-statistic has a $\chi^2(1)$ distribution if the restriction holds. As can be seen in Table 5, we strongly reject the equality of means. Thus, the mean of the VIX Index is overestimated when using option midpoints by 3.698% in monthly variance units or 1.832% in volatility units. Put differently, selling short-term volatility at a 1.832% premium every month would seemingly imply an additional 24.341% return premium in expectation over one year. This is a substantial numerical difference, and it does mean that in reality investors perceive the risk associated with the center of the return distribution as less severe. This finding my also be interpreted as saying that investors are less averse to risk on average. This interpretation of the results draws on Whaley's (2000) description of the VIX as *investor fear gauge*, which makes it a popular indicator of aggregate risk aversion according to Bekaert, Hoerova, and Duca (2013).

Higher-order moments of VIX^{BC} are smaller than the corresponding moments of VIX^m, as well. The sample standard deviations over the entire time period are 46.056 for VIX^m and 41.872 for VIX^{BC}. The model-implied 99%-confidence intervals for the standard deviation of VIX^{BC} do not contain the corresponding estimate for VIX^m, suggesting a rejection of equality of second moments. Yet, if we once again rely on the J-statistic instead, this time having estimated the restricted second moments by GMM, we fail to reject a common variance. Similarly, VIX^{BC} is less skewed and less leptokurtic than VIX^m. Both differences in skewness and kurtosis are strongly significant based on the model-implied standard errors, but only the difference in the skewness is marginally significant based on the J-statistic. The strong persistence measured by the first-order autocorrelation coefficient, AR(1), is the same for the two series.

[Table 5 about here.]

From the two VIX series, we can compute the VRP of Bollerslev et al. (2009) by subtracting the realized variance over the past 22 days (including the overnight squared return). We denote the resulting series by VRP^m and VRP^{BC}, and report them as percentage variances in monthly units. Table 5 presents the corresponding summary statistics. The average daily VRP^{BC} is 6.299 (8.754 in

[7]The long-run variance covariance matrix of the moment conditions is estimated as a HAC estimator, relying on the Bartlett kernel, and using Newey-Wests automatic lag selection criterion.

annual percentage volatility units), whereas the average of VRP^m is larger, equal to 7.890 (9.798 in annual percentage volatility units). If we rely on the same tests for the equality of means as above for VRP, we find a J-statistic of 14.038 and a corresponding p-value of 0.000. We clearly reject that VRP^m and VRP^{BC} have the same mean. Also, the model-implied standard errors suggest that the two averages are numerically different, at least with 90% confidence. In fact, numerically the former exceeds the latter by more than 25%. Thus, if we view VRP as an estimate for the actual variance premium in the market, then the on average positive payoff for the seller of a variance swap contract is substantially smaller than the measure based on midpoints would suggest.

In contrast, the sample standard deviations of the two series are almost the same: 23.173 and 24.586 for VRP^m and VRP^{BC}, respectively, and the GMM-test fails to reject the equality. Thus, while the bias-correction does not affect the variability of VRP much, it does produce a series VRP^{BC} that is more negatively skewed (-4.649) than VRP^m (-3.431), strongly significantly so according to the GMM-test as well as model-implied inference.

[Figure 6 about here.]

5.2 Fear Index

While the differences in VIX and VRP computed from midpoints and bias-corrected quotes are already very sizable, they are dramatic for OIRM series that are highly dependent on OTM option prices. One such example is the Fear Index (FI).

Bollerslev and Todorov (2011) show that the (*ex-ante* negative) VRP can be decomposed into two components: a compensation for bearing continuous or diffusive risk, and a premium for jump risk. Bollerslev et al. (2015) suggest that the known predictability of returns from VRP is mostly driven by the latter jump-risk factor. In doing so, they construct the FI, defined as the left-jump premium (LJP) minus the right-jump premium (RJP). To align the concept of a premium with our definition of VRP in the previous subsection and SRP in the following subsection, we depart from Bollerslev et al. (2015) and specify a premium as the expectation under the $\mathbb{Q}$-measure minus the expectations under the $\mathbb{P}$-measure. Following the derivations in Bollerslev and Todorov (2011) and Bollerslev et al. (2015), it holds that

$$FI \equiv LJP_{t,\tau} - RJP_{t,\tau} \approx \frac{1}{\tau}\,\mathbb{E}_t^{\mathbb{Q}}(LJV_{t,\tau}^{\mathbb{Q}}) - \frac{1}{\tau}\,\mathbb{E}_t^{\mathbb{Q}}(RJV_{t,\tau}^{\mathbb{Q}}), \tag{21}$$

where RJV (LJV) is the left (right) jump variation over horizon τ. The index is positive, which follows from the empirical fact that LJV is bigger than RJV for all our observations[8].

The FI is computed weekly[9]. We report the corresponding summary statistics in Table 5. As can be seen, our estimates for the computation of FI^m, that is the fear index computed from mid-quotes, are reasonably close to the ones in Bollerslev et al. (2015) for their 1996-2013 sample. The only visible differences are in the RJV estimates, but this is negligible since the right tail has next to no impact in the computation of FI[10].

To gauge the impact of the bid-ask asymmetry on the computation of the Bollerslev and Todorov (2011) FI, we further compute a version of the index based on bias-corrected quotes and refer to it as FI^{BC}. That is, we re-compute the tail-shape parameters and the jump variations using $\hat{C}_t$ as input to obtain FI^{BC}. Table 5 presents summary statistics of the resulting time series. As we can see, there is a large difference in the behavior of the fear index based on mid-quotes and $\hat{C}_t$. The average weekly FI computed from $\hat{C}_t$ is 2.914, and the averages of FI^m is 6.113. On average, FI^m thus overstates *fear* by 109.8% when measured in annualized percentage variance units. The sample standard deviations are 6.863 for FI^m and 3.856 for FI^{BC}. The difference in both, first and second moment, are strongly significant according to the GMM J-test and the model-implied confidence intervals. Thus, the unconditional mean and standard deviation of FI^m are both almost twice the magnitude of FI^{BC}. FI^{BC} is also less skewed, less leptokurtic, and has less persistence than FI^m, where the difference in skewness and the AR(1) parameter are marginally significant based on the GMM-test. In contrast, the model-implied standard errors suggest that we reject the equality of these

[8]Since we reversed the definition of a premium, our FI will have the opposite sign of the index in Bollerslev and Todorov (2011) and Bollerslev et al. (2015). Furthermore, in contrast to the original estimates for the FI in Bollerslev and Todorov (2011), Bollerslev et al. (2015) allow for the shape of the jump tails to be time varying. We follow their estimation approach.

[9]The estimation methodology for FI requires the choice of two tuning parameters. We opt for the same parameters as in Bollerslev et al. (2015). Firstly, to compute the left (right) jump tail parameters we only use put (call) options with log-moneyness less than 2.5 (larger than 1) $\times$ the maturity-normalized implied volatility. Secondly, to identify "large" jumps, we use a time-varying cutoff point that is equal to 6.8686 $\times$ the maturity-normalized implied volatility.

[10]In fact, "[...] for the aggregate market portfolio the magnitude of the risk-neutral left jump tail dwarfs that of the right jump tail, so that empirically $LJP_{t,\tau} - RJP_{t,\tau}$ is approximately equal to the $\mathbb{Q}$ expectation of the negative left jump variation only." (Bollerslev et al., 2015)

higher-order moments at a 99% confidence level.

Figure 7 plots the time series and reveals the time-series sensitivity to the asymmetry in the bids and asks. The difference is enormous. The midquote and bias-corrected series are also not very highly correlated, with a correlation coefficient of 0.725 (with a model-implied standard error of 0.002). In conclusion, the fear of an investor in the aggregate market portfolio of rare unpredictable jump events in returns is much more moderate when relying on bias-corrected prices.

[Figure 7 about here.]

In Table 6, we regress FI^m on measures of bid-ask spreads. The first variables are the VIX and the VIX spread ("Spread") defined as $VIX^a - VIX^b$. As seen, these measures explains about 29% of the variation in FI. We also include other spread measures, including the fitted spread $\epsilon_t^a + \epsilon_t^b$ from our model, and the individual estimated mark-ups ϵ_t^a and ϵ_t^b. The table reveals that the various spread measures significantly correlate with the Fear Index computed from midpoints, at least for subsets of the regressions. Moreover, the the dramatic difference in FI^m and FI^{BC} shown in Figure 7 suggests that the FI computation is very sensitive to the bid-ask spreads of OTM puts and calls.

[Table 6 about here.]

5.3 Implied Skewness

The CBOE SKEW Index is a measure for the perceived tail risk in the aggregate market, specifically in the S&P 500 market[11]. Thus, whereas the VIX is heavily impacted by the center of the return distribution, the SKEW explicitly considers the left tail of the return distribution. Tail risk, as defined by the CBOE, is the risk that large negative returns in excess of two standard deviations can occur. Several stock market crashes in the history of the S&P 500 suggest that this risk is substantial.

The SKEW is an option-implied risk measure. More precisely, it is computed from OTM options on the SPX that have 30 days time to maturity. The Index, $SKEW$, is typically larger than 100. If its value is exactly 100, then the risk-neutral expected distribution of returns on the S&P 500 is log-normal. That is, the conditional probability under the equivalent martingale measure of a large

[11] See the CBOE SKEW white paper `https://cdn.cboe.com/resources/indices/documents/SKEWwhitepaperjan2011.pdf`.

negative return is very small. As $SKEW$ increases above the mark of 100, the expected risk of a tail event within the next month increases.

We compute the index from mid-quotes, $SKEW^m$, on a daily basis, following the methodology outlined in CBOE's SKEW White Paper. By the same approach, we also compute $SKEW^{BC}$, the index based on bias-corrected prices $\hat{C}_t$. Table 5 presents summary statistics of the two time series. When we compute the index from mid-quotes it matches the original CBOE SKEW Index very closely, suggesting that we can accurately replicate the index. As we can further see, there is a substantial difference in the properties of the SKEW based on midpoints and bias-corrected prices. As opposed to our findings for the VIX, however, the moments of $SKEW^{BC}$ are all more extreme than the corresponding moments of $SKEW^m$. The unconditional means are 118.585 of $SKEW^m$ vs. 149.907 of $SKEW^{BC}$, and the difference is strongly significant according to the J-statistics and the 99% model-implied confidence intervals. More strikingly, the standard deviation of $SKEW^m$ is only about 9% of the magnitude of $SKEW^{BC}$. Again, the difference between the moments is strongly significant as suggested by both statistical inference approaches. $SKEW^{BC}$ is more skewed and has a much higher kurtosis than $SKEW^m$, and we reject the equality of the moments at a 1% significance level with the GMM-test as well as the model-implied confidence intervals. Lastly, $SKEW^{BC}$ is even marginally more persistent that $SKEW^m$, but this difference is not significant.

The dissimilarity of the $SKEW^{BC}$ and $SKEW^m$ series over time is obvious. The sample correlation between the two skewness indices is only 0.35, with a model implied standard error of 0.006, again suggesting that the series are vastly different. Thus, the probability of outlier returns relative to a log-normal distribution is substantially larger on average than $SKEW^m$ would suggest, and also more volatile. Perceived tail risk, further, is positively skewed, strongly leptokurtic, and persistent.

Analogously to the previous section, we regress $SKEW^m$ on measures of bid-ask spreads. The results are presented in Table 7. When regressing the SKEW computed from midpoints on the VIX and $VIX^a - VIX^b$ ("Spread"), we find that these explain roughly 10% of the variation in SKEW. When using alternative proxies for the spread, such as the fitted spread $\epsilon_t^a + \epsilon_t^b$ from our model, and the individual estimated mark-ups ϵ_t^a and ϵ_t^b, we find that these also correlate significantly with $SKEW^m$. In short, bid-ask spreads seem to impact our OIRM computed from midpoints.

[Table 7 about here.]

Next, we compute the skewness risk premium, SRP. SRP is assumed to be a proxy for the expected return required by an investor to compensate her for the uncertainty in the third moment of the return distribution. To that end, we obtain an estimate for the realized skewness over the last month. More precisely, we compute our measure as

$$RSKEW \equiv \frac{\sqrt{22 \times 78} \sum_{i=0}^{21} \sum_{j=1}^{78} r_{t-i,j}^3}{\left(\sum_{i=0}^{21} \sum_{j=1}^{78} r_{t-i,j}^2\right)^{3/2}},$$

where $r_{t,j}$ denotes that jth 5-minute log return on day t (including the overnight return). This measure is similar to Amaya, Christoffersen, Jacobs, and Vasquez (2015). The skewness premium is then given by $SRP^j = \frac{1}{10}(100 - SKEW^j) - RSKEW$, for $j = \{m, BC\}$.

The average daily SRP computed from $\hat{C}_t$ is -5.255, whereas the average of SRP^m is much closer to zero, equal to -2.114. The GMM J-statistic for the test of equality of means is 367.550 and thus lies far in the rejection region of a $\chi^2(1)$ distribution. Similarly, the model-implied 99% confidence interval for the sample mean of SRP^{BC} of [-6.343,-4.168] does not include the corresponding estimated moment for SRP^m. A plausible scenario is that mid-quotes and bias-corrected option quotes imply different *vol-of-vol*. In this case, the increase in the average SRP^{BC} relative to SRP^m is consistent with previous findings and theoretical equilibrium models. Within a long-run risk-type model in the style of Bansal and Yaron (2004), Sasaki (2016) describes SRP as a linear function of stochastic intensity of jumps in the long-run risk factor and the *vol-of-vol* of the consumption growth rate, where the former has a negative impact and the effect of the latter depends on the magnitude of the correlation between intensity and vol-of-vol, and the risk aversion. Similarly, VRP is assumed to be linear in the same two factors, loading positively on both. Thus, if the variance of variance, i.e. the fourth conditional moment, happens to be lower in bias-corrected prices on average, we would expect $\mathbb{E}(VRP^m) > \mathbb{E}(VRP^{BC})$ and $\mathbb{E}(SRP^m) > \mathbb{E}(SRP^{BC})$ (if risk aversion is large and simultaneously the correlation between the two factors is negative; see e.g. Figures 2-5 in Sasaki, 2016), which is exactly what we find in the data[12].

[Figure 8 about here.]

[12] Note that the *vol-of-vol* likely depends on the OTM options that also affect the FI. If the relation between FI and the fourth conditional moment is positive, then an excessive *vol-of-vol* of mid-quotes would also be consistent with our empirical finding that $\mathbb{E}(FI^m) > \mathbb{E}(FI^{BC})$ in the data

In line with our finding for the $SKEW$, we find that SRP^{BC} also has a higher standard deviation (6.819 vs. 1.787) and a slightly larger AR(1) coefficient (0.929 vs. 0.922) than SRP^{m}, but only the former difference is strongly significant according to both types of statistical inference procedures that we rely on. The greater negative skewness (-3.100 vs. -0.258) that SRP^{BC} exhibits is strongly significant with the p-value from the J-test equal to 1.299×10^{-4} and model-implied 99% confidence interval of [-4.164,-2.054]. SRP^{BC} also has a higher kurtosis (15.865 vs. 7.669) relative to SRP^{m}, and GMM-inference suggests that this difference is significant at a 10% and 5% level, but not at the 1% level. Model-implied inference leads to the same conclusion.

It is striking that the $SKEW$ as well as SRP are so much more disperse, have a substantially higher probability for extreme tail events (right tail for $SKEW$ and left tail for SRP), and are prone to more outliers when the measures are based on bias-corrected prices. We gather from these results that there is a high degree of uncertainty about the perceived tail risk and the associated risk premium in the market, and that it is perhaps extremely difficult to *pin down* this risk with the $SKEW$ measure suggested by the CBOE and SRP. Figure 8 supports this conclusion. It plots a very volatile series SRP^{BC}, but it also shows exceptionally wide model-implied confidence intervals that confirm, once more, the inherent ambiguity in extracting the expected compensation for perceived tail risk. Interestingly, in Figure 8 SRP^{BC} and SRP^{BC} exhibit rather similar dynamics and levels during crisis periods, such as the dot-com bubble burst or the Financial Crisis, whereas the differences are vast mostly during tranquil market periods. The confidence intervals are also narrower in turbulent periods. Thus, the two measures pricing perceived tail risk mostly disagree during times when the market does not actually have a long left tail.

5.4 Implications for Option Traders

It is well known that SPX options, especially puts, carry large negative risk premia (Coval and Shumway, 2001, Bondarenko, 2003). Accordingly, buyers of the VIX (squared) portfolio in Equation (1) can be expected to lose money. To get a sense of what our findings imply for option traders, we examine how much money they would lose with such an investment and how the corresponding premia are affected by bid-ask spreads next.

To investigate the returns to the VIX note that Neuberger (1994) (see also Martin, 2013 and

Heston and Li, 2020) show that the VIX represents a portfolio of SPX options that if held to maturity τ will pay

$$pay_t = -2\ln\left(\frac{S_\tau}{S_t(1+r_f)^{\tau-t}}\right) + 2\left(\frac{S_\tau}{S_t(1+r_f)^{\tau-t}} - 1\right), \tag{22}$$

where r_f is the risk-free rate and S_t is the SPX price. This is an "idealized payoff" as it represents the payoff in the case that the investors can trade a continuum of strikes. The return to this investment is given by $ret_t^j = \frac{pay_t - VIX_t^j}{VIX_t^j}$ for $j = \{a, m, b, BC\}$. Given the overlapping nature of the observations on S_t and the well-known associated problems these present for standard econometric inference, we rely on bootstrap-based inference. To accommodate the temporal dependence in the overlapping returns as well a time-varying variance, we resort to the wild tapered block bootstrap of Hounyo (2014) with block size 22 days.

[Table 8 about here.]

In Table 8 we compute the average returns and the standard deviation (monthly) to the VIX portfolio using various measures of VIX and the idealized payoff in (22) as the terminal payoff to the strategy. In line with the extant literature (see e.g. Eraker and Wu, 2017), we find that the return to such an investment is negative. The table shows that the negative rate of return to the VIX portfolio is extremely large at -42% per month for a long investor who buys the portfolio at the bias-corrected prices. This investor would have realized an annualized Sharpe ratio of -1.33. Assuming that this trader can transact at the prevailing midpoint instead gives a Sharpe ratio of -1.47. These numbers are very sizable, yet largely consistent with the existing literature on returns to options writers (see for example Coval and Shumway, 2001, and Bondarenko, 2003).

It is interesting to note that the difference in the annualized Sharpe ratio between midpoint-based and bias-corrected computations is as large as 0.13. The investment strategy based on VIX^m presumably increases the return loss by a point estimate of approximately 1.7% on average per month relative to an investment in the VIX^{BC}-portfolio, because the investor is "paying" for the negative market-maker inventory during the same period (see Figure 1), who in response demands a large positive ask mark-up without doing much to the bid mark-up (see Figure 4) thereby driving up the spread. Thus, when investing in the VIX computed from mid-quotes, the investor is exposing herself to the market-maker inventory risk.

Lastly, Table 8 also summarizes the idealized return if it were possible to invest in the VIX computed from bid and ask quotes, ret^a and ret^b. The average return of an investment in VIX^b is a very low -39% per month, yet the average return when investing in the VIX based on ask quotes is even lower, equal to -48% monthly. Even more notable is the difference in the Sharpe ratios, suggesting that the risk-adjusted return based on bids exceeds the one based on asks by an annualized 0.55. The 90% confidence interval for the Sharpe ratio of ret^b of $[-1.70, -0.75]$ does not include the Sharpe-ratio point estimate for ret^a of -1.74. Once again, this investment example confirms that bid-ask spreads are enormous in the SPX options market.

6 Concluding Remarks

This paper does three things. First, we document several systemic features of bid-ask spreads in options. Notably, asking prices appear more volatile and sensitive to exogenous shocks than bid prices. Bid-ask spreads depend systematically on market factors such as the level of volatility itself. This mechanically implies that OIRMs computed from midpoints are impacted by the spread. This impact is sizable for the VIX, but it is rather enormous for OIRMs that rely heavily on OTM options prices, such as the SKEW and Fear Indices. This evidence adds to the literature on problems relating to the computation and interpretation of the VIX. Jiang and Tian (2017) analyze a host of problems relating to the implicit discretization of the state-price density in the computation of the VIX index. While the discretization itself gives a coarse approximation, the reliance on truncation points also contributes to errors (see Andersen, Bondarenko, and Gonzalez-Perez (2011) for a discussion). Griffin and Shams (2018) document volume spikes in OTM options surrounding settlement dates for VIX futures and options, suggesting that traders use inexpensive OTM SPX options to manipulate VIX derivatives' payoffs.

Second, we derive a simple model where market maker inventory matters in the determination of spreads. We fit this model to the data and find that it replicates the asymmetric nature of the variation in asking prices, relative to bids. Specifically, it generates asking prices that are more volatile than bids, and generally more responsive to exogenous shocks whenever market makers have negative inventory, and vice versa. Since market makers have historically been sellers of options, our model implies that they would have charged a larger mark-up to meet buyer demand than to meet selling demand. In this way, our model is consistent with observed data. Our model differs from other demand-based option pricing models in that we focus on modeling the asymmetry in bid and ask prices. This is crucial in understanding the impact of the spread on OIRMs.

Third, we document that the asymmetry in bid-ask spreads affects popular OIRMs. OIRMs that rely heavily on out-of-the money options, such as measures for higher-order conditional return moments, are particularly biased. The differences in the means and variances of the Fear Index and the SKEW Index when computed from bias-corrected prices and midpoints are particularly large and strongly statistically significant. We also find that the VIX and the VRP, which approximate investors' fear of market events impacting the center of the return distribution and the corresponding risk premium, are overestimated when computed from midpoints. The same is true for the risk premium for market events affected by unpredictable return jumps, as measured by the Fear Index. Conversely, the SKEW, which is designed to measure investors' fear of market events impacting the tails of the return distribution, is underestimated when based on midpoints, and the associated skewness risk premium is too close to zero.

The large bid-ask spreads in SPX options represent a challenge to the computation and interpretation of OIRMs. Quoted spreads are much lower in SPY options, however. For example, on September 27, 2017, the average spread of the two closest ATM call options with a one-month maturity was about 4.4% for SPX vs. less than 1% for equivalent SPY options. This difference in spreads in part motivates the introduction of a new volatility index, SPIKES[13], which is computed from SPY options and essentially relies on the same formula as the CBOE's VIX. The shorter availability of SPY relative to SPX options and the American exercise feature, which makes it optimal to exercise calls prior to lump-sum SPY dividends, complicate the computation of SPIKES and, more generally, OIRMs that are based on SPY options. We leave these topics as possible avenues of future research.

[13]See `https://www.miaxoptions.com/spikes/overview`.

Appendix 1 Proofs

Proof of Proposition 1. We get the following first-order condition for ϵ_t^a:

$$\lambda(\epsilon_t^a)x + \epsilon_t^a\lambda'(\epsilon_t^a)x - \gamma\lambda'(\epsilon_t^a)\left(V(n_t - x) - V(n_t)\right) \tag{23}$$

or

$$1 - D_a\epsilon_t^a + \gamma\frac{1}{x}D_a\left[V(n_t - x) - V(n_t)\right] = 0 \tag{24}$$

$$\epsilon_t^a = \frac{1}{D_a} + \gamma\frac{1}{x}\left(V(n_t - x) - V(n_t)\right) \tag{25}$$

and similarly

$$\epsilon_t^b = \frac{1}{D_b} + \gamma\frac{1}{x}\left(V(n_t + x) - V(n_t)\right) \tag{26}$$

where

$$\begin{aligned} V(n_t - x) - V(n_t) &= (n_t - x)^2 vega_{vs,t}^2 \mathrm{Var}_t(\triangle VIX_t) - n_t^2 vega_{vs,t}^2 \mathrm{Var}_t(\triangle VIX_{t+1}) && (27)\\ &= \left[(n_t - x)^2 - n_t^2\right] vega_{vs,t}^2 \mathrm{Var}_t(\triangle VIX_{t+1}) && (28)\\ &= x\left[-2n_t + x\right] vega_{vs,t}^2 \mathrm{Var}_t(\triangle VIX_{t+1}) && (29) \end{aligned}$$

giving the result. □

Appendix 2 Structural Model Implications

In the following, we discuss some features of the structural model proposed in Section 3. Figure 9 depicts the theoretical spread, $C_t^{\text{ask}} - C_t^{\text{bid}} = \epsilon_t^a + \epsilon_t^b$, as a function of n_t, keeping other factors constant. As seen, the spread is piecewise linear in n_t, with breakpoints determined by $n^+ = (2\gamma D_a Y_t)^{-1} + \frac{1}{2}x$ and $n^- = -(2\gamma D_b Y_t)^{-1} - \frac{1}{2}x$ for $Y_t = vega_{vs,t}^2 \mathrm{Var}_t(\triangle VIX_t)$. The spread attains a minimum for $n_t \in [n^-, n^+]$. Note that $n_t = 0 \in [n^-, n^+]$, and hence the spread is minimized for the case of zero inventory. This follows because the market maker can not do better than having no inventory. The spread increases with a factor $2\gamma Y_t$ when the inventory exceeds n^+ or is lower than n^-. This symmetry is due to the quadratic utility and would likely be slightly asymmetric if the utility function had a convexity that depended on n_t, instead.

[Figure 9 about here.]

The equilibrium expression for the mark-ups in Equations (16) and (17) reveal that if n_t is positive and large, ϵ_t^a is likely to hit its lower bound of zero while at the same time, ϵ_t^b is large. This attracts buy orders and deters sell orders, implying that the market maker's inventory will tend towards zero, i.e. her desired holding. Conversely, when the inventory is negative, ϵ_t^b is small and ϵ_t^a is large. Note that the possibility that (A_b, D_b) and (A_a, D_a) differ implies that arrival intensities of buy and sell orders differ, both conditionally and unconditionally. For example, if $D_a = D_b$ and $A_b > A_a$ (the arrival intensity of buy orders exceeds that of sell orders), the market maker will be left with an equilibrium steady-state inventory that is negative, $\mathbb{E}(n_t) < 0$. A simulated inventory series where $\mathbb{E}(n_t) < 0$ is shown in the top panel of Figure 10.

[Figure 10 about here.]

The middle plot of Figure 10 shows the mark-ups. As seen, ϵ_t^a is larger than ϵ_t^b most of the time. ϵ_t^a is also more volatile than ϵ_t^b and has a higher right skewness. The behavior of the bid-ask spread, $C_t^{\text{ask}} - C_t^{\text{bid}} = \epsilon_t^a + \epsilon_t^b$, is shown in the bottom panel of Figure 10. It is highly right skewed and persistent.

Appendix 3 The Generalized Hyperbolic Distribution

We provide a brief discussion of the Generalized Hyperbolic (GH) distribution. Proposed by Bardorff-Nielsen (1978), the GH is an extremely general distribution with continuous support on the real line. A number of distributions obtain as special cases, including the normal, Laplace, skewed Laplace, skewed normal, Student-t, skewed Student-t, variance gamma, gamma, inverse gamma, exponential, hyperbolic and normal inverse Gaussian. A full discussion of the parametric restrictions that yield these special cases can be found in Bensaïda and Slim (2016).

We employ a five-parameter version suggested in Bensaïda and Slim (2016), $GH(\mu, s^2, \upsilon, \zeta, \rho)$, where the first two parameters are the mean and variance of the distribution. This has the advantage that it can readily be used to ensure that the residual term in (19) has mean zero. The parameters $\upsilon \subseteq \mathbb{R}$, $\zeta \subseteq \mathbb{R}^+$, $|\rho| < 1$ jointly determine the kurtosis and skewness of the distribution.

[Figure 11 about here.]

Figure 11 demonstrates the impact of changing parameter values in the GH distribution. We plot the GH density for different values of each parameter while keeping two parameters fixed at approximately the estimated values in Table 4. We see that both υ and ζ determine the shape of the distribution. ρ determines the skewness in the sense that if $\rho < 0$ the distribution is left skewed. A large value of ζ gives a skewed normal, or normal distribution if $\rho = 0$. Low values of ζ (and $\upsilon = 1$) gives the skewed Laplace or Laplace distribution if $\rho = 0$.

The lower right corner of Figure 11 shows the estimated error density. We use the posterior samples of the parameters reported in Table 4 to compute the estimated density and then report the point-wise credibility intervals for the density. The results show that the residual terms in our non-linear regression model are right skewed and also heavy tailed.

References

Amaya, Diego, Peter Christoffersen, Kris Jacobs, and Aurelio Vasquez, 2015, Does realized skewness predict the cross-section of equity returns?, *Journal of Financial Economics* 118, 135–167.

Andersen, Torben, Oleg Bondarenko, and Maria T. Gonzalez-Perez, 2011, Exploring Return Dynamics via Corridor Implied Volatility, *Review of Financial Studies* 28(10), 2902–2945.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 59, 1481–1509.

Bardorff-Nielsen, Ole, 1978, Hyperbolic Distributions and Distributions on Hyperbolae, *Scandinavian Journal of Statistics* 5, 151–157.

Bekaert, G., M. Hoerova, and M. Lo Duca, 2013, Risk, Uncertainty and Monetary Policy, *Journal of Monetary Economics* 60, 771–788.

Bensaïda, Ahmed, and Skander Slim, 2016, Highly flexible distributions to fit multiple frequency financial returns, *Physica A* 442, 203–213.

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

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

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22, 4463–4492.

Bollerslev, T., and V. Todorov, 2011, Tails, Fears, and Risk Premia, *Journal of Finance* 66, 2165–2211.

Bollerslev, Tim, Viktor Todorov, and Lai Xu, 2015, Tail Risk Premia and Return Predictability, *Journal of Financial Economics* 118, 113–134.

Bondarenko, Oleg, 2003, Why are Put Options So Expensive?, *working paper, University of Illinois.*

Breeden, Douglas, and Robert L. Litzenberger, 1978, Prices of State-Contingent Claims Implicit in Option Prices, *Journal of Finance* 51, 621–651.

Chen, H., S. Joslin, and S. Ni, 2018, Demand for Crash Insurance, Intermediary Constraints, and Risk Premia in Financial Markets, *Review of Financial Studies* forthcoming.

Chordia, Tarun, Alexander Kurov, Dmitriy Muravyev, and Avanidhar Subrahmanyam, 2018, The Informational Role of Index Option Trading, *working paper, Emory.*

Christoffersen, Peter, Ruslan Goyenko, Kris Jacobs, and Mehdi Karoui, 2017, Illiquidity Premia in the Equity Options Market, *The Review of Financial Studies* 31, 811–851.

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

Coval, D. J., and T. Shumway, 2001, Expected Option Returns, *Journal of Finance* 56, 983–1010.

Dennis, Patrick J, and Stewart Mayhew, 2009, Microstructural biases in empirical tests of option pricing models, *Review of Derivatives Research* 12(3), 169–191.

Eraker, Bjørn, and Yue Wu, 2017, Explaining the negative returns to volatility claims: An equilibrium approach, *Journal of Financial Economics* 125, 72–98.

Gârleanu, N., L. H. Pedersen, and A. M. Poteshman, 2009, Demand-Based Option Pricing, *Review of Financial Studies* 22, 4259–4299.

George, Thomas J., and Francis A. Longstaff, 1993, Bid-Ask Spread and Trading Activity in the S&P 100 Index Options Market, *Journal of Financial and Quantitative Analysis* 28, 381–397.

Griffin, John, and Amin Shams, 2018, Manipulation in the VIX?, *Review of Financial Studies* 4(1), 1377–1417.

Hendershott, Terrence, Dmitry Livdan, and Norman Schürhoff, 2015, Are institutions informed about news?, *Journal of Financial Economics* 117, 249–287.

Hentschel, Ludger, 2003, Errors in Implied Volatility Estimation, *Journal of Financial and Quantitative Analysis* 38, 779–810.

Heston, Steven L., and Shuaiqi Li, 2020, Option Momentum, *Working paper, University of Maryland.*

Ho, T. S. Y., and H. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, *Journal of Financial Economics* 9, 47–73.

Hounyo, U., 2014, The wild tapered block bootstrap, *Working paper, CREATES, Aarhus University.*

Jameson, Mel, and William Wilhelm, 1992, Market Making in the Options Markets and the Costs of Discrete Hedge Rebalancing, *The Journal of Finance* 47, 765–779.

Jiang, George J., and Yisong S. Tian, 2017, Extracting Model-Free Volatility from Option Prices: An Examination of the VIX Index, *Journal of Derivatives* 14(3), 35–60.

Lakonishok, Josef, Inmoo Lee, Neil D. Pearson, and Allen M. Poteshman, 2007, Option Market Activity, *The Review of Financial Studies* 20, 813–857.

Martin, Ian, 2013, Simple Variance Swaps, *working paper, LSE.*

Muravyev, Dmitriy, 2016, Order Flow and Expected Option Returns, *Journal of Finance* 71(2), 673–708.

Muravyev, Dmitriy, and Neil Pearson, 2020, Options trading costs are lower than you think, *Review of Financial Studies* 33(11), 4973–5014.

Neuberger, Anthony, 1994, The Log Contract, *Journal of Portfolio Management* 20(2), 74–80.

Pan, Jun, and Allan Poteshman, 2006, The Information in Option Volume for Future Stock Prices, *Review of Financial Studies* 19(3), 871–908.

Sasaki, Hiroshi, 2016, The skewness risk premium in equilibrium and stock return predictability, *Annals of Finance* 12, 95–133.

Schneider, Paul, and Fabio Trojani, 2018, Divergence and the Price of Uncertainty, *Journal of Financial Econometrics* 9, 1–56.

Stoikov, Sasha, and Mehmet Saglam, 2009, Option market making under inventory risk, *Review of Derivatives Research* 12, 55–79.

Weller, Brian, 2019, Measuring Tail Risks at High Frequency, *The Review of Financial Studies* 32, 3571–3616.

Whaley, R. E., 2000, The Investor Fear Gauge, *Journal of Portfolio Management* 26, 12–17.

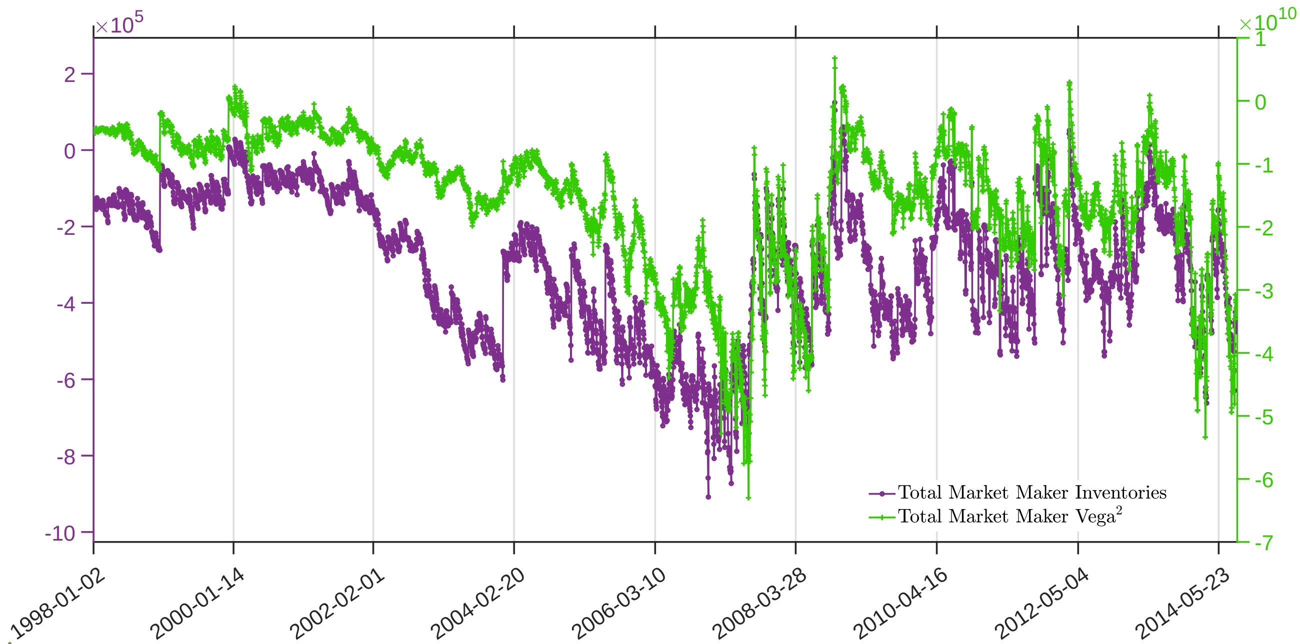

Figure 1: Market-maker inventory and total squared vega position

The plot shows the number of options held by market makers, n_t, defined as the sum of all $n_{i,t}$, on the left y-axis. On the right y-axis, the figure shows the total squared vega of their positions computed as $n_t \times vega_t^2$. The data period is 1998-01-02 to 2014-08-29, based on 4,192 daily observations.

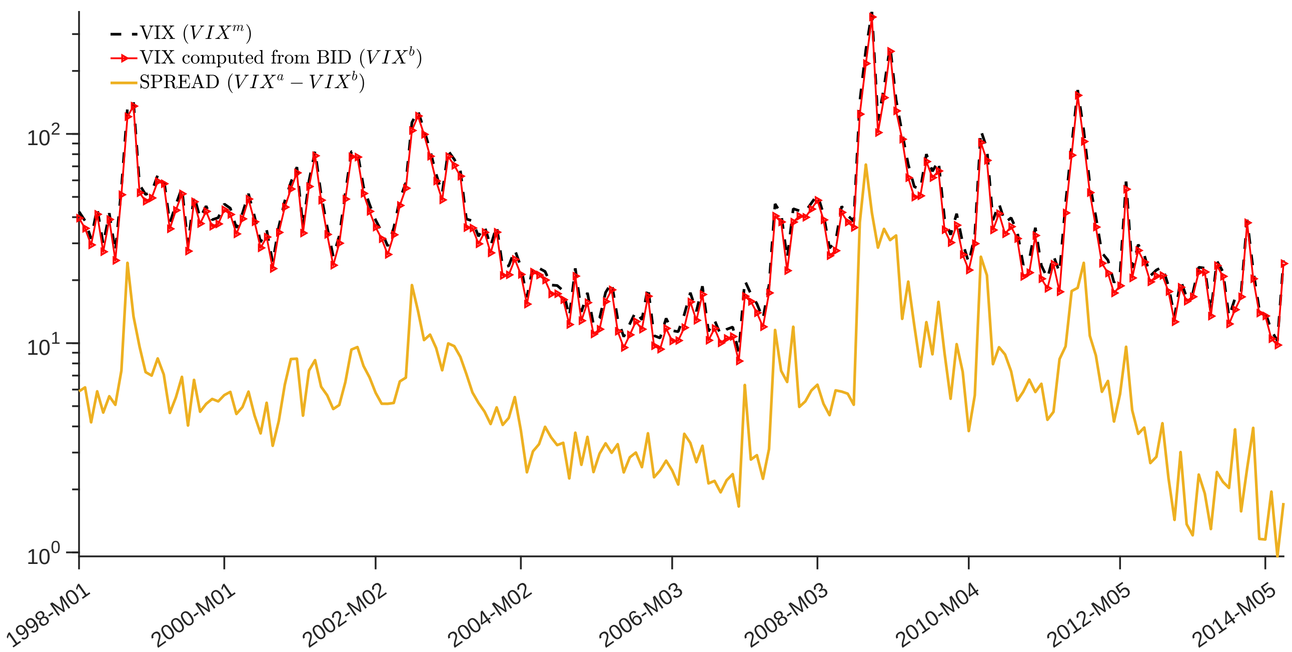

Figure 2: VIX from midpoints, VIX from bids, and VIX bid-ask spreads

The figure plots the VIX computed from the midpoint and from the bid quote, respectively. The yellow line plots the SPREAD, defined as the difference in VIX computed from the asks and bids. The frequency of the three series plotted here, VIX_t^m, VIX_t^b, and $VIX_t^a - VIX_t^b$ is monthly, observing the data on the first trading day of each month for the period 1998-M01 to 2014-M08 (200 observations). The y-axis has logarithmic scale.

Figure 3: Bid-ask spread for SPX at-the-money puts and calls

The top figure plots the bid-ask spread for SPX average ATM puts and calls for the whole sample period from 1998-01-02 to 2014-08-29, based on 4,192 observations. The bottom plot shows the same bid-ask spread for the year 2003, using 252 daily observations, along with vertical bars for the option expiration dates (dark gray) and the last day of the month (light gray).

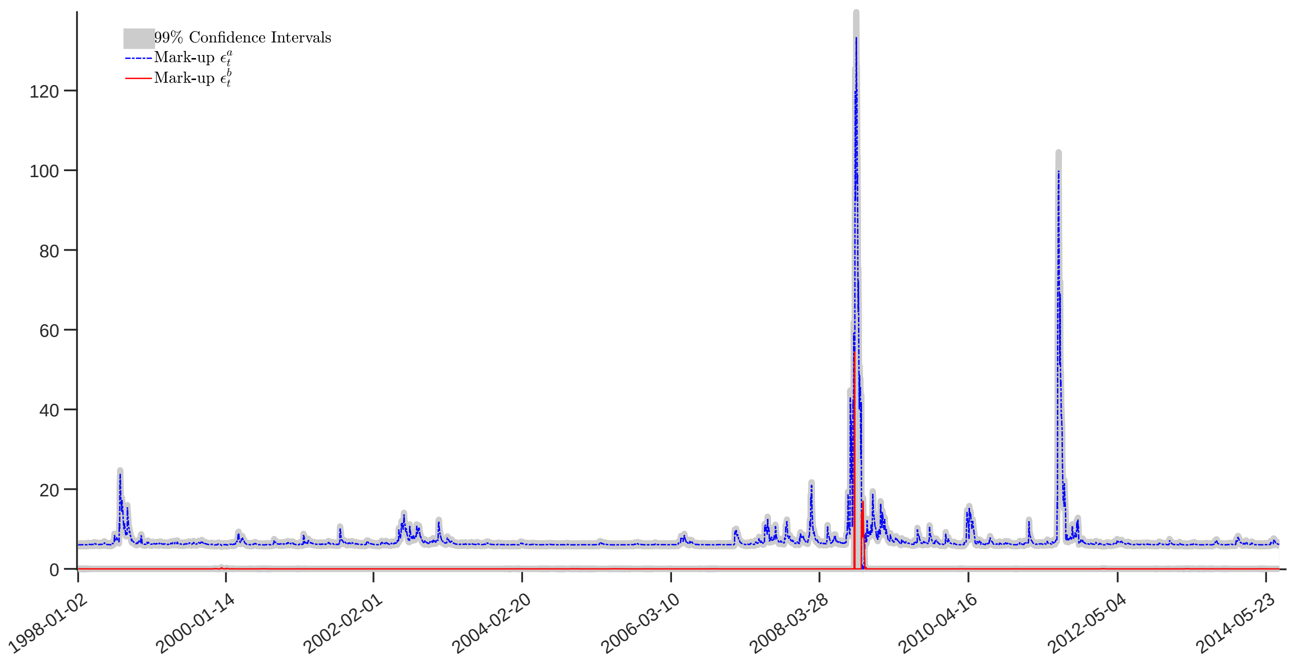

Figure 4: Estimated mark-ups for ask and bid quotes

The figure plots the estimated ask and bid mark-ups from the equilibrium model. These series are based on the MLE estimates for Equation (18), reported in Table 4, for the VIX spread. Also plotted are 99% confidence intervals, which are based on standard errors computed by finite differences and the Delta Method. The sample period is 1998-01-02 to 2014-08-29, with 4,192 daily observations.

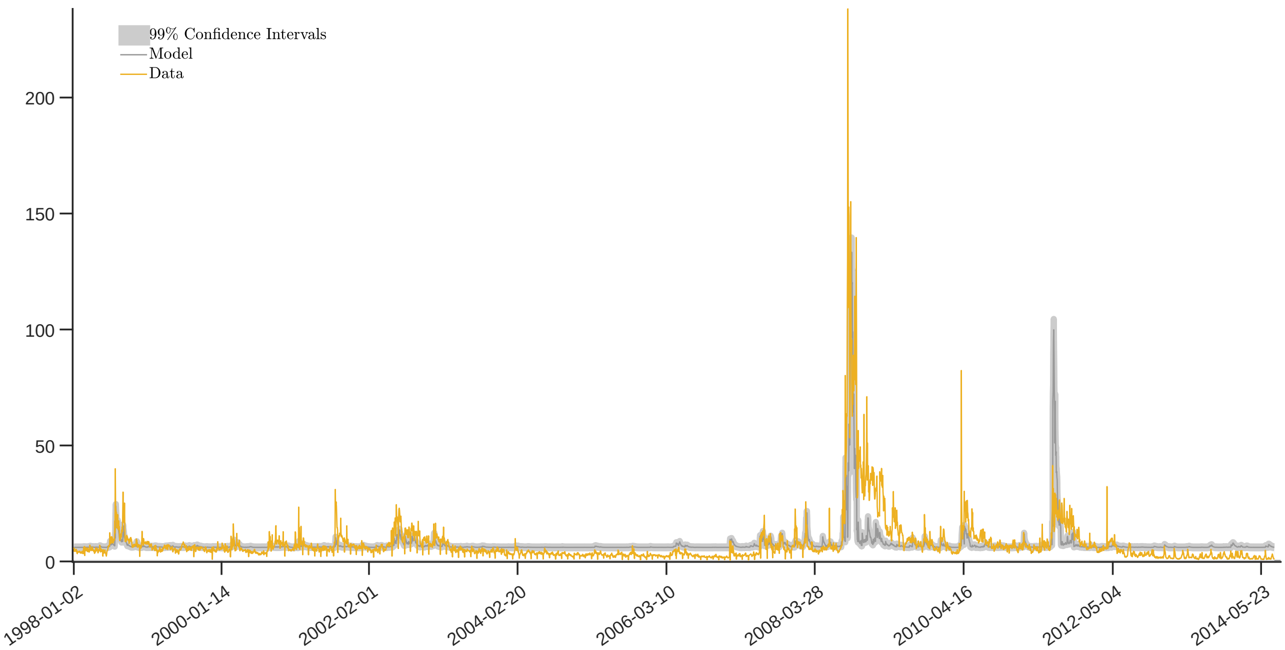

Figure 5: Observed vs. estimated bid-ask spreads

The figure plots the observed and the fitted spread from the equilibrium model. The observed series is $VIX_t^a - VIX_t^b$. The fitted series is based on the MLE estimates for Equation (18), reported in Table 4, for the VIX spread. Also plotted are 99% confidence intervals, which are based on standard errors computed by finite differences and the Delta Method. The sample period is 1998-01-02 to 2014-08-29, with 4,192 daily observations.

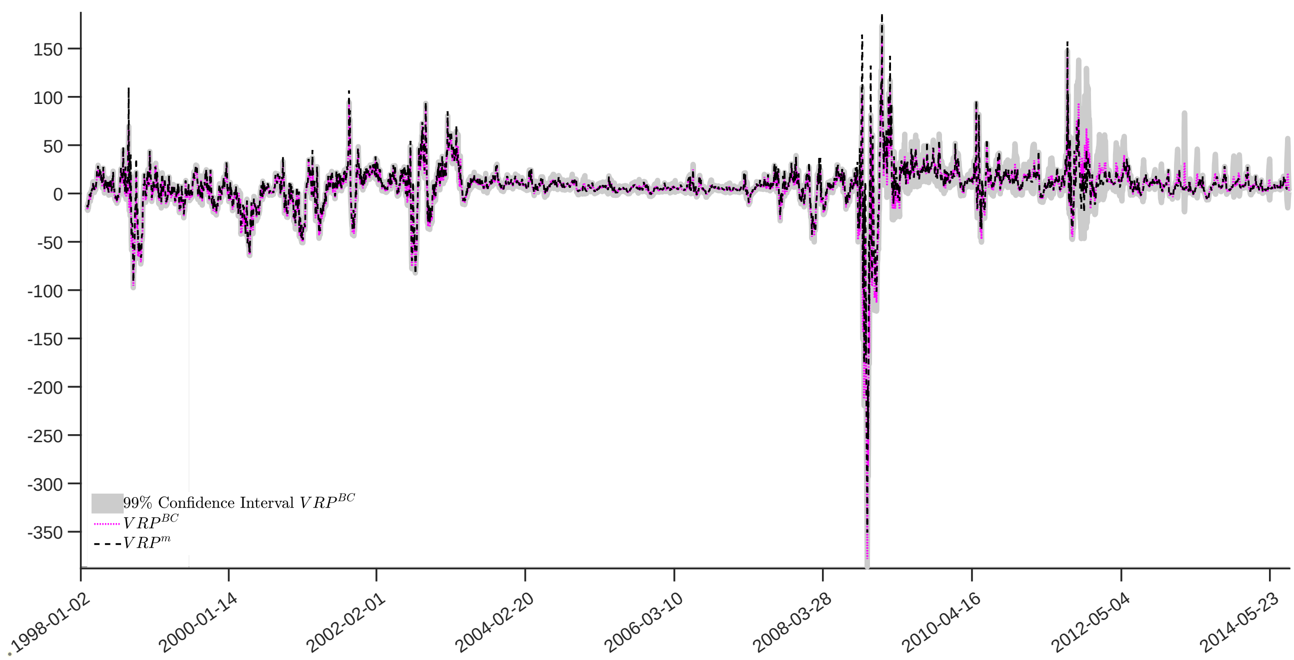

Figure 6: VRP from mid-quotes and bias-corrected quotes

The figure plots the time series of Variance Risk Premia computed from mid-quotes, VRP^m, and bias-corrected quotes, VRP^{BC}. Also plotted are 99% confidence intervals, which are based on standard errors computed by finite differences and the Delta Method. The sample period is 1998-01-02 to 2014-08-29, with 4,192 daily observations.

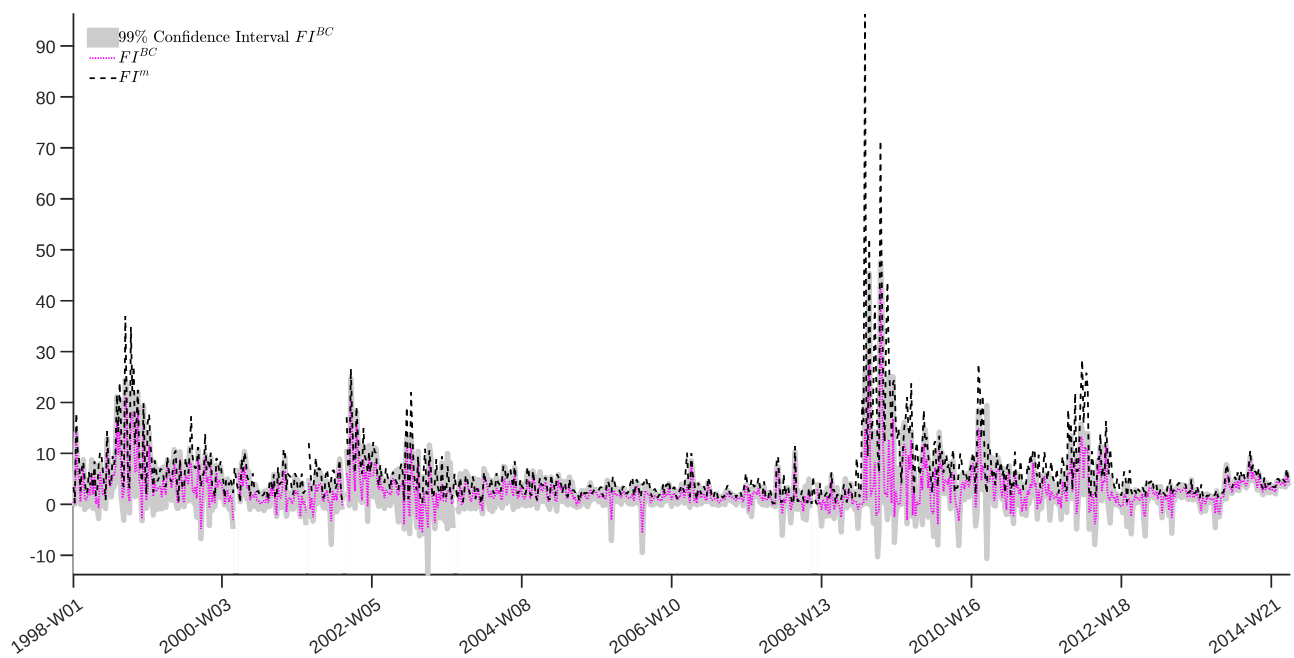

Figure 7: FI from mid-quotes and bias-corrected quotes

The figure plots the time series of Fear Indices computed from mid-quotes, FI^m, and bias-corrected quotes, FI^{BC}. Also plotted are 99% confidence intervals, which are based on standard errors computed by finite differences and the Delta Method. The sample period is 1998-W01 to 2014-W34 ('W' stands for calendar week), with 869 weekly observations.

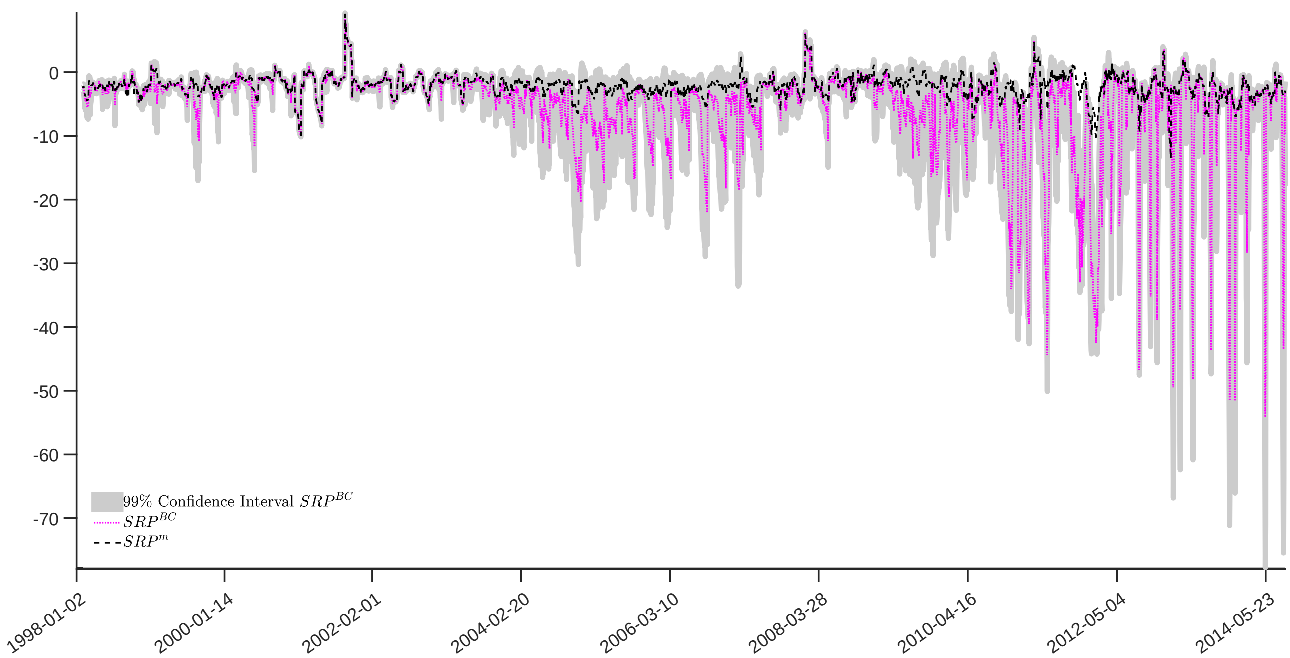

Figure 8: SRP from mid-quotes and bias-corrected quotes

The figure plots the time series of Skewness Risk Premia computed from mid-quotes, SRP^m, and bias-corrected quotes, SRP^{BC}. Also plotted are 99% confidence intervals, which are based on standard errors computed by finite differences and the Delta Method. The sample period is 1998-01-02 to 2014-08-29, with 4,192 daily observations.

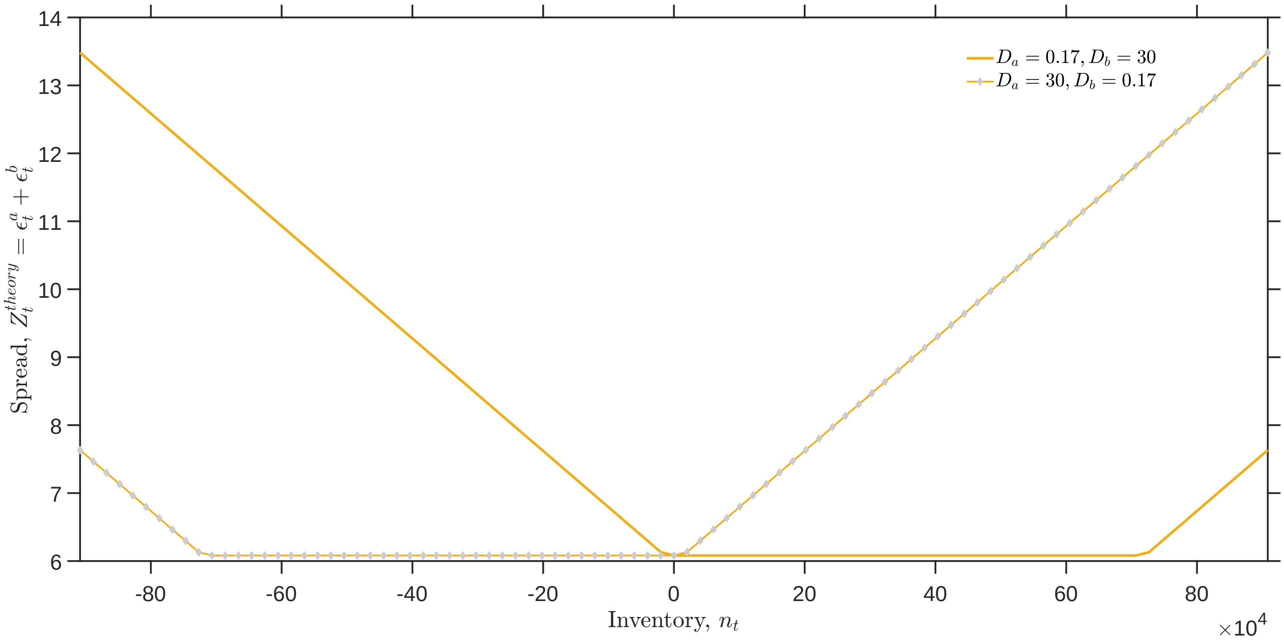

Figure 9: Theoretical bid-ask spread as a function of inventory

The figure plots the theoretical bid-ask spread, $Z_t^{\text{theory}} = \epsilon_t^a + \epsilon_t^b$, as a function of inventory, n_t, resulting from the market-maker model. All other modeling parameters are held constant ($vega_t^2 = 51.79 \times 10^3$, $\text{Var}_t(\Delta VIX_{t+1}) = 181.58$, $x = 0.02$, $\gamma = 0.44$).

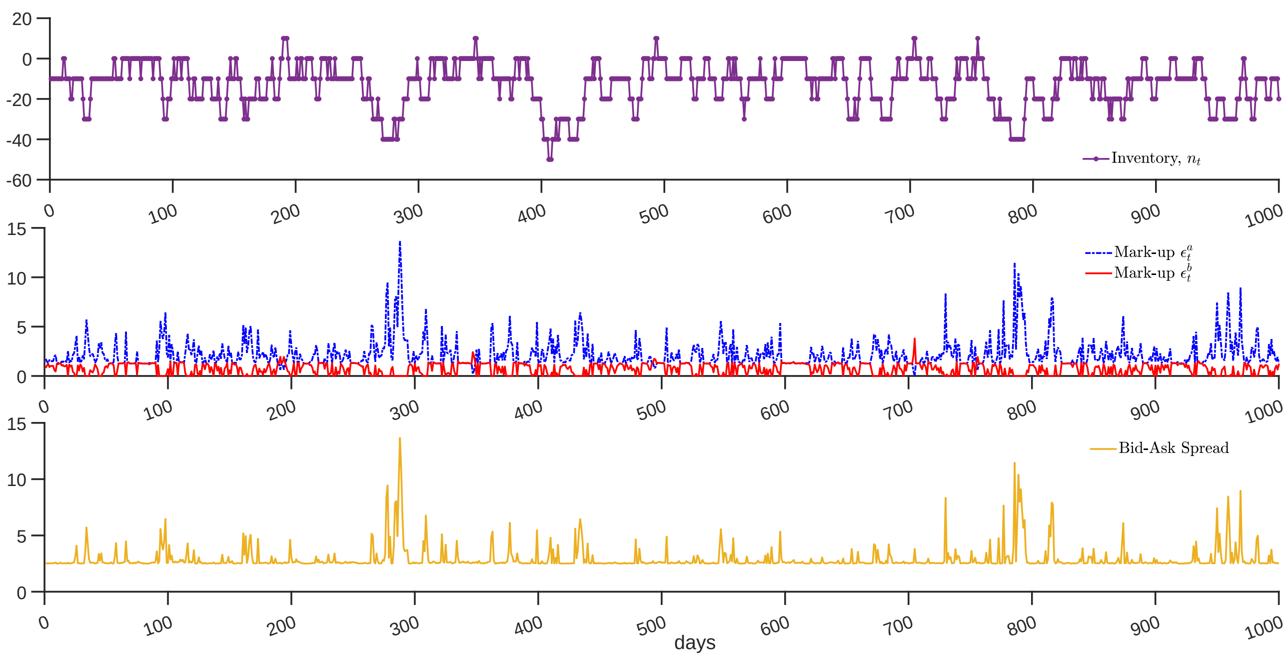

Figure 10: Simulated inventory, mark-ups, and bid-ask spreads

The top figure shows simulated inventory, the middle figure plots simulated mark-ups ϵ_t^a and ϵ_t^b, and the bottom figure depicts the bid-ask spread, $\epsilon_t^a + \epsilon_t^b$. Data are simulated from our model using equations (13) for arrivals of bids and asks and (16)-(17) for the mark-ups. $\mathrm{Var}_t(\Delta VIX_{t+1})$ is simulated as a standard Cox-Ingersoll-Ross process with long-run mean 0.0001, mean-reversion rate 0.008, and volatility parameter 0.16. We set the model parameters as follows: $\gamma = 0.31$, $vega^2 = 10$, $D_a = 1.07$, $D_b = 2.066$, $A_a = 1$, and $A_b = 0.4$.

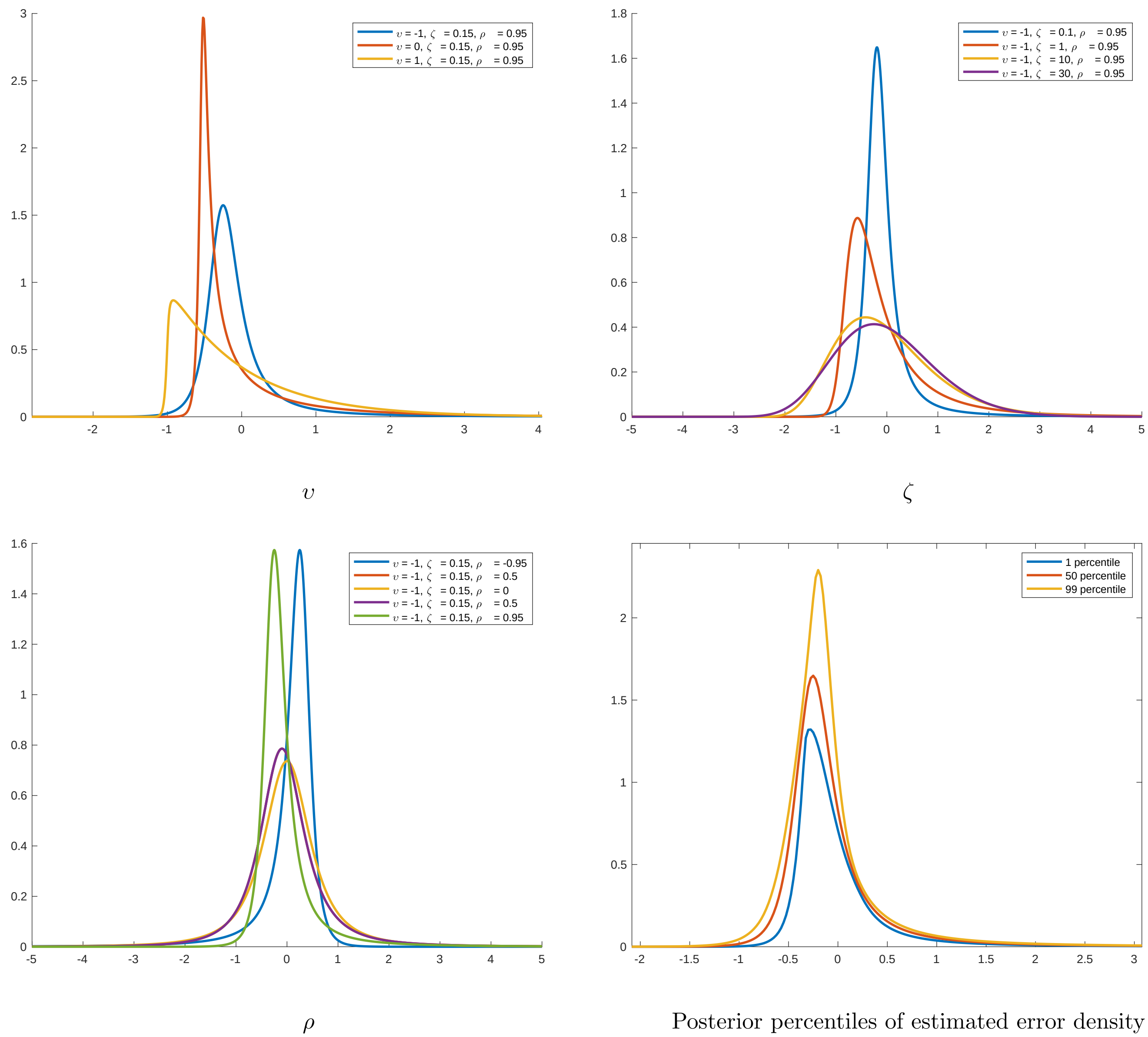

Figure 11: Impact of parameter values in the Generalized Hyperbolic density.

Table 1: Summary Statistics

The table reports descriptive statistics for our data set of 4,192 daily observations from January 2, 1998 to August 29, 2014. VIX is reported as a percentage variance measure scaled into monthly units. Similarly, the risk-free rate, $r_t^{(f)}$, is quoted as a continuously compounded monthly percentage. ALL (ATM) denotes the average prices of all (all at-the-money) option contracts written on the SPX. Inventory n_t is the sum of all holdings of active contracts in the market maker's portfolio. $vega_t^2$ is the average of the squared vega of all these contracts. $\widehat{\mathrm{Var}}_t(\Delta VIX_{t+1})$ is the prediction from the GARCH model (3)-(4). Finally, the Return is the log daily close-to-close price change, and RV is the sum of 77 5-minute intraday squared log returns.

	Average	Std. Dev.	Skewness	Kurtosis	Min.	Max.
VIX^b	41.086	41.074	4.059	27.403	7.426	461.452
VIX^a	48.597	51.189	4.569	33.589	8.981	612.812
$VIX^a - VIX^b$	7.510	11.416	8.038	98.626	0.496	238.223
ATM^b	36.411	21.189	1.085	3.428	2.750	107.143
ATM^a	38.343	21.724	1.068	3.393	3.625	113.269
$\mathrm{ATM}^a - \mathrm{ATM}^b$	1.932	0.944	3.366	31.312	0.125	15.325
ALL^b	136.846	30.098	0.412	2.027	77.926	212.794
ALL^a	139.017	30.543	0.409	2.008	79.721	215.780
$\mathrm{ALL}^a - \mathrm{ALL}^b$	2.170	0.825	3.311	31.855	0.193	13.457
$r_t^{(f)}$	0.185	0.173	0.409	1.599	0	0.524
n_t	-29.93×10^4	17.73×10^4	-0.405	2.419	-90.83×10^4	12.45×10^4
$vega_t^2$	51.79×10^3	16.81×10^3	0.829	3.414	21.83×10^3	117.32×10^3
$\widehat{\mathrm{Var}}_t(\Delta VIX_{t+1})$	181.583	867.640	9.668	112.444	1.139	13.63×10^3
Return	1.73×10^{-4}	1.27×10^{-2}	-0.127	9.845	-0.097	0.109
RV	1.194	2.313	9.421	149.710	0.033	59.068

Table 2: Determinants of the Spread

This table reports the results of OLS regression of Spread, defined as $VIX^a - VIX^b$ onto contemporaneous S&P 500 returns, change in Realized Variance, level of Realized Variance, end of month and expiration dummies, as well as variables related to market-makers' inventories of options: Inventory is the raw inventory series, $vega^2$ is the daily average vega of all options. The results are based on 4191 daily observations for the period 1998-01-05 to 2014-08-29. The RV data are based on 5-minute returns from TAQ. Newey-West standard errors are based on 120 lags.

Constant	4.147^{***}	2.980^{***}	2.260^{***}	3.936^{***}	3.978^{***}
	(0.614)	(0.539)	(0.666)	(0.584)	(0.583)
Return		-0.604^{**}	-0.966^{***}		-0.935^{**}
		(0.270)	(0.387)		(0.457)
RV change			-1.771^{***}	-0.337^{**}	-0.390^{**}
			(0.369)	(0.191)	(0.194)
RV level		3.802^{***}	4.408^{***}		
		(0.754)	(0.836)		
Month End				-1.501^{***}	-1.578^{***}
				(0.421)	(0.415)
Expiration				1.268^{*}	1.225^{*}
				(0.798)	(0.757)
$\lvert\text{Inv}\times vega^2 \times RV\rvert$	2.664^{***}			2.750^{***}	2.734^{***}
	(0.745)			(0.780)	(0.782)
R^2	0.251	0.601	0.674	0.257	0.262

Table 3: Determinants of the VIX^a and VIX^b

This table reports the results of seemingly unrelated regressions (SUR) of VIX^a and VIX^b onto contemporaneous S&P 500 returns, change in Realized Variance, level of Realized Variance, end of month and expiration dummies, as well as variables related to market-makers' inventories of options: Inventory is the raw inventory series, $vega^2$ is the daily average vega of all options. The results are based on 4191 daily observations for the period 1998-01-05 to 2014-08-29. The RV data are based on 5-minute returns from TAQ. Newey-West standard errors are based on 120 lags.

			VIX^b		
Constant	29.502***	25.876***	22.785***	28.467***	28.724***
	(3.532)	(2.722)	(2.312)	(3.302)	(3.281)
Return		−4.298***	−5.850***		−5.726***
		(1.106)	(1.413)		(1.672)
RV change			−7.606***	−2.533***	−2.857***
			(0.792)	(0.729)	(0.728)
RV level		12.786***	15.390***		
		(1.506)	(1.488)		
Month End				0.352	−0.118
				(0.815)	(0.828)
Expiration				1.619	1.356
				(3.516)	(3.325)
$\lvert \text{Inv} \times vega^2 \times RV \rvert$	9.173***			9.803***	9.704***
	(1.931)			(2.100)	(2.116)
R^2	0.230	0.538	0.642	0.244	0.260
			VIX^a		
Constant	33.648***	28.855***	25.045***	32.404***	32.702***
	(4.061)	(2.926)	(2.452)	(3.777)	(3.757)
Return		−4.902***	−6.816***		−6.661***
		(1.364)	(1.791)		(2.122)
RV change			−9.376***	−2.871***	−3.247***
			(0.945)	(0.874)	(0.874)
RV level		16.588***	19.798***		
		(2.009)	(1.985)		
Month End				−1.150	−1.696*
				(1.070)	(1.056)
Expiration				2.887	2.580
				(4.268)	(4.034)
$\lvert \text{Inv} \times vega^2 \times RV \rvert$	11.837***			12.553***	12.438***
	(2.623)			(2.821)	(2.841)
R^2	0.247	0.579	0.681	0.258	0.272

Table 4: Parameter Estimates

The table reports parameter estimates of our equilibrium spread model. Mark-ups ϵ_t^a and ϵ_t^b are given by

$$\epsilon_t^a = \max\left(\frac{1}{D_a} + \gamma\,(-2n_t + x)\,vega_t^2 \mathrm{Var}_t(\triangle \mathrm{VIX}_t), 0\right), \tag{30}$$

$$\epsilon_t^b = \max\left(\frac{1}{D_b} + \gamma\,(2n_t + x)\,vega_t^2 \mathrm{Var}_t(\triangle \mathrm{VIX}_t), 0\right). \tag{31}$$

We estimate $D_a^{-1}, D_b^{-1}, \gamma, s$ using Bayesian MCMC and MLE assuming that the observed VIX spread $Z_t^{\text{data}} - Z_t^{\text{theory}}(\Theta) \sim N(0, s^2)$ (Panel A) or $Z_t^{\text{data}} - Z_t^{\text{theory}}(\Theta) \sim GH(0, s^2, \upsilon, \zeta, \rho)$ (Generalized Hyperbolic - Panel B). The model is estimated with 4,170 daily observations from February 4, 1998 to August 29, 2014. Note that we scale the raw inventory, n_t, and squared vega, $vega_t^2$, series by a constant for the estimation to ensure comparable magnitudes.

	D_a^{-1}	D_b^{-1}	γ	x	s	υ	ζ	ρ	R^2
				Panel A: Normal errors					
				Bayes					
Mean	6.01	0.04	0.44	0.02	8.97				39.2%
St.dev.	(0.14)	(0.03)	(0.01)	–	(0.10)				
				MLE					
MLE	6.01	0.03	0.44	0.02	8.96				39.2%
St.Err.	(0.15)	(0.01)	(0.01)	–	(0.10)				
				Panel B: Generalized Hyperbolic errors					
				Bayes					
Mean	6.43	0.02	0.32	0.02	10.00	−1.04	0.15	0.96	35.8%
St.dev.	(0.15)	(0.01)	(0.01)	–	(1.23)	(0.06)	(0.03)	(0.01)	
				MLE					
MLE	6.45	0.01	0.32	0.02	10.12	−1.05	0.14	0.96	35.8%
St.Err.	(0.09)	(0.00)	(0.01)	–	(0.66)	(0.05)	(0.02)	(0.01)	

Table 5: Indices computed from mid-quotes and Bias-Corrected Quotes

Descriptive Statistics for the VIX, the FI, and the SKEW computed from mid-quotes and bias-corrected prices, for 1998-01-02 - 2014-08-29 with 4,192 daily observations. For the latter series, we report standard errors in square brackets that reflect the estimation uncertainty in Section 4. For comparison, we also add the original CBOE VIX and the CBOE SKEW Indices. For FI computed from mid quotes, we also report the time-varying tail shape parameters, α, and the (left) right jump variation, $(R)LJV$; the latter is in annualized percentage (square) variance units. In parenthesis, we outline the corresponding values reported in Bollerslev et al. (2015) for their 1996-01 - 2013-08 sample. The "GMM-est." denotes the GMM estimate that restricts the respective moment of $(\cdot)^m$ and $(\cdot)^{BC}$ to be the same. The "J-stat." reports the corresponding test statistic for the moment restriction, and *, **, *** denotes a rejection at a 10%, 5%, and 1% level (based on $\chi^2(1)$ critical values).

	"Original" CBOE Squared VIX (scaled)	VIX^m	VIX^{BC}	GMM-est.	J-stat.
Mean	44.48	44.84	43.24	43.97	22.29***
			[0.82]		
Std.	45.06	46.05	41.86	43.48	0.00
			[0.39]		
Skew.	4.28	4.32	3.97	4.13	3.53*
			[0.05]		
Kurt.	30.66	30.37	26.45	28.13	0.00
			[0.62]		
AR(1)	0.97	0.97	0.97	0.97	0.00
			[0.00]		

	VRP^m	VRP^{BC}	GMM-est.	J-stat.
Mean	7.89	6.30	7.25	14.04***
		[0.82]		
Std.	23.17	24.58	23.73	0.00
		[0.27]		
Skew.	-3.43	-4.64	-3.92	16.53***
		[0.23]		
Kurt.	51.30	56.83	53.68	0.00
		[2.80]		
AR(1)	0.86	0.90	0.88	0.97
		[0.00]		

	α_m^-	α_m^+	LJV_m	RJV_m	FI^m	FI^{BC}	GMM-est.	J-stat.
Mean	14.83	56.44	0.48	0.01	6.11	2.91	3.36	156.82***
	(16.23)	(61.81)	(0.45)	(0.02)		[0.10]		
Std.	4.47	16.23	0.56	0.03	6.86	3.86	4.00	102.50***
	(5.33)	(19.68)	(0.54)	(0.05)		[0.11]		
Skew.	0.43	0.39	5.24	9.96	5.29	2.95	3.36	3.82*
	(0.43)	(0.65)	(5.41)	(5.03)		[0.04]		
Kurt.	3.32	3.16	50.89	129.23	51.26	20.03	24.17	2.21
	(3.00)	(3.52)	(48.62)	(32.80)		[0.24]		

Table 5 – *Continued from previous page*

AR(1)	0.7341	0.80	0.53	0.18	0.51	0.35	0.38	3.51*
	(0.59)	(0.67)	(0.69)	(0.11)		[0.00]		

	"Original" CBOE SKEW	$SKEW^m$	$SKEW^{BC}$	GMM-est.	J-stat.
Mean	118.79	118.57	149.91	119.03	1,427.03***
			[4.20]		
Std.	5.80	5.79	63.99	5.80	302,571.76***
			[0.63]		
Skew.	0.70	0.69	3.32	1.11	24.72***
			[0.49]		
Kurt.	3.77	3.76	41.37	4.65	386.27***
			[4.02]		
AR(1)	0.9038	0.91	0.93	0.91	0.01
			[0.01]		

	SRP^m	SRP^{BC}	GMM-est.	J-stat.
Mean	-2.11	-5.26	-2.29	367.55***
		[0.42]		
Std.	1.79	6.82	1.82	3,109.19***
		[0.06]		
Skew.	-0.26	-3.10	-1.18	14.64***
		[0.41]		
Kurt.	7.67	15.86	9.77	5.09**
		[3.19]		
AR(1)	0.92	0.93	0.92	0.00
		[0.00]		

Table 6: Fear Index and the Spread

This table reports the results of OLS regression of the Fear Index computed from midpoints (FI^m) onto the VIX, the VIX bid-ask spread ($VIX^a - VIX^b$), estimated ask and bid mark-ups (ϵ_t^a and ϵ_t^b), as well as their sum. The data period stretches from 1998-01-09 to 2014-08-22, covering 868 weekly observations.

Constant	2.256^{***}	2.140^{***}	1.488^{**}	1.365^{*}	2.159^{***}
	(0.608)	(0.500)	(0.792)	(0.886)	(0.497)
VIX	0.114^{***}	0.061^{**}	0.101^{***}	0.110^{***}	0.059^{***}
	(0.035)	(0.026)	(0.030)	(0.040)	(0.024)
Spread	-0.159^{*}		-0.178^{*}	-0.199^{*}	
	(0.123)		(0.116)	(0.136)	
ϵ_t^a		0.176^{*}		0.188^{**}	
		(0.116)		(0.109)	
ϵ_t^b		-0.387		-1.121^{*}	
		(0.884)		(0.860)	
$\epsilon_t^a + \epsilon_t^b$			0.201^{**}		0.183^{**}
			(0.113)		(0.110)
R^2	0.286	0.292	0.316	0.321	0.291

Table 7: SKEW Index and the Spread

This table reports the results of OLS regression of the SKEW Index computed from midpoints ($SKEW^m$) onto the VIX, the VIX bid-ask spread ($VIX^a - VIX^b$), estimated ask and bid mark-ups (ϵ_t^a and ϵ_t^b), as well as their sum. The data period stretches from 1998-01-05 to 2014-08-29, covering 4,191 daily observations.

Constant	120.478***	119.318***	119.935***	119.988***	119.284***
	(0.938)	(0.829)	(0.849)	(0.848)	(0.827)
VIX	−0.083***	−0.047***	−0.087***	−0.088***	−0.046***
	(0.032)	(0.017)	(0.031)	(0.031)	(0.016)
Spread	0.244**		0.204**	0.207**	
	(0.109)		(0.102)	(0.100)	
ϵ_t^a		0.183***		0.131***	
		(0.055)		(0.032)	
ϵ_t^b		0.342***		0.338***	
		(0.138)		(0.117)	
$\epsilon_t^a + \epsilon_t^b$			0.133***		0.183***
			(0.034)		(0.056)
R^2	0.105	0.090	0.121	0.123	0.089

Table 8: Returns to the VIX Portfolio

This table summarizes the net returns, $ret_t^j = \frac{pay_t}{VIX_t^j} - 1$, $j = \{a, m, b, BC\}$, to investors who buy the VIX portfolio and hold it to expiration, realizing the payoff pay_t given in (22). We report monthly returns, their standard deviations, and annualized Sharpe ratios. The time period covers 1998-01-02 to 2014-07-31 with 4,171 observations, and VIX_t^j is computed from bids, mid-quotes, asks, and bias-corrected prices, respectively. Square brackets contain 90% confidence intervals obtained by the wild tapered block bootstrap of Hounyo (2014).

	ret^a	ret^m	ret^b	ret^{BC}
Mean	-0.482	-0.440	-0.391	-0.423
	[-0.557,-0.402]	[-0.521,-0.354]	[-0.479,-0.296]	[-0.510,-0.330]
Std.	0.962	1.041	1.137	1.100
	[0.702,1.289]	[0.762,1.392]	[0.831,1.521]	[0.793,1.480]
Sharpe Ratio	-1.735	-1.465	-1.190	-1.333
	[-2.422,-1.179]	[-2.058,-0.966]	[-1.704,-0.749]	[-1.900,-0.853]