Inventory Risk, Market Maker Wealth, and the Variance Risk Premium: Theory and Evidence*

Mathieu Fournier
HEC Montréal

Kris Jacobs
University of Houston

December 8, 2015

Abstract

We investigate the role of option market makers in the determination of the variance risk premium and the valuation of index options. A reduced-form analysis indicates that a substantial part of the variance risk premium is driven by inventory risk and market maker wealth. When market makers experience extreme wealth losses, a one standard deviation change in inventory risk leads to a change in the variance risk premium of over 6%. Motivated by these findings, we develop a structural model of a market maker with limited capital who is exposed to market variance risk through his inventory. We derive the endogenous variance risk premium and characterize its dependence on inventory risk and market maker wealth. We estimate the model using index returns and options and find that it performs well, especially during the financial crisis.

JEL Classification: G10; G12; G13.

Keywords: Variance risk premium; inventory risk; financial constraints; option pricing.

*For helpful comments we thank Daniel Andrei, David Bates, Peter Christoffersen, Julien Cujean, Christian Dorion, Redouane Elkamhi, Wayne Ferson, Bruno Feunou, Jean-Sébastien Fontaine, Pascal François, Christian Gouriéroux, Amit Goyal, Ruslan Goyenko, Michael Hasler, Alexandre Jeanneret, Bryan Kelly, Aytek Malkhozov, Tom McCurdy, Dmitriy Muravyev, Chayawat Ornthanalai, Stylianos Perrakis, Norman Schüerhoff, Masahiro Watanabe, Jason Wei, Alan White, seminar participants at the Bank of Canada, HEC Lausanne/Swiss Finance Institute, HEC Montréal, Rotman, conference participants at IFM2, IFSID, and NFA, and market makers at the CBOE. Correspondence to: Mathieu Fournier, HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, QC H3T 2A7. E-mail: mathieu.fournier@hec.ca.

1 Introduction

On average, net demand for index options by end users is positive (Bollen and Whaley, 2004; Gârleanu, Pedersen, and Poteshman, 2009). Market makers in index options are therefore net sellers and build up large negative inventories over time. Through this inventory, market makers are exposed to market variance risk in addition to market return risk.[1] While market makers can effectively hedge return risk using index futures contracts, frictions limit their ability to eliminate this large exposure to market variance (Bates, 2003). Consequently, market makers carry large variance risk exposures. This suggests that their risk bearing capacity may affect supply, which will in turn contribute to the determination of the variance risk premium and trading activity.

We confirm that market makers' risk bearing capacity has a substantial impact on the variance risk premium. We first present the results of a reduced form analysis, using fifteen years of market maker activity for index options and more than one million quotes. We regress the variance risk premium on two variables that are informative about market makers' risk bearing capacity. The first variable captures the exposure of market makers' inventory to market variance risk, and we refer to it as inventory risk. The second variable, market maker wealth, measures market makers' aggregate trading revenues. We construct daily estimates of these variables using data on aggregate CBOE market maker positions.

Our empirical analysis indicates that these risk capacity variables substantially impact on the variance risk premium. When inventory risk varies by one standard deviation, it results in a 1.2% change in the variance risk premium, which is more than twenty times the average daily change in the variance risk premium. Moreover, this effect is magnified when market makers experience dramatic wealth losses. When market makers' loss is at its ninetieth percentile, a one standard deviation change in inventory risk can cause up to 6% variation in the variance risk premium.

Motivated by these findings, we present a structural model of a continuous-time economy with dynamic market variance and a risk-averse representative market maker with limited capital who quotes index option prices. Because market variance fluctuates randomly over time, the market maker is exposed to variance risk which we assume to be unhedgeable. We solve for the endogenous variance risk premium that induces the market maker to clear the index option market. The model provides a number of important new insights.

First, the variance risk premium co-moves with inventory risk to reward the market maker for his risk exposure. When the market maker absorbs net buying pressure, his inventory becomes more negative and so does his inventory risk. Because of risk aversion, the market maker then

[1] An extensive literature documents that the market variance is a source of aggregate risk and that security returns that co-move with market variance contain a variance risk premium. For some seminal contributions to this literature see Bakshi, and Kapadia (2003), Ang, Hodrick, Xing, and Zhang (2006), and Bollerslev, Tauchen, and Zhou (2010).

requires higher compensation which, given his negative exposure, translates into a more negative variance risk premium. Second, the model characterizes the substantial impact of market maker wealth on the variance risk premium. When the market maker incurs losses, his marginal utility of wealth goes up and his required compensation increases. Because the market maker bears negative variance exposure, a higher compensation implies a more negative variance risk premium. By deriving an explicit relation between market maker wealth, inventory risk and the variance risk premium in a stochastic volatility model, these results complement those of Gârleanu, Pedersen, and Poteshman (2009), who show that net option demand exerts pressure on option prices when markets are incomplete.

While the structural model incorporates the optimal strategy of the market maker, it is relatively parsimonious and it can easily be implemented for option valuation. We estimate the model using index returns and a large panel of index put options, and compare its performance to the benchmark Heston (1993) stochastic volatility model. The model performs well both in- and out-of-sample. It performs particularly well during the financial crisis and for pricing out-of-the-money puts, which are more challenging for the benchmark Heston model. Our estimates suggest that during turbulent times, fluctuations in market maker wealth lead to daily changes in option prices of more than 2%.

Our findings contribute to the growing literature on the variance risk premium.[2] In standard stochastic volatility models, the variance risk premium is the product of a time-invariant parameter and the latent spot variance (see, among others, Heston, 1993; Bates, 2000; Pan, 2002; Egloff, Leippold, and Wu, 2010). This specification attributes any discrepancy between the objective distribution of index returns and the risk-neutral probability measure implied by option prices to the marginal investor's preferences over aggregate wealth. Our study departs from this literature by modeling the influence of market makers' risk bearing capacity on the variance risk premium. This model feature is related to the findings of Adrian, Etula, and Muir (2014), who show that intermediaries' leverage ratio is important for explaining the cross-section of equity returns. It is consistent with the lessons from the financial crisis, underlining the non-trivial influence of financial intermediaries' positions and constraints on asset prices (see, for instance, Adrian and Shin, 2010).

Our findings are also closely related to several other strands of literature in option valuation and asset pricing. Bollen and Whaley (2004) demonstrate that net buying pressure positively impacts

[2]See for instance Bakshi and Kapadia (2003), Driessen, Maenhout, and Vilkov (2009), Vilkov (2008), and Carr and Wu (2009) on the structure of the variance risk premium. Egloff, Leippold, and Wu (2010) and Todorov (2010) explain the variance risk premium dynamic using stochastic volatility models with jumps. Ang, Hodrick, Xing, and Zhang (2006) and Cremers, Halling, and Weinbaum (2014) study the cross-section of stock returns, and Schürhoff and Ziegler (2011) study expected option returns. Barras and Malkhozov (2014) compare the variance risk premium from the cross-section of stock returns with the one implied by options. Aït-Sahalia, Karaman, and Mancini (2014), and Dew-Becker, Giglio, Le, and Rodriguez (2015) study the price of variance risk embedded in the term structure of variance swaps.

option implied volatilities and thus prices, but they do not model the channels by which net demand impacts option prices. Leippold and Su (2011) examine the impact of margin requirements on option implied-volatilities in a constant volatility framework. Chen, Joslin, and Ni (2013) investigate the jump premium embedded in index options and its predictive ability for stock market returns, using a model with intermediaries who are constrained exogenously through time-varying risk-aversion. In our model, market variance is stochastic and market maker wealth is endogenous, and we show that both features are needed to explicitly characterize the impact of intermediaries' risk bearing capacity on the variance risk premium.

Finally, with perfectly integrated financial markets, intermediaries' risk exposure and wealth should not affect equilibrium prices. Our findings are inconsistent with this hypothesis, and imply some segmentation of the option market consistent with a limits to arbitrage argument. See Shleifer and Vishny (1997), Gromb and Vayanos (2002), and Brunnermeier and Pedersen (2009) for prominent examples of such theories.

The remainder of the paper is organized as follows. Sections 2 and 3 present a reduced form regression analysis of the relation between the variance risk premium, inventory risk, and market maker wealth. Section 4 introduces the structural model. Section 5 discusses model implications. Section 6 implements and tests the model using return and option data. Section 7 concludes.

2 Inventory Risk, Market Maker Wealth and the Variance Risk Premium: A Reduced-Form Analysis

In this section we present a reduced-form regression analysis of the role of market makers in the determination of the variance risk premium. We first present the data. Subsequently we discuss the construction of the main variables of interest used in the regression analysis. We then formulate our hypotheses regarding the expected sign of the determinants of the variance risk premium in a regression analysis. Finally we briefly discuss additional control variables used in the regression.

2.1 Data

In our regression analysis, we focus on S&P 500 index options (SPX options), the most liquid contracts providing direct exposure to market variance. SPX options trade exclusively on the Chicago Board Option Exchange (CBOE). To construct the aggregate inventory of CBOE market makers, we rely on the Market Data Express Open/Close database. We obtain daily end user (non market maker) order flow for SPX options between January 1, 1996 and December 30, 2011. This data provides daily open/close buy and sell data for firms and customers. On each day, we compute the difference between the sum of the total buys and sells for these two groups combined for each contract, which corresponds to end users' net demand for that contract on that day. Because SPX options are European, the time series of market makers' inventory for each contract can be computed by summing up the negative of daily end users' net demand over time, starting from the first day the contract is quoted. We do not include LEAPS (options with more than one year to maturity) and start the sample period on January 1, 1997 to avoid biases in inventory measurement.[3]

For SPX option prices, we rely on end-of-day data from OptionMetrics between January 1, 1997 and December 30, 2011. We define the price of each contract to be the bid-ask midquote. We filter out contracts that have moneyness (spot price over strike price) less than 0.8 and larger than 1.2, contracts with a midquote less than 3/8, contracts with implied volatility less than 5% and greater than 150%, and contracts with less than ten days to maturity. We estimate maturity-specific interest rates by linear interpolation using zero coupon Treasury yields. The dividend yield is obtained from OptionMetrics.

We then merge the inventory data with the OptionMetrics database. The final sample contains more than one million quotes for SPX puts and calls over the 1997-2011 period.

To construct the variance risk premium, we need daily estimates of realized variance. To this end, we obtain high frequency data for S&P 500 index futures from Tickdata starting on January 1, 1997 and ending on September 30, 2012.[4] We construct daily measures of average realized variance following Zhang, Mykland, and Aït-Sahalia (2005).

We now discuss the variables that are our main focus in the regression analysis.

2.2 The Variance Risk Premium

The variance risk premium captures the difference between the physical and risk-neutral market variance. At time t the (annualized) variance risk premium with a T-day horizon is given by

$$VRP_{t,T} \equiv RV_{t,T} - RNV_{t,T}, \tag{2.1}$$

where $RV_{t,T} \equiv E_t^P[\frac{1}{T/365}\int_t^{t+T/365} V_s ds]$ denotes the expected integrated physical variance and $RNV_{t,T} \equiv E_t^Q[\frac{1}{T/365}\int_t^{t+T/365} V_s ds]$ is the expected integrated risk-neutral variance. Suppose that we want to obtain a model-free estimate of the one-month variance risk premium on day t, that is $VRP_{t,30}$.

[3]Note that we do not include the 1996 data. To correctly measure inventory for a given option, the full time series of end users' order flows for that option must be observed. With a maximum maturity of one year in the sample, we do not have all necessary data for some options quoted during 1996, which were issued in 1995.

[4]For some of the empirical tests, we need estimates of the realized variance up to September 30, 2012 to construct a measure of the 9-month ex-post realized variance up to December 30, 2011.

The first step is to compute $RV_{t,30}$. As in Carr and Wu (2009) and Egloff, Leippold, and Wu (2010), we proxy expected physical variance by ex-post realized variance

$$RV_{t,30} = \frac{365}{30} \cdot \left(\sum_{i=1}^{30} RV_{t+i-1} \right). \tag{2.2}$$

To measure the expected integrated risk-neutral variance, we follow Britten-Jones and Neuberger (2000) and Bollerslev, Tauchen, and Zhou (2009). We compute $RNV_{t,30}$ from a portfolio of SPX call options as

$$RNV_{t,30} = \frac{365}{30} \cdot \left(2 \cdot \int_0^\infty \frac{C(t, 30, Ke^{-r \cdot 30/365}) - C(t, 0, K)}{K^2} dK \right), \tag{2.3}$$

where $C(t, T, K)$ is the price of a call option observed at time t with T days to maturity and strike price K, and r is the risk-free rate. We evaluate (2.3) using the trapezoidal rule. The one-month variance risk premium on day t is given by $VRP_{t,30} = RV_{t,30} - RNV_{t,30}$.

Table 1 presents daily averages for implied volatility, vega, days to maturity, number of quotes, and volume. We also report the yearly averages of the one-month variance risk premium. Market implied volatility is 22.28% on average in our sample. Confirming existing studies (see, among others, Bakshi and Kapadia, 2003; Vilkov, 2008; Carr and Wu, 2009), the variance risk premium is robustly negative for every year in the sample. On average in our sample, the risk-neutral variance exceeds the realized variance by 2.13%.

Figure 1 plots the S&P 500 index in the top panel, and the one-month variance risk premium, expressed in percentages, in the middle panel. As expected, the variance risk premium varies more during periods of high uncertainty, such as the financial crisis. Prior to 2008, the variance risk premium is mostly negative and relatively stable, and it is especially small and stable between 2003 and 2007. To investigate if the large fluctuations in the variance risk premium during the financial crisis are induced by measurement errors, we plot the weekly averages of daily gains and losses of delta-hedged near-the-money options in the bottom panel of Figure 1. This exercise is motivated by Bakshi and Kapadia (2003), who show that the gains and losses from delta-hedged positions in options are informative about the variance risk premium. For a given option f_t^j, the daily dollar gains and losses from delta-hedging it from day $t-1$ to t are given by

$$\Delta Hedge_t^j \equiv f_t^j - \left(\Delta_{t-1}^j S_t + (f_{t-1}^j - \Delta_{t-1}^j S_{t-1}) \cdot (1 + r\Delta t) + \Delta_{t-1}^j S_{t-1} \cdot q\Delta t \right), \tag{2.4}$$

where $\Delta_t^j \equiv \frac{\partial f_t^j}{\partial S_t}$ is option j's delta, S_t denotes the value of the S&P 500, q is the dividend yield, and the time-step Δt is 1/365. Each week, we average the daily $\Delta Hedge_t^j$ for all options with

$0.98 \leqslant S_t/K^j \leqslant 1.02$ to obtain the weekly average gain and loss. Interestingly, the large fluctuations in $VRP_{t,30}$ during the crisis period are also readily apparent from the time series of delta-hedged gains and losses.

2.3 Inventory Risk and Market Maker Wealth

On average, approximately 273 SPX calls and puts with distinct moneyness and maturity are quoted every day. To assess market makers' inventory across contracts, Table 2 reports the daily average of implied volatility, inventory, and delta-hedged gains and losses for different moneyness and maturity categories. Note that market makers' positions are consistently negative across moneyness and maturity. Market makers are short approximately one hundred thousand contracts on a daily basis.

In our analysis, inventory risk captures the aggregate exposure of market makers' inventory of index options to market volatility. At any time t, it is defined by

$$InvRisk_t \equiv \sum_j \Phi_t^{MM,j} \cdot Vega_t^j, \tag{2.5}$$

where $\Phi_t^{MM,j}$ denotes market makers' inventory for option j, and $Vega_t^j \equiv \frac{\partial f_t^j}{\partial \sqrt{V_t}}$ denotes the option vega.[5] The aggregate exposure of market makers to market variance is the sum across all contracts of their inventory times vega. Inventory risk is highly informative about market makers' exposure, because $1\% \times InvRisk$ is the response of inventory in dollar terms to a 1% increase in market volatility. Equation (2.5) indicates that inventory risk is signed. Because vega is always positive, it is the commonality in inventory across contracts that determines the sign of inventory risk. At times when intermediaries act as net sellers and $\Phi_t^{MM,j} < 0$ for most j, we have $InvRisk_t < 0$. In contrast, inventory risk is positive when market makers hold long positions on average.

Figure 2 plots the VIX in the top panel, and the dynamic of inventory risk in the bottom panel. No clear correlation is apparent between VIX and inventory. Consistent with Table 2, Figure 2 indicates that market makers' exposure to S&P 500 volatility is negative most of the time except during the financial crisis and at the start of 2011. Through their inventory, option market makers carry billions of dollars in risk exposure to market variance.

To measure the changes in market maker wealth over time, we first compute their daily profits

[5] In our implementation we compute inventory risk on each day using Black-Scholes vega. See Carr and Wu (2007) and Trolle and Schwartz (2009) for examples of other studies that use Black-Scholes vega as a proxy for option vega in a stochastic volatility setup.

and losses from carrying hedged inventory:

$$P\&L_t \equiv \sum_j \Phi_{t-1}^{MM,j} \cdot \Delta Hedge_t^j. \tag{2.6}$$

where $\Delta Hedge_t^j$ satisfies (2.4). At the end of each day, market makers' aggregate daily profit and loss is the sum of lagged inventory times the delta-hedged gains and losses realized on that day across all contracts. Bid-ask spreads are another source of revenue for market makers. We estimate the daily bid-ask spread revenue earned by market makers as follows:

$$BA_t \equiv \sum_j \min(BO_t^j, SO_t^j) \cdot \left(BidAsk_t^j - 0.36\right), \tag{2.7}$$

where BO_t^j denotes end users' buy orders for option j, SO_t^j denotes end users' sell orders, $BidAsk_t^j$ denotes the option bid-ask spread, and \$0.36 is the transaction fee charged to dealers per contract traded.[6] In our empirical analysis, we define changes in wealth as the sum of (2.6) and (2.7), $\Delta W_t = P\&L_t + BA_t$. Our definition of market makers' wealth is similar to the measure used in Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes (2010) to proxy NYSE specialists' revenues. However, our measure differs from theirs to account for delta-hedging of inventory, which is adopted by most option intermediaries.

Figure 3 plots the daily profits and losses from market makers' delta-hedged inventory in the top panel, the bid-ask spread revenue in the middle panel, and the cumulative daily profits and losses in the bottom panel. Market makers face substantial risks. Their daily profits and losses fluctuate between 265 and -393 million dollars. Consistent with the idea that large positions in options imply substantial risks, the distribution of the daily delta-hedged gains and losses is highly leptokurtic with an excess kurtosis of 73, and asymmetric with a skewness coefficient of -2.55. On aggregate, market makers face a 5% risk of losing 21 million dollars or more on any given day.

Market makers on average earn a profit from delta-hedging their inventory, which generates a positive trend in the bottom panel of Figure 3. In aggregate, market makers earn 9 million dollars monthly from their delta-hedged positions.

Finally, comparing the top panel with the middle panel of Figure 3 reveals the substantial impact market variance risk has on dealers' wealth. Sofianos (1995) investigates NYSE specialists' revenues. He finds that stock market makers on average lose money on their inventories, and that their wealth is almost entirely due to the bid-ask spread. This is in stark contrast with option intermediaries. A large portion of changes in option market makers' wealth is driven by fluctuations in their delta-

[6] This fee includes \$0.33 charged by the CBOE and \$0.03 charged by the Options Clearing Corporation for clearing costs.

hedged inventory. In our sample, the absolute value of $P\&L_t$ is on average 1.5 times bigger than BA_t.

We have established that the representative SPX market maker faces substantial variance risks from carrying large inventories. Option dealers often trade among themselves in order to manage these risks. However, if end users have large net exposure to market variance, this will also be the case for SPX market makers in the aggregate. Because of market makers' large exposure to market variance, it is to be expected that part of their required compensation is embedded in the variance risk premium.

2.4 Methodology and Testable Hypotheses

Our benchmark empirical analysis uses the log-variance risk premium from Carr and Wu (2009), $LogVRP_{t,T} \equiv \ln(RV_{t,T}/RNV_{t,T})$. The distributions of the two variance measures are positively skewed, and the log specification alleviates the impact of extreme values.

Our methodology is adapted from Bollen and Whaley (2004). We regress daily changes in the variance risk premium against the explanatory variables and lagged changes in the dependent variable. For the log specification this gives:

$$\begin{aligned}\Delta LogVRP_{t,T} = Intercept + \beta_1^{Inv} InvRisk_{t-1} + \beta_2^{Inv}\left(\Delta W_t \cdot InvRisk_{t-1}\right) \\ +\beta^c Control_t + \beta^{VRP}\Delta LogVRP_{t-1,T} + \varepsilon_t,\end{aligned} \tag{2.8}$$

where $\Delta LogVRP_{t,T} \equiv LogVRP_{t,T} - LogVRP_{t-1,T}$.

If the variance risk premium captures part of dealers' required compensation, we expect a positive relation between lagged inventory risk and changes in the variance risk premium. The more negative (positive) the inventory risk, the more negative (positive) the variance risk premium, which implies positive returns for dealers. Therefore, we predict that lagged inventory risk should positively impact changes in the variance risk premium, or $\beta_1^{Inv} > 0$.

As their wealth decreases, dealers will require a higher compensation, which will in turn affect the variance risk premium. Because the impact on the variance risk premium will depend on the sign of the market maker's exposure, one must control for inventory risk when analyzing the relation between wealth and the variance risk premium. To capture how dealers dynamically influence the variance risk premium as their wealth fluctuates, we interact contemporaneous changes in wealth with lagged inventory risk to control for the lagged exposure. When market makers experience losses and they are negatively exposed to market variance, $\Delta W_t \cdot InvRisk_{t-1}$ is positive. Given the sign of market makers' exposure, higher compensation is associated with a decrease in the variance risk premium. Thus, the interaction of wealth with inventory risk should be negatively related to the variance risk premium. A similar argument also indicates a negative relation when inventory risk is positive and market maker wealth decreases. We therefore predict $\beta_2^{Inv} < 0$.

Motivated by existing studies, we also include a series of contemporaneous control variables in the regression. We now briefly discuss these control variables.

2.5 Additional Control Variables

Carr and Wu (2009) show that part of the variation in the variance risk premium is contemporaneously related to market index returns. To control for this effect we include the log return on the S&P 500 index, denoted by $S\&P500LogRet_t$, in the regression.

Egloff, Leippold, and Wu (2010), Todorov (2010), and Aït-Sahalia, Karaman, and Mancini (2014), among others, study the impact of jumps on the variance risk premium. We follow Cremers, Halling, and Weinbaum (2014) and construct an aggregate jump factor, $JumpFactor_t$.[7] By construction the jump factor has zero market delta, zero vega, and positive gamma, and thus captures the large fluctuations in the S&P500 index.

Bollen and Whaley (2004) document the effect of net buying pressure on option implied volatility. Through its impact on implied volatility, net buying pressure may also impact the variance risk premium. To disentangle the effect of inventory risk and net buying pressure on the variance risk premium, we also include Bollen and Whaley's net buying pressure variable which we denote $NetByingPressure_t$.[8]

Buraschi, Trojani, and Vedolin (2014) establish that disagreement among investors affects the variance risk premium. Empirically, dispersion of analyst forecasts is often used to gauge investors' disagreement, but this measure is not available at the daily frequency. We use unexpected changes in S&P500 index trading volume as a proxy for disagreement. Every day we calculate the difference between S&P 500 index volume on that day and the average volume over the past 90 trading days. We denote this variable $Disagreement_t$.

[7] On each day, we calculate the returns on two zero-beta at-the-money SPX straddles with maturities T_1 and T_2 with $T_1 < T_2$. We choose T_1 to be between fifteen days and one month, and T_2 between one and two months. Denote the returns on these two straddles by r_t^{S1} and r_t^{S2}. These daily returns are then combined such that $JumpFactor_t \equiv r_t^{S1} - \left(\frac{Vega_t^{S1}}{Vega_t^{S2}}\right) \cdot r_t^{S2}$ where $Vega_t^{S1}$ and $Vega_t^{S2}$ denote the vegas of the two straddles.

[8] The net buying pressure variable is obtained by summing the delta-weighted order imbalances across all contracts. It is calculated as $\sum_j \left(BO_t^j - SO_t^j\right) \cdot \frac{abs\left(\Delta_t^j\right)}{Volume_t}$ where $Volume_t$ is the aggregate volume of SPX options, and $abs(.)$ denotes the absolute value.

3 Regression Results

This section presents the results from the regression analysis. We first discuss the benchmark regression results. We emphasize how results during the financial crisis differ from results for the entire sample period. Finally we highlight our results for the term structure of variance risk premia and we present results from robustness exercises.

3.1 Explaining the Time Variation in the Variance Risk Premium

Table 3 presents the estimates based on equation (2.8), which regresses daily changes in the log-variance risk premium on inventory risk and market maker wealth. We also include the control variables discussed above. We report the results for the full sample in columns (1) and (2). In columns (3) to (6), we also report results for two sub-samples of similar length, 1997-2004 and 2005-2011. This allows us to assess the impact of the financial crisis on the regression results. Note that each variable is standardized to have unit variance in order to facilitate the interpretation of the coefficients. We report the Newey-West p-value with 8 lags to capture autocorrelation in the residuals.

Columns (2), (4), and (6) in Table 3 establish the importance of inventory risk and market maker wealth for explaining the variance risk premium. Both variables are statistically significant with the anticipated sign. In the two subsamples, including these variables increases the adjusted R-square by 9% relative to the results in columns (3) and (5). Given an average adjusted R-square of 42%, this corresponds to a $0.09/0.42 = 21\%$ increase in explanatory power.

Interestingly, the impact of inventory risk is greater for the sample period that includes the financial crisis. A one standard deviation decrease in inventory risk leads to a 1.20% decrease in the variance risk premium. When inventory risk equals its 2005-2011 sample average, a one standard deviation decrease in market maker wealth is associated with a 2.29% decrease in the variance risk premium. The impact of inventory risk is magnified when market makers experience dramatic losses. Conditioning on the 95^{th} percentile of the distribution of market maker loss, which corresponds to $\Delta W_t =$-$72,898,000$ million dollars, a one standard deviation decrease in inventory risk results in a 6% decrease in the log-variance risk premium the next day.

The relation between S&P 500 log-returns and the variance risk premium is strongly significant. The variance risk premium tends to decrease when index returns decrease. In columns (1), (3), and (5) in Table 3, S&P 500 log returns and lagged changes to the variance risk premium jointly account for 80% of the explanatory power.[9]

[9] Bivariate regressions of changes in the log variance risk premium on log index returns and lagged changes produce an adjusted R-square of 30% on average for the three sample periods (the full sample and two subsamples).

Aggregate jumps are negatively related to changes in the variance risk premium. By construction, the returns to the jump factor are high when the S&P 500 drops sharply. The estimate thus suggests that large negative jumps result in a more negative variance risk premium. This is consistent with Todorov's (2010) analysis of the impact of market jumps on the variance risk premium.

Bollen and Whaley (2004) find that net buying pressure increases implied volatility. Through its effect on implied volatility, high buying pressure should therefore result in a lower variance risk premium. Table 3 indicates that net buying pressure is indeed negatively related to the variance risk premium, but the relation is significant only for one of the subsamples in column (5).

Consistent with Buraschi, Trojani, and Vedolin (2014), the loadings estimated on disagreement are consistently negative. When investors' disagreement increases, the difference between realized and implied volatility tends to become more negative.

3.2 The Term Structure of Variance Risk Premia

The term structure of variance risk premia is a topic of substantial recent interest. It has been studied in Egloff, Leippold, and Wu (2010), Aït-Sahalia, Karaman, and Mancini (2014), and Dew-Becker, Giglio, Le, and Rodriguez (2015) among others. To quantify the impact each variable has on the term structure of variance risk premia, we construct measures of the variance risk premium for various horizons. In addition to the one-month horizon, we analyze four additional horizons: 60, 90, 180, and 270 days.

Table 4 reports results obtained for the log variance risk premium using the full sample. Several interesting findings emerge. The effect of inventory risk and fluctuations in market maker wealth is robust across horizons. Interestingly, the estimated coefficients display a term structure effect. For most variables, the magnitude of their impact on the variance risk premium decreases as the horizon increases. The effect of inventory risk and market maker wealth is most prominent for short-term variance risk premia. Aït-Sahalia, Karaman, and Mancini (2014) conclude that investors' fear of a market crash is mostly captured in short-term variance risk premia. Consistent with Aït-Sahalia et al., Dew-Becker, Giglio, Le, and Rodriguez (2015) also find that the price of variance risk is mostly negative at short horizons. Because inventory risk and market maker wealth matter the most for short-term variance risk premia, our results complement both of these studies.

Relative to the other variables, for horizons of sixty days or more, the impact of inventory risk and market maker wealth exceeds that of aggregate jumps, but it is smaller than the impact of index returns.

3.3 Robustness

So far, we have used future realized variance to proxy expected physical variance. We now investigate the robustness of our results when we instead use the forecast of a predictive model as an estimate of the expected physical variance. We adopt a HAR-RV dynamic based on Corsi (2009) to model realized variance. Using rolling windows of 252 observations, we estimate the model for each maturity on each day. We then use the one-step-ahead model forecast as an estimate of $RV_{t,T}$. Appendix A provides further details.

Panel A of Table A.1 in the online Appendix presents the average of the daily parameter estimates, the p-values, and the R-squares. The high p-values indicate that it is difficult to precisely estimate all parameters, but the high R-squares demonstrate the model's ability to forecast future market variance. Using the model prediction for $RV_{t,T}$, we construct measures of the variance risk premium for each horizon. Panel B of Table A.1 presents descriptive statistics on the variance risk premium implied by the HAR-RV dynamic.

Based on these variance risk premia, we regress the changes in the log variance risk premium on the explanatory variables for each horizon. Details on the regressions are provided in Table A.2 in the online Appendix. The impact of inventory risk and market maker wealth is robust to the computation of expected physical variance. The adjusted R-squares range from 40% to 49%. In Table 5, we present the average coefficients, p-values, and adjusted R-squares across horizons. The first column reports the results from Table 4 and the second column reports the averages based on Table A.2. Results are clearly very similar.

In Table 6, we assess the robustness of our empirical analysis when the variance risk premium is measured as $RV_{t,T} - RNV_{t,T}$, where $RV_{t,T}$ is once again proxied by future realized variances, as in the results for the log variance risk premium in Table 4. The impact of inventory risk and fluctuations in market maker wealth is once again robust. Univariate regressions of $\Delta LogVRP_{t,T}$ on $\Delta VRP_{t,T}$ yield a factor loading of 10 on average across horizons. Multiplying the parameter estimates in Table 6 by 10 indeed brings the results close to those of Table 4. For instance, the estimates for the one-month horizon become 0.60 for inventory risk and -1.7 for the interaction variable, similar to those in Table 4.

Recall that our computation of inventory delta-hedged profits and losses relies on option midquotes. Because market makers carry net short positions on average, intermediaries trade at the ask price more often than they trade at the bid price. We now assess if this impacts the results. We use end-of-day ask prices to calculate intermediaries' delta-hedged profits and losses on each day. Based on these daily estimates we compute the daily changes in market maker wealth. Table A.3 in the online Appendix reports the regression results. On average across horizons, the coefficient for the interaction of inventory risk with changes in wealth is -2.15. This is very close to the average estimate of -2.13 obtained when using midquotes. We conclude that our results are robust to the measurement of market maker wealth.

3.4 The Financial Crisis

The index option market functions as an insurance market for market risk. A clientele of institutional investors primarily buys SPX options, which causes market maker inventory risk to be negative on average. On November 20, 2008, the VIX reached a high of 80.86%. Around the same time, end users were heavily shorting SPX options. These large net sell orders resulted in a significant increase in inventory risk during the month of November. On aggregate, market makers accumulated more than 2.5 billion dollars positive net exposure to market volatility. By November 20, CBOE market makers carried more than 538 billion dollars in long option positions, accounting for 18% of the total capitalization of SPX options at the time.[10]

To understand the risk and reward associated with the positive exposure to market volatility during the financial crisis, Table 7 presents descriptive statistics for delta-hedged near-the-money option returns. Because these options are close to the money, they are highly sensitive to market variance. For comparison, we report statistics for the full sample as well as the financial crisis.

The delta-hedged positions are usually negative but earned 6.26% per month during the crisis period. Thus, when market makers' exposure to market variance was positive, long positions in near-the-money options were profitable on average. Note that the risk exposure from these options is very high. During the financial crisis, the volatility of daily returns peaked at 95%, and its excess kurtosis was about 11. This further emphasizes that index option market makers take on substantial risks, which allows end users to hedge against and speculate on market volatility.

In summary, the evidence presented thus far supports the notion that part of the variance risk premium captures index option market makers' compensation for exposure to market variance. Motivated by these findings, we now develop a structural theoretical model with dynamic variance and a risk-averse representative market maker who endogenously quotes index options, which affects the variance risk premium.

[10]To obtain an estimate of the total market capitalization of SPX options, we multiply open interest by the midquote for each option series, and sum across all series.

4 A Model of Inventory, Market Maker Wealth, and the Variance Risk Premium

We consider a continuous-time economy in which the underlying source of uncertainty is driven by two independent Brownian motions Z^S and Z^V.[11] Market participants have a finite investment horizon T, and can invest in the market index S_t, which evolves according to

$$\frac{dS_t}{S_t} = (\mu - q)\, dt + \sqrt{V_t}\left(\sqrt{1-\rho_V^2} dZ_t^S + \rho_V dZ_t^V\right), \quad \text{with } S_0 \text{ known,} \tag{4.1}$$

where μ is the market premium, q is the dividend yield, and V_t is the market variance, which follows the CEV dynamic

$$dV_t = \kappa(\theta - V_t)dt + \delta V_t^\eta dZ_t^V, \quad \text{with } V_0 \text{ known,} \tag{4.2}$$

where θ denotes the unconditional variance, κ is the speed of mean reversion, δ is the volatility of volatility, and η determines the elasticity of variance.[12] In (4.1), ρ_V captures the correlation between the innovations to the market return and market variance. In addition to the market index, market participants can invest in a risk-free bond

$$\frac{dB_t}{B_t} = r dt, \quad B_0 = 1, \tag{4.3}$$

with constant interest rate r. The economy is endowed with a stochastic discount factor (SDF) which reflects aggregate preferences. As in the portfolio literature (see Detemple, Garcia, and Rindisbacher, 2003, 2005; Detemple and Rindisbacher, 2010; Elkamhi and Stefanova, 2011), the form of the SDF is exogenously given. This SDF follows

$$\frac{d\xi_t}{\xi_t} = -r dt - \phi_t^S dZ_t^S - \phi_t^V dZ_t^V, \quad \xi_0 = 1, \tag{4.4}$$

where ϕ_t^S and ϕ_t^V are the market prices of risk. The (instantaneous) variance risk premium is the product of the market price of variance risk and the quantity of variance risk, that is $VRP_t = \phi_t^V \cdot (\delta V_t^\eta)$. Therefore, the variance risk premium and the price of variance risk are substitutes as ϕ_t^V can be replaced by $VRP_t/(\delta V_t^\eta)$ in (4.4).

The variance risk premium is determined through trading activity in index options. We denote

[11] The information available to agents consists of the trajectories generated by the two Brownian motions (the Brownian filtration $\mathcal{F}$). The underlying probability space is $(\Omega, \mathcal{F}, P)$, where P is the physical probability measure.

[12] Equation (4.2) nests a large variety of stochastic volatility models studied in the existing literature. For instance, Heston (1993) is obtained when $\eta = 1/2$, and Jones (2003) discusses the model with $\eta > 1$.

European index calls and puts by f_t^j where j identifies a particular option. Two types of agents interact in the option market. End users have an exogenous need to get exposure to index options. We denote end users' net demand for option j by $\Phi_t^{EU,j}$. To meet this demand, a representative market maker provides liquidity for index options. Since the physical dynamics of the market index, the market variance, and the SDF are all exogenous, in this framework the demand and supply for index options only affects ϕ_t^V and VRP_t.

In the next proposition, we present the pricing rule used by the market maker to quote options.

Proposition 1 *Given (4.1), (4.2), and (4.4), applying Ito's lemma to f_t^j implies the following dynamic for the price of option j under the P-measure*

$$\begin{aligned} df_t^j &= d\Delta Rep_t^j + \vartheta_t^j \cdot dF_t^V \\ &= d\Delta Rep_t^j + \vartheta_t^j \cdot \left(VRP_t dt + \delta V_t^\eta dZ_t^V\right), \end{aligned} \tag{4.5}$$

where ΔRep_t^j corresponds to the delta replication of f_t^j, $\vartheta_t^j \equiv Vega_t^j / \left(2\sqrt{V_t}\right)$ is the sensitivity of option j to the market variance risk factor F_t^V, $VRP_t \equiv \frac{1}{dt}\left(E_t^P[V_{t+dt}] - E_t^Q[V_{t+dt}]\right) = (\delta V_t^\eta) \cdot \phi_t^V$ is the (instantaneous) variance risk premium, and $\delta V_t^\eta dZ_t^V$ is the aggregate variance risk.

Proof. See Appendix B. ■

Under the Black-Scholes (1973) assumptions, index options can be perfectly replicated by holding the appropriate amount of the market index and the risk-free bond. When the market variance is stochastic, perfect replication is no longer achievable through the trading of S_t and B_t only. As a result, the price dynamic of index options can be decomposed into two components. As in Black and Scholes (1973), the first component, denoted $d\Delta Rep_t^j$, corresponds to the delta replication of f_t^j. In addition to $d\Delta Rep_t^j$, the entire cross-section of index options is affected by the market variance risk factor.

When $\Phi_t^{MM,j}$ denotes the market maker's inventory of option j, the market clearing condition for index options is

$$\Phi_t^{MM,j} + \Phi_t^{EU,j} = 0 \text{ for all } j \Rightarrow InvRisk_t = \sum_j \Phi_t^{MM,j} Vega_t^j = -\sum_j \Phi_t^{EU,j} Vega_t^j. \tag{4.6}$$

When end users' exposure to market volatility, $\sum_j \Phi_t^{EU,j} Vega_t^j$, does not cancel out across index options, neither does the market maker's inventory risk. Consequently, the representative market maker will be non-trivially dynamically exposed to the market variance risk factor.

For tractability reasons, we do not endogenize end users' trading motives for index options. Instead, building on the work of Amihud and Mendelson (1980) and Gârleanu, Pedersen, and

Poteshman (2009), we model the fluctuations of inventory risk exogenously. As apparent in Figure 2, the time series of inventory risk shares several common features with volatility, such as clustering, autocorrelation, and reversal. This observation is consistent with the market microstructure literature, which finds that stock market makers' inventory mean reverts (see Madhavan and Sofianos, 1998; Hansch, Naik, and Viswanathan, 1998; Naik and Yadav, 2003). To capture these statistical properties, we define the following mean-reverting dynamic for inventory risk

$$dInvRisk_t = \lambda(\alpha - InvRisk_t)dt + \psi V_t dt + \sigma InvRisk_t \left(\sqrt{1-\rho_{Inv}^2} dZ_t^S + \rho_{Inv} dZ_t^V \right), \qquad (4.7)$$

where $InvRisk_t$ satisfies (4.6), λ captures the speed of mean reversion, α captures the level of inventory risk, ψ captures the sensitivity of inventory risk to market variance, ρ_{Inv} measures its correlation with market variance innovations, and σ is the volatility parameter. Arguably, fluctuations in market variance should affect the aggregate net demand for index options and thus inventory risk. To account for this, we allow the dynamic (4.7) to depend on V_t and dZ_t^V.

We determine the variance risk premium through the maximization of the market maker's expected utility of terminal wealth. For a given admissible investment strategy $\tilde{\pi}_t \equiv \left(\pi_t^B, \pi_t^S, \{\pi_t^{f,j}\}\right)$, the market maker's self-financing wealth dynamic is

$$\frac{dW_t}{W_t} = \pi_t^B \cdot \frac{dB_t}{B_t} + \pi_t^S \cdot \left(\frac{dS_t}{S_t} + q dt \right) + \sum_j \pi_t^{f,j} \cdot \frac{df_t^j}{f_t^j}, \quad \text{with} \quad W_0 = w, \qquad (4.8)$$

where each π is expressed as a percentage of wealth, qdt accounts for the reinvestment of dividends, and w denotes the initial endowment. When w is low, the market maker is more financially constrained. The problem faced by the market maker can be written as

$$\max_{\tilde{\pi}_t} \; E^P[U(W_T)] \quad \text{subject to} \quad \begin{array}{l} \cdot \; (4.7) \\ \cdot \; (4.8) \text{ with } W_t \geqslant 0\text{: } t \in [0,T], \end{array} \qquad (4.9)$$

where $U(.)$ is the market maker's utility function. In this model, the representative market maker determines his trading strategy given market prices. Our objective is to invert this mapping and infer the variance risk premium in (4.5) that induces the market maker to clear the index option market.

5 Model Implications

We now discuss the most important model implications. First we discuss the structure of the variance risk premium. Then we characterize the optimal wealth of the market maker, and finally we present the model's risk-neutral dynamics.

5.1 The Structure of the Variance Risk Premium

The following proposition illustrates how the model variance risk premium depends on inventory risk and market maker wealth.

Proposition 2 *At time $t \in [0, T]$, if the market maker is myopic with $U(W_t) = \ln(W_t)$, the variance risk premium is given by*

$$VRP_t = \rho_V \delta V_t^{\eta} \left(Sharpe_t\right) + 0.5(1-\rho_V^2)\delta^2 V_t^{2\eta-0.5}\left(\frac{InvRisk_t}{W_t}\right), \tag{5.1}$$

where $Sharpe_t \equiv \frac{\mu - r}{\sqrt{V_t}}$ is the market Sharpe ratio, $InvRisk_t = -\sum_j \Phi_t^{EU,j} Vega_t^j$ captures the sensitivity of the market maker's inventory to the variance risk factor F_t^V defined in Proposition 1, and W_t is market maker wealth.

Proof. *See Appendix C.* ■

This proposition provides several insights. The decomposition in (5.1) splits up the variance risk premium into two components. The first component is a function of the market Sharpe ratio. Innovations in market variance are correlated with index returns. Consequently, the variance risk premium inherits the properties of the market Sharpe ratio. The greater $abs(\rho_V)$, the higher the dependence of the variance risk premium on $Sharpe_t$. Equation (5.1) provides a potential explanation for why we obtain statistically significant estimates when regressing changes in the variance risk premium on $S\&P$ 500 log returns.

When $abs(\rho_V) < 1$, index options cannot be perfectly hedged by trading S_t and B_t. Consequently, the market is incomplete from the market maker's perspective. The market maker then requires compensation in addition to $\rho_V \delta V_t^{\eta}\left(Sharpe_t\right)$. This additional premium is proportional to the ratio of inventory risk to wealth. Since $(1-\rho_V^2)\delta^2 V_t^{2\eta-0.5} \geqslant 0$ and $W_t \geqslant 0$, the variance risk premium is positively impacted by inventory risk. The more negative the exposure of the market maker to market variance, the lower the variance risk premium. Consequently, the model can explain the positive estimate obtained by regressing changes in the variance risk premium on market maker exposure to market variance.

Several studies have shown that the premium for market variance risk is negative on average (see, among others, Bakshi and Kapadia, 2003; Vilkov, 2008; Carr and Wu, 2009). Given (5.1), the variance risk premium is negative when

$$\frac{InvRisk_t}{W_t} < \frac{-\rho_V \left(Sharpe_t\right)}{0.5(1-\rho_V^2)\delta V_t^{\eta-0.5}}. \tag{5.2}$$

For the empirically relevant case $\rho_V < 0$ and $Sharpe_t > 0$, a sufficient condition for the variance risk premium to be negative is negative inventory risk. Since index option market makers typically have a negative exposure to market variance, the model provides a potential explanation for the negative variance risk premium found in existing studies.

As is apparent from the middle and bottom panels of Figure 1, the variance risk premium is occasionally positive. When inequality (5.2) is not satisfied, $\rho_V < 0$ and $Sharpe_t > 0$, a positive inventory risk exposure results in a positive variance risk premium. Hence, the model can also explain positive variance risk premiums when option market makers are positively exposed to market variance.

Finally, note that positive inventory risk does not result in a positive variance risk premium as long as (5.2) is satisfied.

5.2 Market Maker Optimal Wealth

Most classical inventory models assume that dealers have access to unlimited capital (see, among others, Ho and Stoll, 1981, 1983; Mildenstein and Schleef, 1983). Recently, Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2009) relax this assumption. In Brunnermeier and Pedersen (2009), market makers' limited funding capacity determines how much liquidity they provide. When market makers' margin requirements are close to the available capital, intermediaries provide less liquidity, which in turn affects price. Similar predictions are obtained by Gromb and Vayanos (2002). In our model, the intermediary's financial constraint also has important pricing implications.

In Proposition 2, inventory risk is normalized by the market maker's wealth. Consequently, inventory risk will matter most for the variance risk premium at low values of W_t. In contrast, the effect of inventory risk vanishes when the market maker is unconstrained, that is, when his wealth goes to infinity. This result explains the strongly significant estimates in the regression analysis for the interaction of inventory risk with changes in market makers' wealth.

In the model, the market maker's wealth is endogenous and if $U(W_t) = \ln(W_t)$ we have

$$W_t = 1/\left(\gamma\xi_t\right), \tag{5.3}$$

where γ is the shadow price of the market maker's financial constraint $W_0 = E^P[\xi_T W_T]$. Since the market maker's marginal utility is strictly increasing, γ is uniquely defined by (5.3) and the intermediary's financial constraint. Together, these results imply $\gamma = 1/W_0$, which gives $W_t = W_0/\xi_t$. Consequently, at a given point in time, the market maker's wealth is proportional to the ratio of his initial endowment to the SDF. Given (4.4), we can apply Ito's lemma to W_t to obtain the dynamics of the market maker's optimal wealth

$$dW_t = \left(r + \left(\phi_t^S\right)^2 + \left(\phi_t^V\right)^2\right) W_t dt + \phi_t^S W_t dZ_t^S + \phi_t^V W_t dZ_t^V. \tag{5.4}$$

The market maker's wealth dynamic is driven by the two aggregate shocks and their prices of risk. We now characterize the risk-neutral distribution of the market return, which is affected by market maker wealth and inventory risk through their impact on the price of variance risk.

5.3 Risk-Neutral Dynamics

In our empirical analysis, we proceed by estimating the model based on option data. Option valuation requires discounting the payoff at maturity under the risk-neutral measure using the risk-free rate. Therefore, all underlying processes need to be risk-neutralized.

Given the SDF (4.4), we have $dZ_t^S = d\tilde{Z}_t^S - \phi_t^S dt$ and $dZ_t^V = d\tilde{Z}_t^V - \phi_t^V dt$ where $\tilde{Z}_t^S$ and $\tilde{Z}_t^V$ are risk-neutral. Using this result in (4.1), (4.2), (4.7), and (5.4) characterizes the economy's dynamics under the pricing measure. We refer to Appendix D for a detailed discussion of the risk-neutral processes. We now discuss the model's implications for risk-neutral market variance and skewness.

Consider the market risk-neutral variance first, which follows the non-affine dynamic

$$dV_t = \kappa(\theta - V_t)dt - VRP_t dt + \delta V_t^{\eta} d\tilde{Z}_t^V, \tag{5.5}$$

where VRP_t satisfies Proposition 2. When end users' demand for index options increases, inventory risk decreases. A decrease in inventory risk implies a more negative variance risk premium on average. From (5.5), changes in the risk-neutral market variance are negatively impacted by the variance risk premium. An increase in end users' demand will therefore result in a higher risk-neutral variance. This model prediction is related to Bollen and Whaley (2004), who document that changes in the implied volatility of OTM index puts are positively affected by end users' net buying pressure. However, our model suggests that only the variance exposure of market maker's total risk exposure should impact changes in risk-neutral volatility.

The nonlinearities in the dynamics of the risk-neutral variance, inventory risk, and market maker wealth have interesting implications for risk-neutral market skewness. Under the pricing measure, the market maker's wealth and inventory risk jointly satisfy

$$dW_t = rW_t dt + \phi_t^S W_t d\tilde{Z}_t^S + \phi_t^V W_t d\tilde{Z}_t^V \tag{5.6}$$

$$\begin{aligned} dInvRisk_t = {} & \lambda(\alpha - InvRisk_t)dt + \psi V_t dt - \sigma InvRisk_t \left(\sqrt{1-\rho_{Inv}^2}\phi_t^S + \rho_{Inv}\phi_t^V\right) dt \\ & +\sigma InvRisk_t \left(\sqrt{1-\rho_{Inv}^2} d\tilde{Z}_t^S + \rho_{Inv} d\tilde{Z}_t^V\right). \end{aligned} \tag{5.7}$$

When $d\tilde{Z}_t^V > 0$, variance risk is high and market returns are low for the empirically relevant case $\rho_V < 0$. If inventory risk and the variance risk premium are both negative, high variance risk reduces market maker wealth since $\phi_t^V d\tilde{Z}_t^V < 0$ in (5.6). Lower wealth in turn implies a more negative ratio of inventory risk to wealth, which further decreases the variance risk premium. This feedback effect between the variance risk premium and the ratio of inventory risk to wealth amplifies the increase in risk-neutral variance in bad times. This mechanism also allows the model to generate substantial negative skewness, which is appealing given the challenge standard stochastic volatility models face in explaining the cross-section of out-of-the-money index puts.

Overall, the predictions delivered by the model are consistent with empirical stylized facts. In the next section, we estimate the model and quantitatively assess the importance of inventory risk and market maker wealth for the valuation of index options.

6 Model Estimation and Model Fit

In this section, we first describe our estimation methodology. Subsequently, we report on parameter estimates and model fit. Finally we use the estimated parameters to assess the economic impact of inventory risk and market maker wealth on index option prices. For ease of notation we henceforth refer to the inventory risk and wealth model as the IRW model.

6.1 Estimation Methodology

Several approaches are available to estimate stochastic volatility models. Aït-Sahalia and Kimmel (2007) and Jones (2003) use bivariate time series of returns and at-the-money implied volatility. Pan (2002) uses GMM to estimate the objective and risk-neutral parameters using returns and option prices. Christoffersen, Jacobs, and Mimouni (2010) adopt a particle filtering approach to estimate various alternatives to the Heston (1993) model based on returns and a large panel of option prices.

We adopt a two-step procedure to estimate the model. In a first step, we use index returns to filter the physical parameters of the market variance dynamic (4.2) along with the spot variance

V_t. In a second step, we take the filtered spot variances and the variance parameters under the physical measure as given, and we estimate the dynamic of inventory risk and market maker wealth using a large panel of SPX put prices. Both $InvRisk_t$ and W_t are latent variables in the model. However, to avoid overfitting we constrain $InvRisk_t$ to equal its observed value (2.5). In addition, we set market maker initial wealth to $W_0 = w$ at the beginning of every day. Based on these initial values and the filtered V_t, we use the results in Propositions 2 and Appendix D to simulate the economy and infer the dynamics (5.7)-(5.6) that are consistent with observed index option prices. For estimation purposes, we set the time-step Δt to 1/365 and we set the expected return net of the dividend yield $\mu - q$ to its sample average $\frac{1}{T-1}\sum_{t=2}^{T}\left(S_t - S_{t-1}\right)/S_{t-1} \times 365$, where T denotes the last day in the sample. We now describe the two steps in more detail.

Step 1: Filtering the Variance Dynamic Using S&P 500 Returns

We need to estimate the structural parameters $\Theta^V \equiv \{\kappa, \theta, \delta, \rho_V, \eta\}$ in (4.2) along with the vector of spot variances $\{V_t\}_{t=1,2,...,T}$. To this end, we adopt the particle filtering algorithm (PF henceforth). The PF offers a convenient approach for estimating stochastic volatility models. It was recently used by Johannes, Polson, and Stroud (2009), Christoffersen, Jacobs, and Mimouni (2010), and Malik and Pitt (2011), among others.

Let $\left\{V_t^j\right\}_{j=1}^N$ denote the smooth resampled particles where N defines the number of particles which is set to $10,000$. Using the algorithm described in Appendix E, each day we estimate the likelihood of observing S_{t+1} given V_t^j and S_t, and denote it $\tilde{P}_t^j\left(V_t^j, \Theta^V\right)$. Based on the likelihood of each particle calculated on each day, we use the MLIS criterion to estimate Θ^V

$$\hat{\Theta}^V = \arg\max \sum_{t=1}^{T} \mathcal{L}_t, \tag{6.1}$$

where $\mathcal{L}_t \equiv \ln\left(\frac{1}{N}\sum_{j=1}^{N}\tilde{P}_t^j\left(V_t^j, \Theta^V\right)\right)$ is the daily model log-likelihood. On day t, the filtered spot variance is obtained by averaging the smooth particles

$$\hat{V}_t = \frac{1}{N}\sum_{j=1}^{N} V_t^j. \tag{6.2}$$

Next, we discuss the estimation of inventory risk and market maker wealth based on SPX options.

Step 2: Estimating Inventory Risk and Market Maker Wealth

The model does not allow for closed-form solution for option prices. Consequently, we rely on Monte-Carlo methods for estimating the inventory risk and market maker wealth dynamics embedded in SPX options. Taking $\hat{\Theta}^V$ and $\left\{\hat{V}_t, Inv\hat{R}isk_t\right\}_{t=1,2,\ldots,T}$ as given, where $Inv\hat{R}isk_t$ corresponds to (2.5), we estimate $\Theta^{Inv} \equiv \{\lambda, \alpha, \psi, \sigma, \rho_{Inv}\}$ and w by minimizing the sum of implied volatility squared errors (SIVSE)

$$\left\{\hat{\Theta}^{Inv}, \hat{w}\right\} = \arg\min \sum_{j,t}^{N_J} \left(IV_{j,t} - IV_{j,t}^M \left(\Theta^{Inv}, w, \hat{\Theta}^V, \hat{V}_t, Inv\hat{R}isk_t\right)\right)^2, \tag{6.3}$$

where N_J is the total number of observations, $IV_{j,t}$ is option j's implied volatility on day t, and $IV_{j,t}^M$ denotes model-implied volatility. We use the Black-Scholes model to calculate implied volatilities for both market and model prices. When calculating model prices, we use the algorithm described in Appendix F, using $10,000$ Monte-Carlo paths.

6.2 Parameter Estimates

We first discuss the estimates of the parameters characterizing the variance dynamic, and then the estimates associated with inventory risk and market maker wealth.

6.2.1 The Variance Dynamic

Panel A of Table 8 presents descriptive statistics based on daily S&P 500 returns. The average market return net of dividend yield in the 1997-2011 sample period is 5.86%. This low average return is partly due to the sharp drop in the S&P 500 index during the crisis period (see Figure 1). The sample variance is 4.58% annually, which corresponds to a 21% average volatility.

Panel B of Table 8 reports the estimated parameters for the CEV dynamic (4.2). The sample MLIS is $11,746$. The estimated θ is close to the sample variance in Panel A. The estimate of ρ_V is large and negative. In the model, a large and negative ρ_V is important to generate sufficient negative skewness in the S&P 500 return distribution. Large fluctuations in the market variance (i.e. high volatility of dV_t) helps the model generate additional variability in the return process. This is required to capture the kurtosis of the S&P 500 return distribution. Based on the parameter estimates in Table 8, the variance of dV_t is on average equal to $d\langle V, V\rangle_{V=\theta} = \delta\theta^\eta dt = 0.06dt$. Thus, the high volatility of volatility δ is required by the model to generate enough variation in the variance given $\eta = 0.90$. Finally, note how the index data require a slow speed of mean reversion in the variance. The 2.91 estimate corresponds to a daily variance persistence of $1 - \kappa/365 = 0.99$.

For comparison, Panel C of Table 8 reports the parameters obtained for the Heston model, which imposes $\eta = 1/2$. The model MLIS is close to the likelihood obtained for the CEV dynamic. However, the two models require different structural parameters to explain the data. For instance, note the differences in mean reversion speed and unconditional variance. The Heston dynamic requires a higher speed of mean reversion but has a lower long-term variance. Moreover, the two models also display different volatility of volatility. This is partly driven by the difference in η between the two models. Because $\sqrt{\theta} = \sqrt{0.0408} = 0.20$ in the Heston model is substantially higher than $\theta^{\eta} = 0.0458^{0.90} = 0.06$ in the CEV model, the Heston model requires a smaller volatility of volatility parameter than the CEV model to fit the S&P 500 return distribution. For the Heston model, on average $d\langle V, V\rangle_{V=\theta} = \delta\sqrt{\theta}dt = 0.04dt$, slightly lower than the CEV model.

Figure 4 plots the time series of filtered spot volatilities $\sqrt{\hat{V}_t}$ for both models, annualized and expressed in percentages. As expected, these time series of physical spot volatilities share common features with the VIX in Figure 2. Comparing the two models, the filtered spot volatilities display similar patterns most of the time. However, during the crisis the filtered spot variances from the CEV model are substantially higher than the ones from the Heston model. The most likely explanation is the higher elasticity of the CEV model, which requires a higher level of spot volatility to generate sufficient kurtosis during the crisis period.

6.2.2 The Inventory Risk and Market Maker Wealth Dynamics

We use the OptionMetrics volatility surface data for calibrating the inventory risk and market maker wealth parameters. To speed up estimation, we restrict attention to put options observed on the first Wednesday of each month with moneyness between 0.9 and 1.1, and with 2, 3, and 6 months to maturity. The resulting option sample for the 1997-2011 period consists of $6,292$ put contracts.

Panel A of Table 9 reports the estimated coefficients for $\hat{\Theta}^{Inv}$ obtained by minimizing the sum of squared errors (6.3). The estimate of the mean reversion parameter is 10.73. A high speed of mean reversion is necessary to explain the abrupt reversal in inventory risk observed during the crisis period, as indicated in the bottom panel of Figure 2. The daily inventory persistence is $1 - \lambda/365 = 0.97$. This persistence implied by options is close to the persistence of 0.98 obtained when fitting an AR(1) model on the spot inventory risk measures. The high persistence in the variance risk exposure of index option market makers is in line with the evidence in Madhavan and Smidt (1993), who show that the inventory of NYSE specialists can deviate from its long-run mean for several weeks.

As apparent from the bottom Panel of Figure 2, inventory risk is negative most of the time. In the structural model, this stylized fact is captured by $\hat{\alpha}$ which is large and negative. Note however that the unconditional expectation of inventory risk also depends on $\hat{\psi}$. Interestingly, the loading of inventory risk on the lagged spot variance is positive. Inventory risk thus tends to increase with market uncertainty. This result is consistent with time series regressions of daily changes in inventory risk on the VIX, which also yields a positive factor loading.

The estimate of the inventory risk volatility parameter is 16.55%, which is of a similar order of magnitude than the volatility of volatility parameter for the index. The instantaneous correlation between inventory risk and market variance is nearly zero. While changes in inventory risk increase conditionally with market variance through ψ, the estimate of ρ_{inv} indicates that inventory risk is contemporaneously nearly independent of market variance innovations.

The w parameter captures the dealer's initial wealth. The estimated wealth level is approximately 440 million dollars. This estimate is comparable in magnitude to the daily delta-hedged profits and losses documented in Figure 3. Quantitatively, it represents approximately 4 years of cumulative daily profits and losses.

6.3 Model Fit

We first discuss option fit for the sample used to estimate the models. Subsequently we address model fit in a larger sample, as well as the model's performance in a more stringent exercise that requires forecasting of the state variables.

6.3.1 In-Sample Fit

For the OptionMetrics volatility surface data used in estimation, which consist of $6,292$ put contracts for 1997-2011, the sum of squared pricing errors for the IRW model using the optimized parameters in Panel A of Table 9 is 6.68. To benchmark this performance, we fit the Heston (1993) model on option prices using a similar methodology but imposing $VRP_t = h \cdot V_t$, where h is a constant to be estimated. This specification for the variance risk premium is consistent with most of the existing literature, including Heston (1993), Bates (2000), and Pan (2002).

Panel B of Table 9 indicates that the variance risk premium parameter for the Heston model is -1.08. The fit obtained using the Heston model is not quite as good as the fit of the IRW model. The Heston sum of squared implied volatility errors is approximately 20% higher than that of the IRW model.

6.3.2 Out-of-Sample Fit

Because option prices are not available analytically, model estimation is rather time-intensive, and we choose the option sample to make estimation feasible. Now we compare the performance of the

IRW model with that of the Heston model using a much larger sample. We use all put options available in OptionMetrics for 1997-2011 with moneyness between 0.9 and 1.1, and with 2, 3, and 6 months to maturity. The resulting sample consists of $131,838$ observations, considerably larger than the sample used in estimation.[13] Note that this exercise is also beneficial because it is out-of-sample. Less parsimonious models often obtain better in-sample fit at the cost of poor out-of-sample performance.

To evaluate model performance, we compute the percentage implied volatility RMSE (IVRMSE) defined as

$$IVRMSE \equiv \sqrt{\frac{1}{N_J}\sum_{j,t}^{N_J}\left(IV_{j,t} - IV_{j,t}^{M}\right)^2} \times 100, \tag{6.4}$$

where $IV_{j,t}^{M}$ is the implied volatility based on the model price.

Table 10 presents the results for the IRW and Heston models. We report the IVRMSE for all contracts averaged by year in Panel A, by moneyness in Panel B, and by maturity in Panel C. The average yearly IVRMSE for the IRW model is 3.14%, which is satisfactory, especially because the sample includes the financial crisis.

The fit provided by the two models is not very different, but the IRW model dominates the Heston model, especially in the second part of the sample. The Heston model performs relatively well in 1997-1999, outperforming the IRW model by 0.43%. In contrast, the IRW does relatively well in 2009-2011, outperforming the Heston model by 1%. Since 2003, the IRW model has outperformed the Heston by more than 0.91%.

Panel B of Table 10 indicates that the inventory model achieves a better fit for ATM puts relative to the Heston model. ATM options are most sensitive to changes in the variance risk premium. This suggests that the Heston model, which uses the specification $VRP_t = h \cdot V_t$, is unable to adequately capture variations in the variance risk premium. We further investigate this by regressing the daily implied volatility root mean squared errors against the empirical proxy of the one-month variance risk premium and the daily log-likelihood $\mathcal{L}_t$ of the physical returns. For the IRW model, we obtain

$$IVRMSE_t = \underset{(0.00)}{3.14} + \underset{(0.90)}{0.02} \cdot VRP_{t,30} - \underset{(0.01)}{0.27} \cdot \mathcal{L}_t + \varepsilon_t, \tag{6.5}$$

where the regressors are standardized, and the Newey-West p-values reported in parentheses are calculated with 8 lags. The adjusted R-square of the regression is 1.77%. Mispricing in the IRW model seems largely unrelated to the realization of the variance risk premium. For the Heston

[13] Approximately 12 put observations are available for each maturity on each day. Given the 3776 days of the sample period, and given that we consider three distinct maturities, it yields to a total of $131,838$ put observations.

model, we get

$$IVRMSE_t = \underset{(0.00)}{3.56} - \underset{(0.00)}{0.81} \cdot VRP_{t,30} - \underset{(0.75)}{0.02} \cdot \mathcal{L}_t + \varepsilon_t, \tag{6.6}$$

with an adjusted R-square of 17.89%. This result is striking. In contrast to the IRW model, a large part of the pricing error associated with the Heston model can be attributed to its inability to capture the fluctuations in the variance risk premium.

Panel B of Table 10 also indicates that the IRW model fits OTM puts better than the Heston model. This result is partly due to the feedback effects between the variance risk premium and the ratio of inventory risk to wealth. When the variance risk premium is negative, the ratio of inventory risk to wealth tends to decrease when volatility increases. This further reduces the variance risk premium and increases the level of the risk-neutral variance during bad times. Because a high risk-neutral volatility during market declines results in negative skewness of the index return distribution, this feedback effect improves the pricing of OTM puts.

Panel C of Table 10 also indicates that the IRW model achieves better pricing performance across maturities. The term structure of risk-neutral volatility is directly affected by the term structure of variance risk premia. This finding is consistent with the results in Table 4. Because inventory risk and market maker wealth matter for the term structure of variance risk premia, they are also important for the term structure of risk-neutral volatility.

Figure 5 plots the daily one-month variance risk premium and IVRMSE for both models. Panel A suggests that the IRW model can deliver a wide range of variance risk premia. In contrast, the one-month variance risk premium implied by the Heston model is always negative. This ability of the IRW model to generate substantial variation in variance risk premia is key to its improved performance.

Overall, these results strongly suggest that the IRW outperforms the Heston model in our sample period. We now turn to a more stringent out-of-sample exercise that uses one-day ahead forecasts of the latent state variables. For the IRW model, each day we forecast spot inventory and spot variance based on the dynamics (4.2)-(4.7). For the Heston model, each day we forecast the spot variance. We assess model fit based on these forecasts and the parameter estimates. Table 11 presents the results. The IRW continues to outperform the Heston model.

These empirical results suggest that accounting for inventory risk and market maker wealth is critical to understand and model variation in the variance risk premium and explain index option prices. Next we quantify the impact of changes in inventory risk and market maker wealth on SPX option prices.

6.4 Pricing Impact

Little is known in the existing literature about how market makers adjust option midquotes when they absorb large buy orders, which causes their exposure to market variance to become more negative and their inventory risk to decrease. Related to this, the magnitude of the impact of market makers' losses and gains on index option prices is also an open question.

Figure 6 addresses the impact of inventory risk and market maker wealth on quotes and prices. We plot the model-implied dollar sensitivity of SPX put options to a decrease in the state variables. We compute the model sensitivities $\partial P_t/\partial InvRisk_t$ and $\partial P_t/\partial W_t$ using the estimated parameters $\hat{\Theta}^V$, $\hat{\Theta}^{Inv}$, and $\hat{w}$, and set $r = 4\%$, $q = 0$, $S_t = 1183$, and $V_t = \hat{\theta}$. Inventory risk is set to $InvRisk_t = -9.03E + 09$, and market maker wealth is initialized to $W_t = \hat{w}$. Based on the resulting model sensitivities, we then compute the dollar response of each option as $\Delta InvRisk_t \cdot \partial P_t/\partial InvRisk_t$ and $\Delta W_t \cdot \partial P_t/\partial W_t$, and we plot the results across moneyness. The circles in Figure 6 depict the dollar response to an average decrease in the state variables. The diamonds depict the dollar response to a 90^{th} percentile decrease in each latent variable.

Figure 6 provides several insights. First, a decrease in inventory risk results in an increase in index option prices. This is consistent with the theoretical prediction in Proposition 2. When market makers' risk exposure decreases, they require a more negative variance risk premium, which translates into an increase in index option prices. Similarly, market makers' losses result in an increase in the price of index options. Interestingly, market maker wealth has the biggest impact on option prices.

Note also that the effect of inventory risk and market maker wealth on SPX options across strike prices is nonlinear, and is most prominent for at-the-money options. This is consistent with Table 2, which indicates that these options also make up most of market makers' inventory.

When all variables are set equal to their average values, the average decrease in inventory risk and the average wealth loss increases prices by between 10 and 50 dollars. Given an average option price of $6,500$ dollars, this corresponds to a 0.15%-0.77% daily increase in price.

During turbulent times, when market makers' aggregate loss is at its 90^{th} percentile, it causes a 150 dollar increase in price, which corresponds to a 2.31% daily increase. These results further highlight the non-trivial role of market maker wealth in the determination of index option prices through their effect on the variance risk premium.

7 Summary and Conclusions

We investigate how inventory risk and market maker wealth jointly determine the value of index options through their effects on the variance risk premium. We first conduct an exploratory regression analysis using daily data on aggregate market makers' index option positions at the CBOE. We regress the variance risk premium on measures of inventory risk and market maker wealth, and find that inventory exposure to market variance and changes in market makers' wealth explain the variance risk premium. A one standard deviation decrease in inventory risk causes a 1.2% decrease in the variance risk premium.

Motivated by these findings, we develop a structural model in which market variance is stochastic and a representative market maker with limited capital accumulates inventory over time by absorbing end users' net demand for index options. Starting from the market maker's optimal trading strategy, we derive an explicit formula linking the variance risk premium to inventory risk and market maker wealth. The model provides interesting insights on the structure and the composition of the variance risk premium.

Finally, we estimate the structural model using S&P 500 returns and option data. Overall, the model performs well, particularly during the financial crisis, and our findings suggest that accounting for inventory risk and market maker wealth may lead to more accurate pricing of index options. The estimation results confirm that changes in market maker wealth and inventory risk have a non-negligible impact on index option prices.

Several issues are left for future research. First, it would be interesting to develop and test the implications of alternative inventory risk dynamics for option valuation. Second, existing findings in the option literature suggest that extending the dynamics of the model, for instance by allowing for jumps in the prices, may result in a better model. Third, the model can be used to predict future option prices given estimates of the market maker's wealth. Finally, an investigation of the pricing implications of inventory risk and market maker wealth for other derivative markets would also be of significant interest.

Appendix

This appendix starts by presenting the strategy used to forecast for integrated physical variance. It then collects the proofs of the propositions and discusses the algorithms used for estimating the model.

A. Forecasting Expected Physical Variance

Suppose we want to estimate the T-days expected integrated physical variance on date t_0 that is $RV_{t_0,T}$. Using a rolling-window of 252 days with $t \in \{t_0 - 252, ..., t_0 - 1\}$, we run the following HAR-RV model in the spirit of Corsi (2009)

$$\begin{aligned}\ln(RV_{t,T}) &= a_0 + a_1 \ln(RV_{t-1,1}) + a_2 \ln(RV_{t-6,6}) + \\ &\quad a_3 \ln(RV_{t-30,30}) + a_4 \ln(RV_{t-60,60}) + a_5 \ln(RV_{t-90,90}) \\ &\quad + a_6 \ln(RV_{t-120,120}) + \varepsilon_{t,T}, \end{aligned} \tag{7.1}$$

where $RV_{t,T}$ is the T-days realized variance at time t. We can write the model in matrix form. We have $\ln(RV_{t,T}) = X_{t-1} \cdot A + \varepsilon_{t,T}$ where X_{t-1} contains the explanatory variables known on day $t-1$ and A is the matrix of parameters. Using the OLS estimate for $\hat{A}$ and setting the regressors to their values on the day t_0, the model prediction for $RV_{t_0,T}$ on that day is $\exp(X_{t_0} \cdot \hat{A} + \frac{\hat{\sigma}_\varepsilon^2}{2})$ where $\hat{\sigma}_\varepsilon^2$ is the variance of the residuals. We repeat this procedure on each day and for each horizon.

B. Proof of Proposition 1

For ease of notation, we define $a_t^P \equiv E_t^P[\frac{dV_t}{dt}]$ and $a_t^Q \equiv E_t^Q[\frac{dV_t}{dt}]$ the physical and risk-neutral market variance drifts respectively. Applying Ito's lemma to f^j implies the following dynamic of index options under the P-measure

$$\begin{aligned} df_t^j = &\left(\frac{\partial f_t^j}{\partial t} + \frac{\partial f_t^j}{\partial S_t} S_t (\mu - q) + \frac{\partial f_t^j}{\partial V_t} a_t^P + \frac{\partial^2 f_t^j}{\partial S_t \partial V_t} \rho_V \delta S_t V_t^{\eta+0.5} + \frac{1}{2}\left(\frac{\partial^2 f_t^j}{(\partial S_t)^2} V_t S_t^2 + \frac{\partial^2 f_t^j}{(\partial V_t)^2} (\delta V_t^\eta)^2 \right)\right) dt \\ &+ \frac{\partial f_t^j}{\partial S_t} S_t \sqrt{V_t} \left(\sqrt{1-\rho_V^2} dZ_t^S + \rho_V dZ_t^V \right) + \frac{\partial f_t^j}{\partial V_t} \delta V_t^\eta dZ_t^V. \end{aligned} \tag{7.2}$$

We also know that given the dynamics (4.1) to (4.4), the price of any derivative f^j must satisfy the PDE

$$r f_t^j = \frac{\partial f_t^j}{\partial t} + \frac{\partial f_t^j}{\partial S_t} S_t (r - q) + \frac{\partial f_t^j}{\partial V_t} a_t^Q + \frac{\partial^2 f_t^j}{\partial S_t \partial V_t} \rho_V \delta S_t V_t^{\eta+0.5} + \frac{1}{2}\left(\frac{\partial^2 f_t^j}{(\partial S_t)^2} V_t S_t^2 + \frac{\partial^2 f_t^j}{(\partial V_t)^2} (\delta V_t^\eta)^2 \right). \tag{7.3}$$

Combining (7.2) with (7.3), we obtain

$$\begin{aligned} df_t^j &= \left(r f_t^j + \frac{\partial f_t^j}{\partial S_t} S_t (\mu - r - q) + \frac{\partial f_t^j}{\partial V_t} \left(a_t^P - a_t^Q \right) \right) dt + \frac{\partial f_t^j}{\partial S_t} S_t \sqrt{V_t} \left(\sqrt{1-\rho_V^2} dZ_t^S + \rho_V dZ_t^V \right) \\ &\quad + \frac{\partial f_t^j}{\partial V_t} \delta V_t^\eta dZ_t^V. \end{aligned} \tag{7.4}$$

Since the delta replication of f^j satisfies $\Delta Rep_t^j = \theta_t^B \cdot B_t + \theta_t^S \cdot S_t$ where θ^B and θ^S denote the units of bond and market index to hold. For an investment horizon dt, we have $\theta_t^S \cdot S_t = \frac{\partial f_t^j}{\partial S_t} S_t$ and $\theta_t^B \cdot B_t = f_t^j - \frac{\partial f_t^j}{\partial S_t} S_t$. Consequently, the replication portfolio when dividends are reinvested evolves as

$$\begin{aligned} d\Delta Rep_t^j &= \theta_t^B \cdot dB_t + \theta_t^S \cdot (dS_t + qS_t dt) \\ &= \left(f_t^j - \frac{\partial f_t^j}{\partial S_t} S_t \right) \cdot \frac{dB_t}{B_t} + \left(\frac{\partial f_t^j}{\partial S_t} S_t \right) \cdot \left(\frac{dS_t}{S_t} + qdt \right), \end{aligned} \tag{7.5}$$

with $\Delta Rep_t^j = f_t^j$. Combining (7.4) and (7.5), we get

$$\begin{aligned} df_t^j &= d\Delta Rep_t^j + \vartheta_t^j \cdot \left(VRP_t dt + \delta V_t^\eta dZ_t^V \right) \\ &= d\Delta Rep_t^j + \vartheta_t^j \cdot dF_t^V, \end{aligned} \tag{7.6}$$

where $VRP_t \equiv a_t^P - a_t^Q = \frac{1}{dt} \left(E_t^P[V_{t+dt}] - E_t^Q[V_{t+dt}] \right) = \delta V_t^\eta \cdot \phi_t^V$, $\vartheta_t^j \equiv \frac{\partial f_t^j}{\partial V_t}$, and $dF_t^V = VRP_t dt + \delta V_t^\eta dZ_t^V$ is the market variance risk factor. We can now express ϑ_t^j in terms of sensitivity to volatility

$$\vartheta_t^j = \frac{\partial f_t^j}{\partial V_t} = \frac{\partial f_t^j}{\partial \sqrt{V_t}} \cdot \frac{\partial \sqrt{V_t}}{\partial V_t} = Vega_t^j \cdot \frac{1}{2\sqrt{V_t}}, \tag{7.7}$$

which completes the proof.

C. Proof of Proposition 2

We adopt the following strategy. First, we solve the market maker's (unconstrained) portfolio allocation when the market clearing condition is not imposed. Based on the investment strategy obtained, we then require it to satisfy the market clearing condition and infer the structure of the variance risk premium.

The static maximization

$$\max_{\tilde{\pi}_t} \; E^P[\ln(W_T)] \quad \text{subject to} \quad \begin{aligned} &\cdot\, W_0 \geqslant E^P[\xi_T W_T] \\ &\cdot\, W_t \geqslant 0, \end{aligned} \tag{7.8}$$

is the dual problem of the unconstrained utility maximization (4.9) (see, among others, Karatzas, Lehoczky, and Shreve, 1987, and Cox, and Huang, 1989). To solve this, we form the lagrangian

$$\begin{aligned} L(\gamma) &= E^P[\ln(W_T)] + \gamma(W_0 - E^P[\xi_T W_T]) \\ &= E^P[\ln(W_T) - \gamma\xi_T W_T] + \gamma W_0, \end{aligned} \tag{7.9}$$

where γ is the lagrangian coefficient of the static budget constraint $W_0 = E^P[\xi_T W_T]$. A point wise maximization of (7.9) implies the following FOC condition

$$\frac{1}{W_t} = \gamma\xi_t. \tag{7.10}$$

Since the previous equation is valid for any t and $\xi_0 = 1$, the lagrangian coefficient satisfies $\gamma = 1/W_0$. Thus, optimally we have $W_t = \frac{W_0}{\xi_t}$. Given the definition of $\pi_t^{f,j}$, we also have $\pi_t^{f,j} W_t = \Phi_t^{MM,j} f_t^j \Longleftrightarrow \pi_t^{f,j} = \Phi_t^{MM,j} f_t^j / W_t$. Using this with (7.5) and (7.6) in Appendix B, we can write the market maker's aggregate position in index options as

$$\begin{aligned} & \sum_j \pi_t^{f,j} \cdot \frac{df_t^j}{f_t^j} \\ = & \sum_j \frac{\Phi_t^{MM,j}}{W_t} \cdot df_t^j \\ = & \frac{dB_t}{B_t} \cdot \sum_j \frac{\Phi_t^{MM,j}}{W_t} \cdot \left(f_t^j - \frac{\partial f_t^j}{\partial S_t} S_t \right) + \left(\frac{dS_t}{S_t} + q dt \right) \cdot \sum_j \frac{\Phi_t^{MM,j}}{W_t} \cdot \left(\frac{\partial f_t^j}{\partial S_t} S_t \right) + \frac{dF_t^V}{2\sqrt{V_t}} \cdot \frac{InvRisk_t}{W_t}, \end{aligned} \tag{7.11}$$

where we use the definition (2.5) to uncover the market maker's inventory risk. We can now use previous result to express the wealth process (4.8) as

$$\begin{aligned} \frac{dW_t}{W_t} &= \pi_t^B \cdot \frac{dB_t}{B_t} + \pi_t^S \cdot \left(\frac{dS_t}{S_t} + q dt \right) + \sum_j \pi_t^{f,j} \cdot \frac{df_t^j}{f_t^j} \\ &= \bar{\pi}_t^B \cdot \frac{dB_t}{B_t} + \bar{\pi}_t^S \cdot \left(\frac{dS_t}{S_t} + q dt \right) + \frac{InvRisk_t}{W_t} \cdot \frac{dF_t^V}{2\sqrt{V_t}}, \end{aligned} \tag{7.12}$$

where $\bar{\pi}_t^B = \pi_t^B + \sum_j \frac{\Phi_t^{MM,j}}{W_t} \left(f_t^j - \frac{\partial f_t^j}{\partial S_t} S_t \right)$ and $\bar{\pi}_t^S = \pi_t^S + \sum_j \frac{\Phi_t^{MM,j}}{W_t} \left(\frac{\partial f_t^j}{\partial S_t} S_t \right)$ represent market maker's total investment in the bond and in the index respectively.

In this economy, the discounted wealth process satisfies

$$\frac{d(\xi_t W_t)}{\xi_t W_t} = \frac{dW_t}{W_t} + \frac{d\xi_t}{\xi_t} + \frac{d\langle \xi, W \rangle_t}{\xi_t W_t}, \tag{7.13}$$

where $d\langle .,.\rangle_t$ is the covariance operator. Applying the previous equation on the SDF dynamic (4.4) and the wealth process (7.12), we obtain

$$\frac{d(\xi_t W_t)}{\xi_t W_t} = \left(\bar{\pi}_t^S \sqrt{V_t}\sqrt{1-\rho_V^2} - \phi_t^S\right) dZ_t^S + \left(\bar{\pi}_t^S \sqrt{V_t}\rho_V + 0.5\delta V_t^{\eta-0.5}\frac{InvRisk_t}{W_t} - \phi_t^V\right) dZ_t^V, \quad (7.14)$$

for which we have imposed the martingale conditions $\xi_t S_t + \int_0^t q\xi_u S_u du = E_t^P[\xi_T S_T + \int_0^T q\xi_u S_u du]$ and $\xi_t f_t^j = E_t^P[\xi_T f_T^j]$. We can now integrate (7.14) to express $\xi_T W_T$ in its integral form

$$\xi_T W_T = W_0 + \int_0^T \xi_t W_t \left(\left(\bar{\pi}_t^S \sqrt{V_t}\sqrt{1-\rho_V^2} - \phi_t^S\right) dZ_t^S + \left(\bar{\pi}_t^S \sqrt{V_t}\rho_V + 0.5\delta V_t^{\eta-0.5}\frac{InvRisk_t}{W_t} - \phi_t^V\right) dZ_t^V\right). \quad (7.15)$$

By application of the Clark-Ocone formula to $\xi_t W_t$, we also have the following expression for $\xi_T W_T$ in terms of its Malliavin derivatives

$$\begin{aligned} \xi_T W_T &= E_0^P[\xi_T W_T] + \int_0^T E_t^P[D_t^S(\xi_T W_T)]dZ_t^S + \int_0^T E_t^P[D_t^V(\xi_T W_T)]dZ_t^V \\ &= W_0 + \int_0^T E_t^P[D_t^S(\xi_T W_T)]dZ_t^S + \int_0^T E_t^P[D_t^V(\xi_T W_T)]dZ_t^V, \end{aligned} \quad (7.16)$$

where $D_t^i(X)$ is the time t Malliavin derivative of X with respect to Z^i for $i \in \{S, V\}$.[14] This representation of $\xi_T W_T$ can be combined with (7.15) to obtain explicit formulas for $\bar{\pi}^S$ and $InvRisk_t/W_t$. Because both (7.16) and (7.15) uniquely defined $\xi_T W_T$ the integrands in both equations must be equal. This implies

$$\bar{\pi}_t^S \sqrt{V_t}\sqrt{1-\rho_V^2} - \phi_t^S = E_t^P[D_t^S(\xi_T W_T)] \quad (7.17)$$

$$\bar{\pi}_t^S \sqrt{V_t}\rho_V + 0.5\delta V_t^{\eta-0.5}\left(\frac{InvRisk_t}{W_t}\right) - \phi_t^V = E_t^P[D_t^V(\xi_T W_T)]. \quad (7.18)$$

Together, the two previous equations define the market maker's optimal investment strategy. We can now impose the market clearing condition to (7.17) and (7.18). The market clearing condition imposes $\Phi_t^{MM,j} = -\Phi_t^{EU,j}$ for all j and thus $InvRisk_t = -\sum_j \Phi_t^{EU,j} Vega_t^j$ in aggregate. Solving for $\bar{\pi}^S$ in (7.17) and using the result in (7.18), we get

$$\left(E_t^P[D_t^S(\xi_T W_T)] + \phi_t^V\right) - m\left(E_t^P[D_t^V(\xi_T W_T)] + \phi_t^S\right) = 0.5\delta V_t^{\eta-0.5}\left(\frac{InvRisk_t}{W_t}\right), \quad (7.19)$$

[14] Malliavin derivatives have also been used to obtain explicit formulas for optimal investment strategies in Detemple, Garcia, and Rinsdisbacher (2003) and Detemple and Rinsdisbacher (2010) among others. We refer to the Appendix D in Detemple, Garcia, and Rindisbacher (2003) for an introduction to Malliavin calculus and a presentation of the Clarck-Ocone formula. We refer to Di Nunno, Økskendal, and Proske (2009) for an extensive treatment of Malliavin Calculus applied to Finance.

where $m \equiv \rho_V/\sqrt{1-\rho_V^2}$.[15] By the properties of Malliavin derivatives, we have for $i \in \{S, V\}$

$$D_t^i(\xi_T \cdot W_T) = D_t^i(\xi_T \cdot \frac{W_0}{\xi_T}) = D_t^i(W_0) = 0, \tag{7.20}$$

where we used the optimality condition $W_T = W_0/\xi_T$, and the fact that the Malliavin derivative of an adapted process is 0 (i.e. $D_t(X_s) = 0$ when $s < t$). Therefore, $E_t^P[D_t^S(\xi_T W_T)] = E_t^P[D_t^V(\xi_T W_T)] = 0$. Using this into (7.19), we see that

$$\phi_t^V - m\phi_t^S = 0.5\delta V_t^{\eta-0.5}\left(\frac{InvRisk_t}{W_t}\right). \tag{7.21}$$

relates market maker's optimal inventory risk to the two prices of risks. The market index no-arbitrage imposes

$$\xi_t S_t + \int_0^t q\xi_u S_u du = E_t^P[\xi_T S_T + \int_0^T q\xi_u S_u du] \Leftrightarrow \frac{\mu - r}{\sqrt{V_t}} = \sqrt{1-\rho_V^2}\cdot\phi_t^S + \rho_V \cdot \phi_t^V. \tag{7.22}$$

We can use previous equation in order to express the market price of risks ϕ^S in terms of the market premium and ϕ^V

$$\phi_t^S = \frac{\mu - r}{\sqrt{V_t(1-\rho_V^2)}} - m\cdot\phi_t^V. \tag{7.23}$$

Combining (7.21) and (7.23), we can express the market price of variance risk as

$$\phi_t^V = \rho_V\left(\frac{\mu - r}{\sqrt{V_t}}\right) + 0.5(1-\rho_V^2)\delta V_t^{\eta-0.5}\left(\frac{InvRisk_t}{W_t}\right). \tag{7.24}$$

Given that the variance risk premium satisfies $VRP_t = (\delta V_t^\eta)\cdot\phi_t^V$, we finally get

$$VRP_t = \rho_V \delta V_t^\eta\,(Sharpe_t) + 0.5(1-\rho_V^2)\delta^2 V_t^{2\eta-0.5}\left(\frac{InvRisk_t}{W_t}\right), \tag{7.25}$$

where $Sharpe_t = \left(\frac{\mu-r}{\sqrt{V_t}}\right)$, and $InvRisk_t = -\sum_j \Phi_t^{EU,j} Vega_t^j$.

D. Risk-Neutral Dynamics

When the SDF follows (4.4), the Girsanov theorem implies that $dZ_t^S = d\tilde{Z}_t^S - \phi_t^S dt$ and $dZ_t^V = d\tilde{Z}_t^V - \phi_t^V dt$. Using this result in (4.1), (4.2), (4.7), and (5.4) yields the risk-neutral processes.

[15] Note that m is well-defined whenever $abs(\rho_V) \neq 1$.

$$dS_t = (r-q)\,S_t dt + \sqrt{V_t}S_t\left(\sqrt{1-\rho_V^2}d\tilde{Z}_t^S + \rho_V d\tilde{Z}_t^V\right)$$
$$dV_t = \kappa(\theta - V_t)dt - VRP_t dt + \delta V_t^{\eta} d\tilde{Z}_t^V$$
$$dInvRisk_t = \lambda(\alpha - InvRisk_t)dt + \psi V_t dt - \sigma InvRisk_t\left(\sqrt{1-\rho_{Inv}^2}\phi_t^S + \rho_{Inv}\phi_t^V\right)dt$$
$$+\sigma InvRisk_t\left(\sqrt{1-\rho_{Inv}^2}d\tilde{Z}_t^S + \rho_{Inv}d\tilde{Z}_t^V\right)$$
$$dW_t = rW_t dt + \phi_t^S W_t d\tilde{Z}_t^S + \phi_t^V W_t d\tilde{Z}_t^V,$$

where $\tilde{Z}_t^S$ and $\tilde{Z}_t^V$ are independent Brownian motions under the risk-neutral measure, $\phi_t^S = Sharpe_t/\sqrt{1-\rho_V^2}$ $\rho_V\phi_t^V/\sqrt{1-\rho_V^2}$ is obtained by imposing the no-arbitrage condition (7.22), $\phi_t^V = VRP_t/\left(\delta V_t^{\eta}\right)$, and VRP_t satisfies Proposition 2.

E. Particle Filter Estimation

The following algorithm describes the way we evaluate the likelihood $\tilde{P}_t^j\left(V_t^j, \Theta^V\right)$ of observing S_{t+1} given the smooth resampled particles V_t^j, and the structural parameters Θ^V. For estimation purposes, we set the number of particles denoted N to $10,000$.

Using the Euler discretization for $dln(S_t)$ and (4.2), one can simulate the state of the N raw particles $\left\{\tilde{V}_t^j\right\}_{j=1}^N$ forward given $\left\{V_{t-1}^j\right\}_{j=1}^N$ according to

$$Z_t^{V,j} = \left(\frac{\ln\left(\frac{S_t}{S_{t-1}}\right) - \left(\mu - q - \frac{V_t^j}{2}\right)\Delta t}{\sqrt{V_t^j}} - \sqrt{1-\rho_V^2}Z_t^{S,j}\right)/\rho_V \tag{7.26}$$

$$\tilde{V}_t^j = V_{t-1}^j + \kappa(\theta - V_{t-1}^j)\Delta t + \delta\left(V_{t-1}^j\right)^{\eta} Z_t^{V,j}, \tag{7.27}$$

where $Z_t^{S,j}$ is $N(0,\sqrt{\Delta t})$, and μ is fixed to the sample average. Using the set of raw particles, the likelihood of observing S_{t+1} given $\tilde{V}_t^j$ and S_t is

$$\tilde{P}_t^j\left(\tilde{V}_t^j, \Theta^V\right) = \frac{1}{\sqrt{2\pi\tilde{V}_t^j}}\exp\left(-\frac{\left(\ln\left(\frac{S_{t+1}}{S_t}\right) - \left(\mu - q - \frac{\tilde{V}_t^j}{2}\right)\Delta t\right)^2}{2\tilde{V}_t^j}\right). \tag{7.28}$$

Based on the set of normalized weights

$$\breve{P}_t^j\left(\tilde{V}_t^j,\Theta^V\right)=\frac{\tilde{P}_t^j\left(\tilde{V}_t^j,\Theta^V\right)}{\sum_j \tilde{P}_t^j\left(\tilde{V}_t^j,\Theta^V\right)}, \tag{7.29}$$

and the raw $\tilde{V}_t^j$, the method of Pitt (2002) can be applied to resample the smoothed particles $\left\{V_t^j\right\}_{j=1}^N$ and evaluate their corresponding weights $\tilde{P}_t^j\left(V_t^j,\Theta^V\right)$.[16]

F. Risk-Neutral Pricing

Suppose that we want to price an index put option on day t with T days to maturity and strike price K based on $N = 10,000$ simulations. For each simulation n, we initiate the state variables S_t, V_t, and $InvRisk_t$ to their respective values on the day of the pricing. Moreover, we initialize the market maker wealth to w. For a given path n, the forward state of the discretized processes in Appendix D given the information on day t is

$$\ln(S_{t+1}^n)=\ln(S_t^n)+(r-q-V_t^n/2)\,\Delta t+\sqrt{V_t^n}\left(\sqrt{1-\rho_V^2}\,\tilde{Z}_{t+1}^{S,n}+\rho_V\,\tilde{Z}_{t+1}^{V,n}\right) \tag{7.30}$$

$$V_{t+1}^n=V_t^n+\kappa(\theta-V_t^n)\Delta t-VRP_t^n\Delta t+\delta\left(V_t^n\right)^\eta\tilde{Z}_{t+1}^{V,n} \tag{7.31}$$

$$\begin{aligned} InvRisk_{t+1}^n \;=\;& InvRisk_t^n+\lambda(\alpha-InvRisk_t^n)\Delta t+\psi V_t^n\Delta t+\sigma InvRisk_t^n\left(\sqrt{1-\rho_{Inv}^2}\tilde{Z}_{t+1}^{S,n}+\rho_{Inv}\tilde{Z}_{t+1}^{V,n}\right)\\ &-\sigma InvRisk_t^n\left(\sqrt{1-\rho_{Inv}^2}\phi_t^{S,n}+\rho_{Inv}\phi_t^{V,n}\right)\Delta t \end{aligned} \tag{7.32}$$

$$\ln(W_{t+1}^n)=\ln(W_t^n)+\left(r-\left(\left(\phi_t^{S,n}\right)^2+\left(\phi_t^{V,n}\right)^2\right)/2\right)\Delta t+\phi_t^{S,n}\tilde{Z}_{t+1}^{S,n}+\phi_t^{V,n}\tilde{Z}_{t+1}^{V,n}, \tag{7.33}$$

where $\tilde{Z}_{t+1}^{S,n}$ and $\tilde{Z}_{t+1}^{V,n}$ are independent $N(0,\sqrt{\Delta t})$. In the previous system, we set

$$VRP_t^n=\rho_V\delta\left(V_t^n\right)^\eta Sharpe_t^n+0.5(1-\rho_V^2)\delta^2\left(V_t^n\right)^{2\eta-0.5}\left(\frac{InvRisk_t^n}{W_t^n}\right), \tag{7.34}$$

[16] The method proposed in Pitt (2002) involves smoothing the $\breve{P}_t^j$ to a continuous CDF from which the set of smooth particle V_t^j can be resampled.

where $Sharpe_t^n = \left(\frac{\mu - r}{\sqrt{V_t^n}}\right)$. Moreover, the prices of risks are calculated according to

$$\phi_t^{S,n} = Sharpe_t^n / \sqrt{1-\rho_V^2} - \rho_V \phi_t^{V,n} / \sqrt{1-\rho_V^2} \quad \text{and} \quad \phi_t^{V,n} = VRP_t^n / \left(\delta \left(V_t^n\right)^\eta\right). \tag{7.35}$$

Simulating the system forward from day t to T, the price of the index put option on day t is equal to

$$P\left(\Theta^{Inv}, w, \Theta^V, V_t, InvRisk_t\right) = \sum_{n=1}^{N} \frac{\max\left(K - S_T^n, 0\right) \cdot \exp(-r \cdot (T-t)/365)}{N}, \tag{7.36}$$

where Θ^V and Θ^{Inv} are the structural parameters of the market variance and inventory risk processes respectively, and w is the market maker's wealth parameter.

References

[1] Adrian, T., Etula, E., Muir, T. (2014) Financial intermediaries and the cross-section of asset returns, Journal of Finance, 69, 2557–2596.

[2] Adrian, T., Shin, H.S. (2010) The changing nature of financial intermediation and the financial crisis of 2007–2009, Annual Review of Economics, 2, 603–618.

[3] Aït-Sahalia, Y., Karaman, M., Mancini, L. (2013) The term structure of variance swaps and risk premia, Working Paper, Princeton University.

[4] Aït-Sahalia, Y., Kimmel, R. (2007) Maximum likelihood estimation of stochastic volatility models, Journal of Financial Economics, 83, 413–452.

[5] Amihud, Y., Mendelson, H., (1980) Dealership market: Market making with inventory, Journal of Financial Economics, 8, 31–53.

[6] Ang, A., Hodrick, R., Xing, Y., Zhang, X. (2006) The cross-section of volatility and expected returns, Journal of Finance, 61, 259–299.

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

[8] Bakshi, G., Madan, D. (2006) A Theory of volatility spread, Management Science, 52, 1945–1956.

[9] Barras, L., Malkhozov, A. (2014) When variance risk has two prices: Evidence from the equity and option markets, Working Paper, McGill University.

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

[11] Bates, D. (2003) Empirical option pricing: A retrospection, Journal of Econometrics, 116, 387–404.

[12] Black, F., Scholes, M. (1973) Valuation of options and corporate liabilities, Journal of Political Economy, 81, 637–654.

[13] Bollen, N.P., Whaley, R.E. (2004) Does net buying pressure affect the shape of implied volatility functions?, Journal of Finance, 59, 711–53.

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

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

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

[17] Buraschi, A., Trojani, F., Vedolin, A. (2014) When uncertainty blows in the orchard: Comovement and equilibrium volatility risk premia, Journal of Finance, Forthcoming.

[18] Carr, P., Wu, L. (2007) Stochastic skew in currency options, Journal of Financial Economics, 86, 213–247.

[19] Carr, P., Wu, L. (2009) Variance risk premiums, Review of Financial Studies, 22, 1311–1341.

[20] Chen, H., Joslin, S., Ni, S. (2013) Demand for crash insurance and stock returns, Working Paper, MIT.

[21] Christoffersen, P., Jacobs, K., Mimouni, K. (2010) Volatility dynamics for the S&P 500: evidence from realized volatility, daily returns, and option prices, Review of Financial Studies, 23, 3141–3189.

[22] Comerton-Forde, C., Hendershott, T., Jones, C., Moulton, P., Seasholes, M. (2010) Time variation in liquidity: The role of market maker inventories and revenues, Journal of Finance 65, 295–331.

[23] Corsi, F. (2009) A simple approximate long-memory model of realized volatility, Journal of Financial Econometrics, 7, 174–196.

[24] Cox, J.C., Huang, C.F. (1989) Optimal consumption and portfolio plicies when asset prices follow a diffusion process, Journal of Economic Theory, 49, 33–83.

[25] Cremers, M., Halling, M., Weibaum, D. (2014) Aggregate jump and volatility risk in the cross-section of stock returns, Journal of Finance, Forthcoming.

[26] Detemple, J. B., Garcia, R., Rindisbacher, M. (2003) A Monte Carlo method for optimal portfolios, Journal of Finance, 58, 401–446.

[27] Detemple, J. B., Garcia, R., Rindisbacher, M. (2005) Intertemporal asset allocation: a comparison of methods, Journal of Banking and Finance, 29, 2821–2848.

[28] Detemple, J. B., Rindisbacher, M. (2010) Dynamic asset allocation: portfolio decomposition formula and applications, Review of Financial Studies, 23, 25–100.

[29] Dew-Becker, I., Giglio, S., Le, A., Rodriguez, M. (2015), The price of variance risk, Working Paper.

[30] Di Nunno, G., Øksendal, B., Proske, F. (2009) Malliavin calculus for Lévy processes with applications to finance, Springer-Verlag.

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

[32] Egloff, D., Leippold, M., and Wu, L. (2010) The term structure of variance swap rates and optimal variance swap investments, Journal of Financial and Quantitative Analysis, 45, 1279–1310.

[33] Elkamhi, R., Stefanova, D. (2011) Dynamic correlation or tail dependence hedging for portfolio selection, Working Paper, University of Toronto.

[34] Gârleanu, N., Pedersen, L.H., Poteshman, A. (2009) Demand-based option pricing, Review of Financial Studies, 22, 4259–4299.

[35] Gromb, D., Vayanos, D. (2002) Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs, Journal of Financial Economics, 66, 361–407.

[36] Hansch, O., Naik, N., Viswanathan, S., (1998) Do inventories matter in dealership markets? Evidence from the London Stock Exchange, Journal of Finance, 53, 1623–1655.

[37] Heston, S. (1993) Closed-form solution for options with stochastic volatility, with application to bond and currency options, Review of Financial Studies, 6, 327–343.

[38] Ho, T., Stoll, H., (1981) Optimal dealer pricing under transactions and return uncertainty, Journal of Financial Economics, 9, 47–73.

[39] Ho, T., Stoll, H., (1983) The dynamics of dealer markets under competition, Journal of Finance, 38, 1053–1074.

[40] Johannes, M., Polson, N., Stroud, J. (2009) Optimal filtering of jump-diffusions: Extracting latent states from asset prices, Review of Financial Studies, 22, 2759–2799.

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

[42] Karatzas, I., Lehoczky, J.P., Shreve, S. (1987) Optimal portfolio and consumption decisions for a small investor on a finite time-horizon, Journal on Control and Optimization, 25, 1557–1586.

[43] Leippold, M., Su, L. (2011) Collateral smile, Working Paper, Swiss Finance Institute.

[44] Madhavan, A., Smidt, S., (1993) An analysis of daily changes in specialist inventories and quotations, Journal of Finance, 48, 1595–1628.

[45] Madhavan, A., Sofianos, G., (1998) An empirical analysis of NYSE specialist trading, Journal of Financial Economics, 48, 189–210.

[46] Malik, S., Pitt, M. (2011) Particle filters for continuous likelihood evaluation and maximisation, Journal of Econometrics, 165, 190–209.

[47] Mildenstein, E., Schleef, H., (1983) The optimal pricing policy of a monopolistic market maker in the equity market, Journal of Finance, 38, 218–231.

[48] Naik, N., Yadav, P., (2003) Do dealer firms manage inventory on a stock-by-stock or a portfolio basis? Journal of Financial Economics, 69, 325–353.

[49] Pan, J. (2002) The jump risk premia implicit in options: Evidence from an integrated time series study, Journal of Financial Economics, 63, 3–50.

[50] Pitt, M. (2002) Smooth particle filters for likelihood evaluation and maximization, Working Paper, University of Warwick.

[51] Schürhoff, N., Ziegler, A. (2011) Variance risk, financial intermediation, and the cross-section of expected option returns. Working Paper, University of Zurich.

[52] Shleifer, A., Vishny, R. (1997) The limits of arbitrage, Journal of Finance, 36, 35–55.

[53] Sofianos, G. (1995) Specialist gross trading revenues at the New York Stock Exchange, NYSE Working Paper.

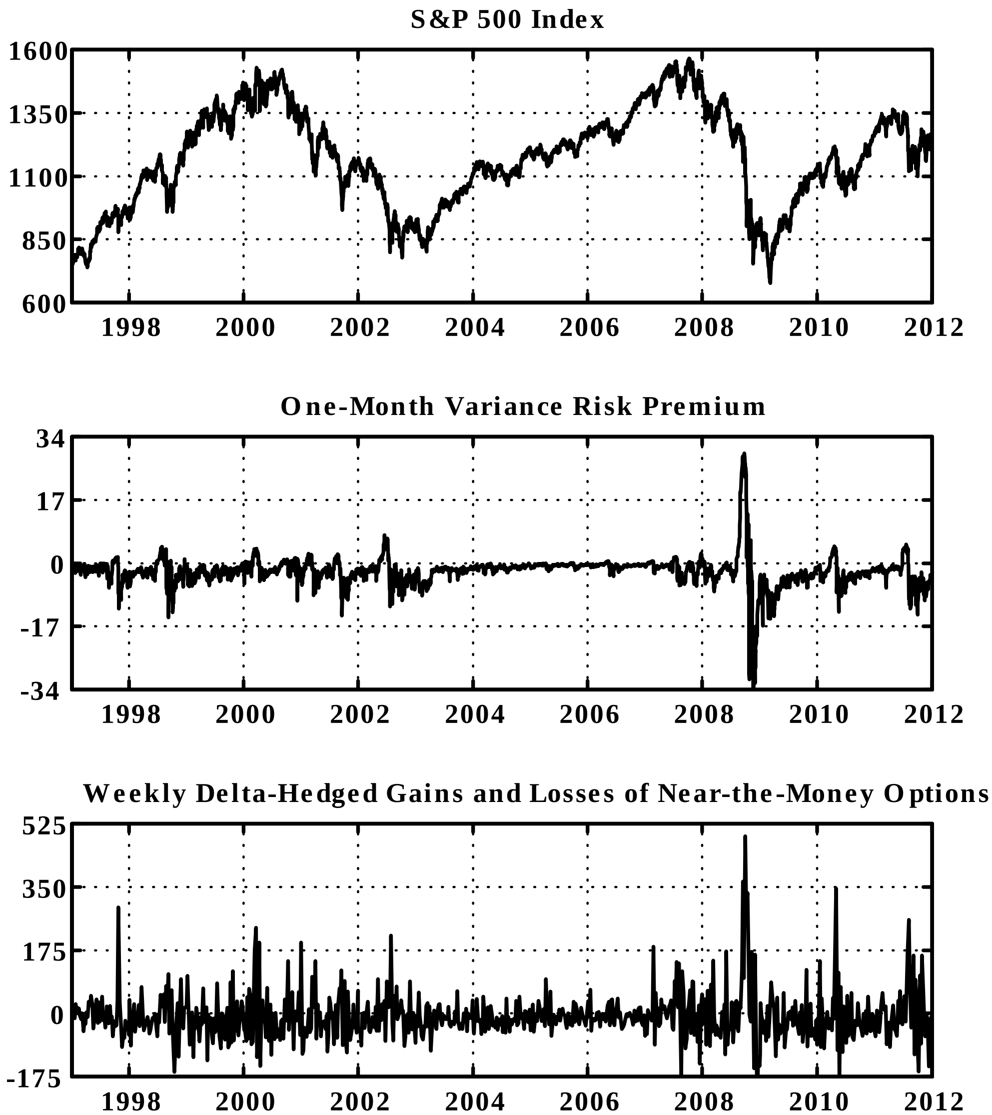
[54] Todorov, V. (2010) Variance risk premia dynamics: The role of jumps, Review of Financial Studies, 23, 345–383.

[55] Trolle, A., Schwartz, E. (2009) Unspanned stochastic volatility and the pricing of commodity derivatives, Review of Financial Studies, 22, 4423–4461.

[56] Vilkov, G. (2008) Variance risk premium demystified, Working Paper, INSEAD.

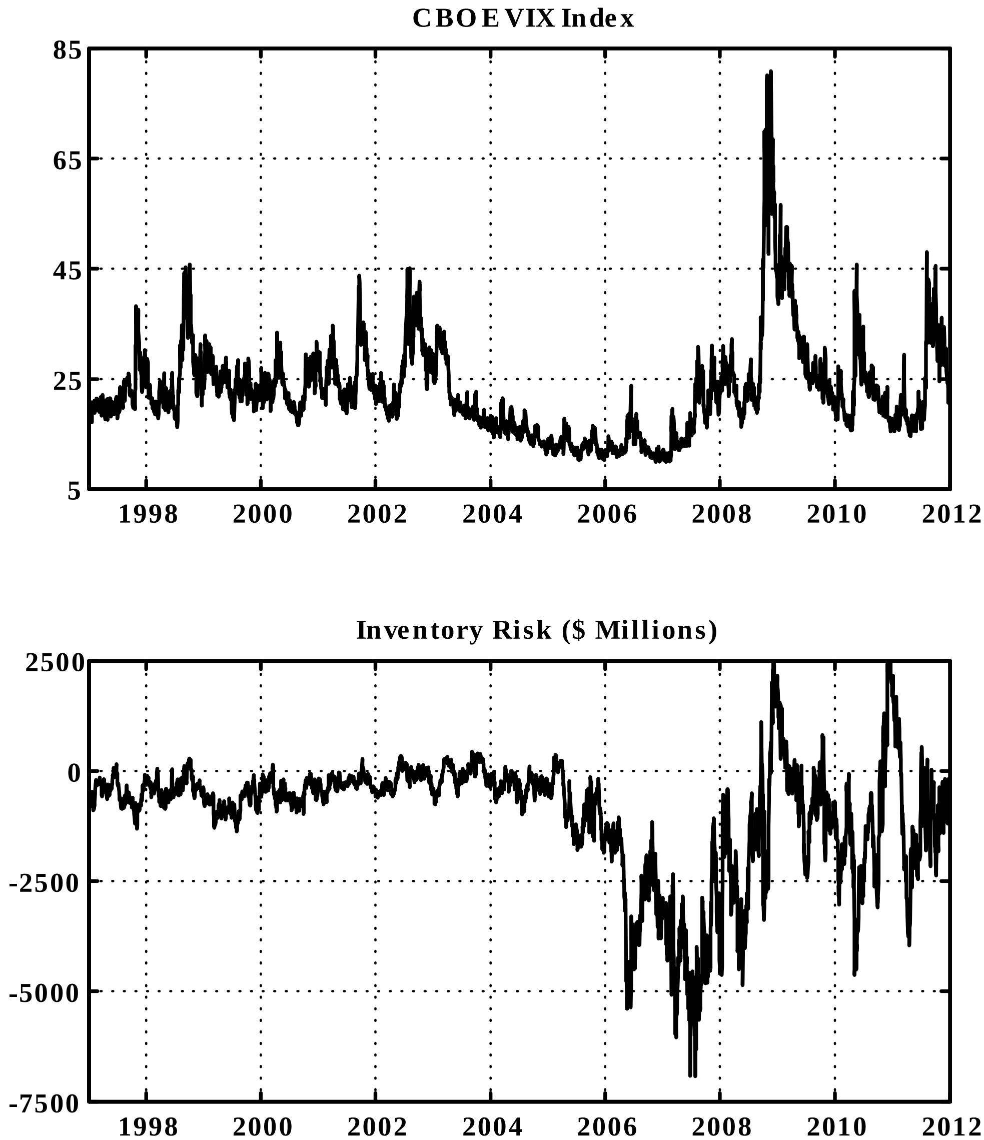
[57] Zhang, L., Mykland, P., Aït-Sahalia, Y. (2005) A Tale of two time scales: Determining integrated volatility with noisy high-frequency data, Journal of the American Statistical Association, 100, 1394–1411.

Figure 1: The S&P 500 Index, theVariance Risk Premium, and Delta-Hedged Gains and Losses



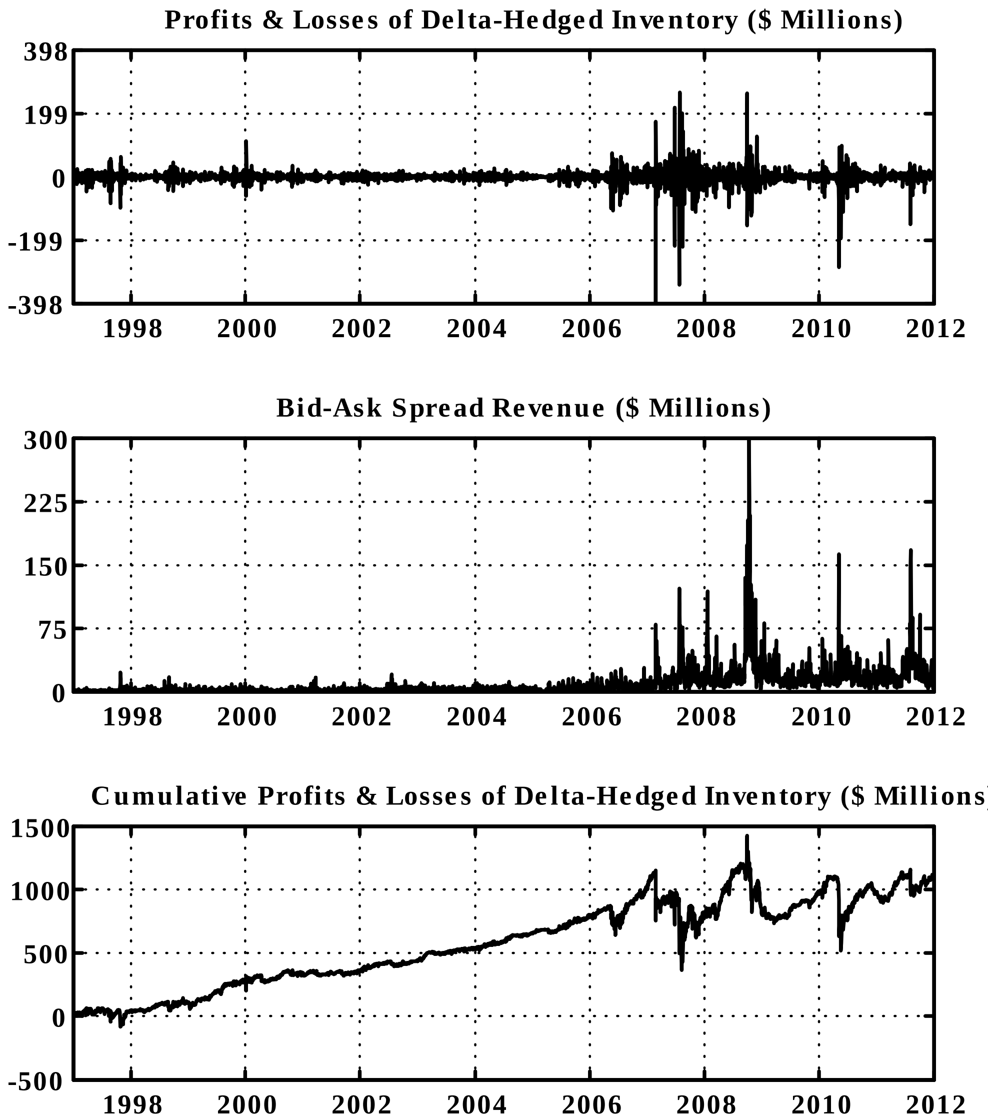
Notes to Figure: The top panel plots the time series of the S&P 500 index. The middle panel plots the one-month variance risk premium expressed in percentages and measured as the difference between the one-month ex-post realized variance and the one-month expected risk-neutral variance. The bottom panel graphs the weekly average of daily delta-hedged gains and losses for all options with moneyness (S/K) between 0.98 and 1.02.

Figure 2: The VIX Index and Market Maker Inventory Risk



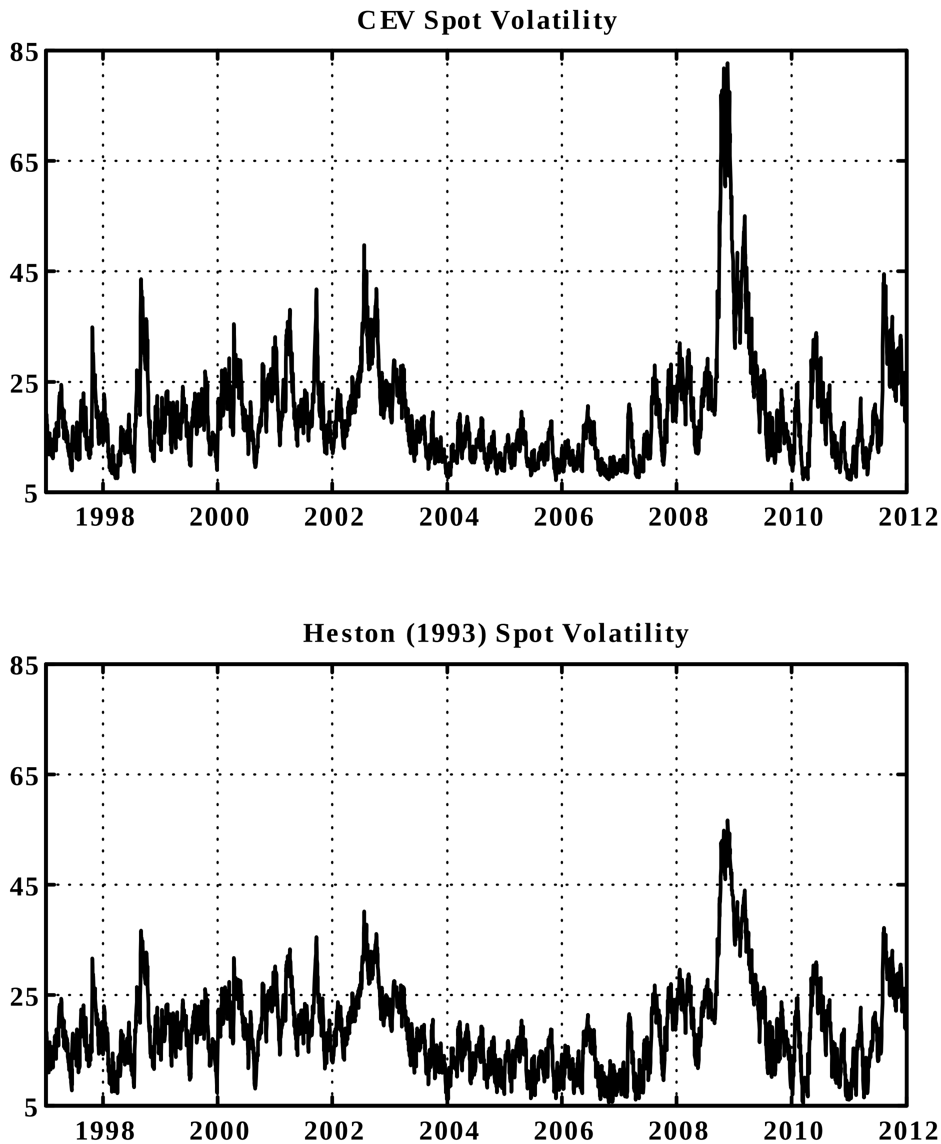
Notes to Figure: The top panel plots the CBOE VIX index, which represents the implied volatility of an at-the-money option with exactly 30 days to maturity expressed in percentages. The bottom panel plots CBOE market makers' inventory risk dynamic, measured as the vega-weighted sum of inventories across all contracts expressed in millions of dollars.

Figure 3: Market Makers' Daily and Cumulative Profits and Losses, and Market Makers' Bid-Ask Spread Revenue



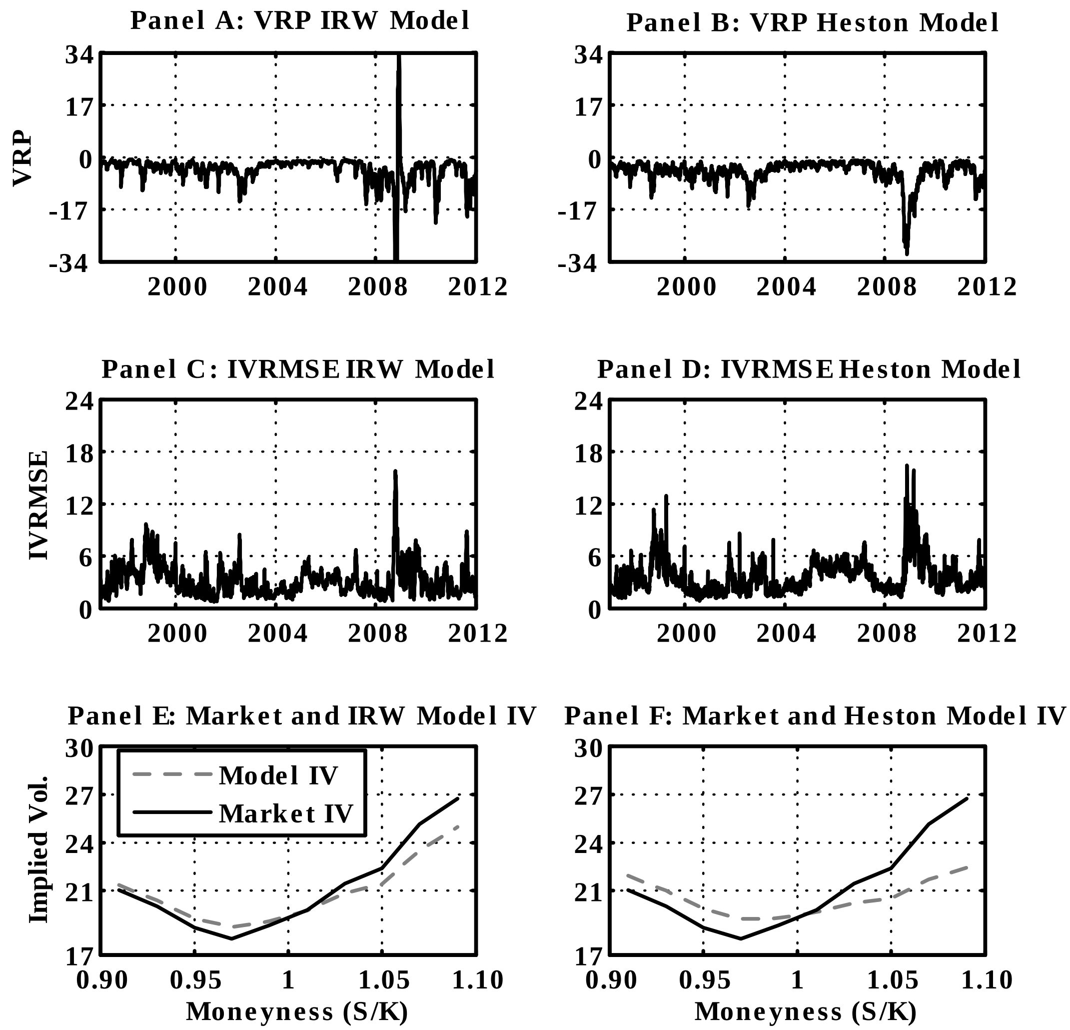
Notes to Figure: The top panel plots the daily profits and losses from market makers' delta-hedged inventory expressed in millions of dollars. The middle panel graphs the cumulative profits and losses from market makers' delta-hedged inventory in millions of dollars. The bottom panel plots the SPX market makers' bid-ask spread revenue expressed in millions of dollars.

Figure 4: Filtered Spot Volatilities Estimated Using Daily S&P 500 Returns



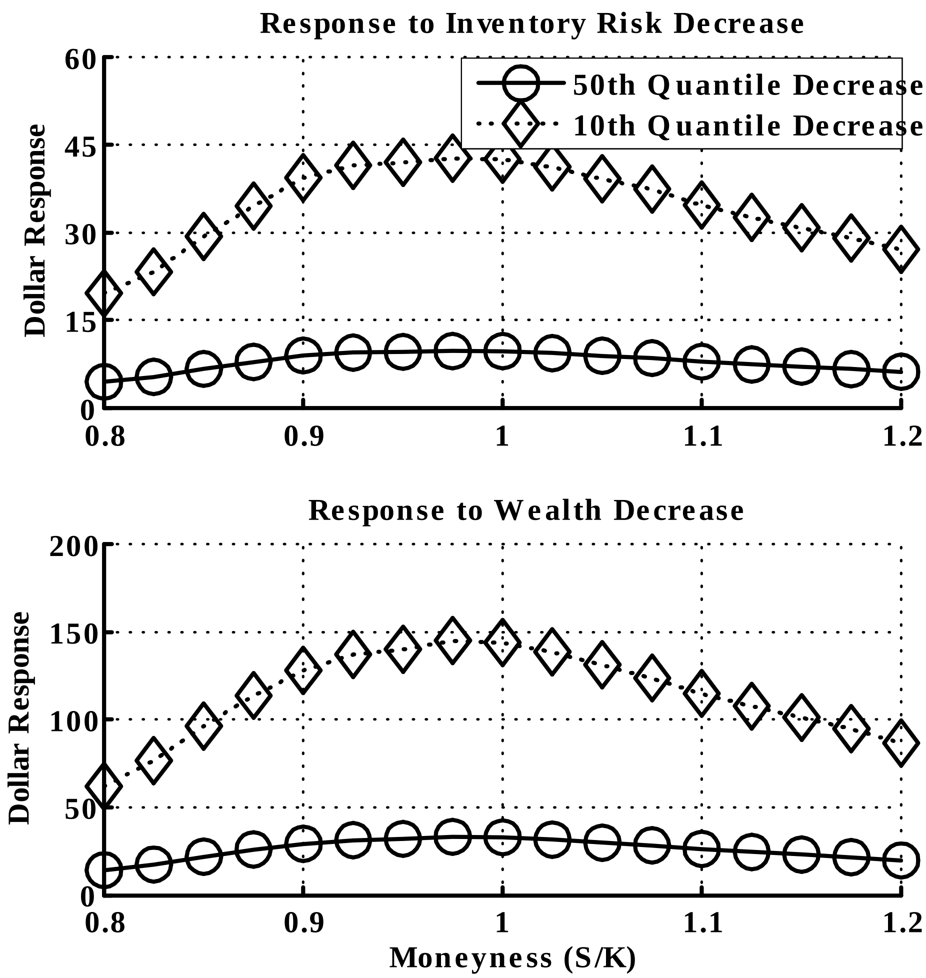
Notes to Figure: The figure plots the daily spot volatilities filtered from S&P 500 daily returns using particle filtering. In the top panel, we plot the spot volatilities estimated using the CEV dynamic. In the bottom panel, we graph the filtered spot volatilities estimated using the Heston (1993) model. For both panels, the daily volatilities are annualized and expressed in percentages.

Figure 5: Model-Implied Variance Risk Premiums, IVRMSEs, and Implied Volatility Smiles



Notes to Figure: Panels A-D plot the daily one-month variance risk premium (VRP) and IVRMSE, expressed in percentages, for the IRW and Heston models. To obtain the models' one-month VRP on each day, we simulate 10,000 paths, calculate the 30 days integrated VRP for each path, and take the average. For the IRW model, the VRP is calculated using estimated parameters and latent variables. For the Heston model, the instantaneous VRP is set to $h \cdot V_t$ where $h = -1.08$. Panels E and F plot the market-implied (solid) and model-implied (dashed) volatility smile.

Figure 6: The Impact of Inventory Risk and Market Maker Wealth on Option Prices



Notes to Figure: We plot the dollar response of SPX puts with 90 days to maturity to a decrease in inventory risk (top panel) and to a decrease in market maker wealth (bottom panel). To calculate the model sensitivies $\frac{\partial P}{\partial InvRisk}$ and $\frac{\partial P}{\partial W}$, we use the estimated parameters $\hat{\Theta}^V$, $\hat{\Theta}^{Inv}$, and $\hat{w}$, and set $r = 4\%$, $q = 0$, $S_t = 1183$, and $V_t = \hat{\theta}$. Inventory risk and market maker wealth are set to $InvRisk_t = -9.03E + 09$ and $W_t = \hat{w}$. Based on these sensitivities, we then calculate the dollar response of each option as $\Delta InvRisk_t \cdot \frac{\partial P}{\partial InvRisk}$ and $\Delta W_t \cdot \frac{\partial P}{\partial W}$, and plot the result across moneyness. The circles plot the dollar response to an average decrease in the state variables. The diamonds plot the dollar response to a decrease in the state variables equal to the 90^{th} percentile of the sample distribution of decreases.

Table 1: Descriptive Statistics for SPX Options

Year	Implied Volatility (%)	Vega	Days to Maturity	Quotes	Volume	$VRP_{t,30}$ (%)
1997	22.18	133	102	214	303	-1.95
1998	25.14	164	101	181	324	-2.35
1999	24.96	217	110	178	290	-2.60
2000	22.95	223	111	138	273	-0.93
2001	23.93	186	116	159	316	-2.64
2002	25.20	149	109	157	399	-2.45
2003	21.64	140	109	153	503	-2.71
2004	16.44	156	109	186	525	-1.17
2005	14.23	148	95	171	799	-0.63
2006	14.57	151	85	217	1,109	-0.68
2007	18.69	203	101	343	1,073	-1.17
2008	29.39	181	101	457	826	-1.29
2009	29.51	136	98	474	709	-5.61
2010	22.30	160	106	516	708	-2.63
2011	23.04	170	95	545	713	-3.08
Average	22.28	168	103	273	591	-2.13

Notes to Table: For each year, we report the average of implied volatility, vega, days to maturity, number of quotes, and volume for SPX puts and calls combined. We also report the average of the one-month variance risk premium measured as the difference between the one-month ex-post realized variance and the one-month expected risk-neutral variance. Option implied volatility and vega are computed using the Black-Scholes model.

Table 2: Implied Volatility, Market Maker Inventory, and Delta-Hedged Gains and Losses by Moneyness and Maturity for SPX Options

Days to Maturity		*Moneyness (S/K)* 0.80 to 0.85	0.85 to 0.95	0.95 to 1.05	1.05 to 1.15	1.15 to 1.20	Sum of Inventory by Days-to-Maturity
10 to 30	IV (%)	43.80	21.27	21.00	29.06	38.48	
	Inventory	-1,419	-2,010	-15,100	-1,856	-3,616	*-24,000*
	ΔHedge ($)	14.27	5.68	-8.16	-10.29	-5.87	
31 to 60	IV (%)	27.81	17.71	20.73	26.47	31.36	
	Inventory	-1,226	-4,309	-12,574	-7,897	-777	*-26,784*
	ΔHedge ($)	14.59	2.22	-9.08	-10.92	-11.26	
61 to 90	IV (%)	21.99	17.60	20.83	25.60	29.10	
	Inventory	-499	-1,376	-6,567	-8,672	-946	*-18,061*
	ΔHedge ($)	11.25	2.67	-5.08	-7.39	-13.54	
91 to 120	IV (%)	20.10	18.85	22.33	26.65	29.93	
	Inventory	-332	-1,288	-6,752	-9,241	-2,324	*-19,937*
	ΔHedge ($)	11.18	2.89	-5.24	-7.05	-19.65	
121 to 365	IV (%)	17.52	18.63	21.32	24.22	26.21	
	Inventory	464	1,628	-1,615	-8,900	-2,159	*-10,581*
	ΔHedge ($)	7.14	-0.53	-3.48	-4.65	-14.57	
Sum of Inventory by Moneyness		*-3,012*	*-7,354*	*-42,608*	*-36,566*	*-9,823*	***Total Inventory -99,363***

Notes to Table: For each SPX option moneyness and maturity category, we compute the average of the implied volatility (IV) and market makers' inventory. We compute averages for each category on each day and then average across days. We also report ΔHedge, which denotes the sample average of the daily delta-hedged gains and losses across all options in each moneyness and maturity category. The top right column reports the sum of inventory for each maturity category and the bottom row reports the sum of inventory by moneyness category.

Table 3: Explaining Time Variation in the One-Month Log Variance Risk Premium

	Dependent Variable: $\Delta \text{LogVRP}_{t,30} \times 100$					
	1997-2011		*1997-2004*		*2005-2011*	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Intercept	-0.26	-0.06	-0.43	-0.23	-0.10	-0.01
	(*0.17*)	(*0.80*)	(*0.11*)	(*0.55*)	(*0.70*)	(*0.98*)
InvRisk(*t-1*)		0.96 ***		0.85 ***		1.20 ***
		(*0.00*)		(*0.00*)		(*0.00*)
ΔW(*t*) × InvRisk(*t-1*)		-3.39 ***		*-5.82* ***		*-5.11* ***
		(*0.00*)		(*0.00*)		(*0.00*)
S&P500LogRet(*t*)	7.53 ***	7.46 ***	7.32 ***	6.84 ***	7.97 ***	7.91 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
JumpFactor(*t*)	*-4.89* ***	*-4.06* ***	-4.12 ***	-2.28 ***	-5.74 ***	-4.06 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
NetBuyingPressure(*t*)	-0.22	-0.14	-0.03	-0.16	-0.62 **	-0.31
	(*0.26*)	(*0.45*)	(*0.92*)	(*0.53*)	(*0.02*)	(*0.17*)
Disagreement(*t*)	*-0.46* ***	*-0.61* ***	0.01	-0.26	-0.44	-0.74 **
	(*0.07*)	(*0.01*)	(*0.98*)	(*0.50*)	(*0.19*)	(*0.03*)
$\Delta \text{LogVRP}_{t-1,30}$	-3.25 ***	-2.97 ***	-5.02 ***	-3.94 ***	-0.99 ***	-0.39
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.01*)	(*0.22*)
Adj. R^2 (%)	39	43	34	43	49	57
N	3774	3774	2011	2011	1763	1763

Notes to Table: We present the full sample and subsample results from regressing daily changes in the one-month log-variance risk premium, denoted $\Delta \text{LogVRP}_{t,30}$, on lagged inventory risk, the interaction of changes in market maker wealth with lagged inventory risk, and the control variables: S&P 500 log-return, S&P500 jump, net buying pressure for index options, and investors' disagreement. On each day, we estimate inventory risk by summing vega-weighted inventories across options. Changes in market maker wealth are measured as the sum of delta-hedged inventory profits and losses (2.6) and bid-ask spread revenue (2.7). All regressors are standardized to have unit variance. The *p*-values are in parentheses and are computed using Newey-West with 8 lags.

Table 4: Explaining Time Variation in the Log Variance Risk Premium. Various Maturities

	Dependent Variable: $\Delta LogVRP_{t,T} \times 100$				
	T=30	*T=60*	*T=90*	*T=180*	*T=270*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Intercept	-0.06	-0.09	-0.16	-0.03	0.01
	(*0.80*)	(*0.57*)	(*0.21*)	(*0.72*)	(*0.94*)
InvRisk(*t-1*)	0.96 ***	0.64 ***	0.39 ***	0.41 ***	0.25 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
ΔW(*t*) × InvRisk(*t-1*)	-3.39 ***	-2.51 ***	-2.03 ***	-1.46 ***	-1.28 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
S&P500LogRet(*t*)	7.46 ***	6.39 ***	5.58 ***	4.08 ***	3.52 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
JumpFactor(*t*)	*-4.06* ***	-2.37 ***	-1.85 ***	-1.10 ***	-0.79 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
NetBuyingPressure(*t*)	-0.14	-0.17	-0.15	-0.07	0.06
	(*0.45*)	(*0.21*)	(*0.17*)	(*0.34*)	(*0.62*)
Disagreement(*t*)	*-0.61* ***	-0.45 **	-0.53 ***	-0.33 ***	-0.41 ***
	(*0.01*)	(*0.03*)	(*0.00*)	(*0.00*)	(*0.00*)
$\Delta LogVRP_{t-1,T}$	-2.97 ***	-1.48 ***	-1.20 ***	-0.53 ***	-0.16
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.23*)
Adj. R^2 (%)	43	47	49	53	20
N	3774	3774	3774	3774	3774

Notes to Table: The table presents results from regressing daily changes in the log variance risk premium, denoted $\Delta LogVRP_{t,T}$, on the explanatory variables. We consider the variance risk premium for five different horizons T. All regressors are standardized to have unit variance. The p-values are in parentheses and are computed using Newey-West with 8 lags. The sample period is 1997-2011.

Table 5: Explaining Time Variation in the One-Month Log Variance Risk Premium. Various Definitions of the Risk Premium

	Dependent Variable: $\Delta LogVRP_{t,T} \times 100$		
	$RV_{t,T}$ proxied by Future Realized Variances (1)	*$RV_{t,T}$ proxied by HAR-RV Model Forecast (2)*	*Parameter Difference (1) - (2)*
Intercept	-0.07	-0.10	0.03
	(*0.65*)	(*0.53*)	
InvRisk(*t-1*)	0.53 ***	0.55 ***	-0.02
	(*0.00*)	(*0.00*)	
ΔW(*t*) × InvRisk(*t-1*)	-2.13 ***	-2.28 ***	0.15
	(*0.00*)	(*0.00*)	
S&P500LogRet(*t*)	5.41 ***	5.45 ***	-0.04
	(*0.00*)	(*0.00*)	
JumpFactor(*t*)	*-2.03* ***	*-1.93* ***	-0.10
	(*0.00*)	(*0.00*)	
NetBuyingPressure(*t*)	-0.10	-0.32	0.22
	(*0.36*)	(*0.16*)	
Disagreement(*t*)	*-0.47* ***	*0.18*	-0.65
	(*0.01*)	(*0.35*)	
$\Delta LogVRP_{t-1,T}$	-1.27 **	-2.16 *	0.89
	(*0.05*)	(*0.01*)	
Adj. R^2 (%)	42	44	

Notes to Table: We report parameter estimates and *p*-values from regressing daily changes in the log-variance risk premium on the explanatory variables.The first column reports the average parameter estimates and *p*-values from Table 4, across horizons T. In the second column, we present the average parameter estimates and *p*-values obtained when the log-variance risk premia are constructed based on the HAR-RV model prediction for $RV_{t,T}$. Detailed results for each horizon are provided in Table A.2 in the Online Appendix. All regressors are standardized to have unit variance and the *p*-values are computed using Newey-West with 8 lags. The sample period is 1997-2011.

Table 6: Time Variation in the Term Structure of Variance Risk Premia

	Dependent Variable: $\Delta VRP_{t,T} \times 100$				
	T=30	*T=60*	*T=90*	*T=180*	*T=270*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Intercept	0.00	0.00	0.00	0.00	0.02
	(*0.86*)	(*0.81*)	(*0.74*)	(*1.00*)	(*0.38*)
InvRisk(*t-1*)	0.06 ***	0.04 ***	0.02 *	0.02 ***	0.02 ***
	(*0.00*)	(*0.01*)	(*0.07*)	(*0.01*)	(*0.01*)
ΔW(*t*) × InvRisk(*t-1*)	-0.17 ***	-0.11 ***	-0.09 ***	-0.06 ***	-0.06 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
S&P500LogRet(*t*)	0.83 ***	0.66 ***	0.54 ***	0.36 ***	0.30 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
JumpFactor(*t*)	*-0.19* ***	-0.09 ***	-0.06 ***	-0.03 **	-0.01
	(*0.00*)	(*0.00*)	(*0.01*)	(*0.02*)	(*0.32*)
NetBuyingPressure(*t*)	-0.01	-0.02	-0.02 *	-0.01	0.02
	(*0.44*)	(*0.13*)	(*0.08*)	(*0.19*)	(*0.48*)
Disagreement(*t*)	*-0.06* **	-0.05 *	-0.04 **	-0.03 ***	-0.04 **
	(*0.05*)	(*0.06*)	(*0.03*)	(*0.01*)	(*0.05*)
$\Delta VRP_{t-1,T}$	-0.31 ***	-0.15 ***	-0.10 ***	-0.05 ***	0.00
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.37*)
Adj. R^2 (%)	44	50	53	57	4
N	3774	3774	3774	3774	3774

Notes to Table: We present the results from regressing daily changes in the variance risk premium, denoted $\Delta VRP_{t,T}$, on the explanatory variables. As in Table 4, we consider five horizons T to investigate the term structure of variance risk premia. All regressors are standardized to have unit variance. The p-values are in parentheses and are computed using Newey-West with 8 lags.

Table 7: The Return Distribution of Delta-Hedged Near-the-Money Options

Frequency	Min (%)	Max (%)	Average Daily per Period Return × 30 (%)	Annualized Volatility (%)	Skewness	Excess Kurtosis
			Full Sample			
Daily	-12.19	35.46	-8.11	71.11	2.33	10.72
Weekly	-3.92	9.02	-8.37	32.65	1.27	3.42
Monthly	-2.08	4.84	-8.21	17.27	1.44	5.21
			Sample: August 9, 2007 - April 2, 2009			
Daily	-10.43	35.46	6.26	95.68	2.55	11.47
Weekly	-3.92	7.94	3.31	42.07	1.34	2.93
Monthly	-1.38	4.84	3.48	28.38	1.95	3.58

Notes to Table: We report the minimum, maximum, average, volatility, skewness, and excess kurtosis for the return distribution of delta-hedged near-the-money options. Every day, we delta-hedge one long position in each option with moneyness between 0.98 and 1.02, assuming daily rebalancing. We then average the daily returns over all the options on that day to obtain a single return for that day. To obtain weekly and monthly return measures, we average the daily returns over each week and month respectively. We report the statistics for the return distribution calculated using the full sample, as well as for a sample period corresponding to the financial crisis. For the financial crisis, the sample starts on August 9th, 2007, when BNP Paribas froze three of their funds due to valuation issues, and ends on April 2nd, 2009, when the G20 agreed on a global stimulus package worth five trillion dollars.

Table 8: Return Statistics and Parameter Estimates for the CEV and Heston Volatility Models

Panel A: Annualized Statistics for Daily S&P 500 Returns

Mean	Variance
5.860%	4.583%

Panel B: The CEV Model

Parameter Estimates, *p* -values, and Standard Errors					MLIS Objective Value	Filtered Spot Variances	
κ	θ	δ	ρ_V	η		Mean	Variance
2.91	4.56%	98.35%	-0.60	0.90			
(0.002)	(0.000)	(0.012)	(0.000)	(0.000)	11,746.29	4.433%	142.728%
0.922	0.009	0.394	0.129	0.061			

Panel C: The Heston Model

Parameter Estimates, *p* -values, and Standard Errors					MLIS Objective Value	Filtered Spot Variances	
κ	θ	δ	ρ_V	η		Mean	Variance
5.32	4.08%	18.82%	-0.47	0.50			
(0.000)	(0.000)	(0.014)	(0.002)		11,722.79	4.084%	52.913%
0.701	0.003	0.077	0.152				

Notes to Table: Panel A presents the descriptive statistics for the sample of daily S&P 500 returns. Panel B presents results for the physical variance dynamic in equation (4.2). Panel C presents results for the Heston (1993) model with the physical variance dynamic in equation (4.2) with $\eta = 1/2$. For Panels B and C, the structural parameters and daily spot variances are obtained by maximum likelihood importance sampling (MLIS) on S&P 500 daily returns. We set the difference μ-q equal to the sample average of 5.86%. All parameters and statistics are in annual units. The *p* -values in parentheses are based on standard errors computed using the outer product of the gradient evaluated at the optimal parameter values.

Table 9: Parameter Estimates and Model Fit for the IRW and Heston Models

Panel A: Parameter Estimates and Model Fit for the IRW Model

Inventory Risk and Wealth Parameters						Sum of Implied Volatility Squared Errors
λ	α	ψ	σ	ρ_{Inv}	w ($ Millions)	
10.73	-1.00E+10	2.28E+11	16.55%	-5.14E-04	440.92	6.68

Panel B: Parameter Estimates and Model Fit for the Heston (1993) Model

Price of Variance Risk	Sum of Implied Volatility Squared Errors
h	
-1.08	7.99

Notes to Table: Panel A presents the estimates for the inventory risk dynamic (4.7) and the market maker initial wealth parameter when the wealth dynamic evolves according to (4.8). To estimate both dynamics, the variance risk premium is determined according to Proposition 2, and the state variables evolve according to processes derived in Appendix D. Panel B reports the results for the Heston (1993) model where the variance risk premium is defined as the product of a constant h and the spot variance. For each model, we also report the sum of the implied volatility squared errors based on 6,292 SPX put observations. All parameters are in annual units. For some of the parameters, we use the scientific notation E+/-*n* to denote the power of 10.

Table 10: Model Fit

Year	IRW Model	Heston Model	Difference
	Panel A: Subsample IVRMSE		
1997-1999	4.317	3.890	0.427
2000-2002	2.613	2.411	0.202
2003-2005	2.545	3.358	-0.813
2006-2008	3.115	4.027	-0.912
2009-2011	3.114	4.120	-1.006
Average	3.141	3.561	-0.420

Moneyness	IRW Model	Heston Model	Difference
	Panel B: IVRMSE by Moneyness		
S/K≤0.95	3.666	3.542	0.124
0.95<S/K≤1.05	3.424	3.675	-0.251
S/K>1.05	4.128	4.292	-0.164
Average	3.739	3.836	-0.097

Months to Maturity	IRW Model	Heston Model	Difference
	Panel C: IVRMSE by Maturity		
2 months	3.829	3.912	-0.082
3 months	3.452	3.800	-0.348
6 months	3.417	4.004	-0.586
Average	3.566	3.905	-0.339

Notes to Table: We present the fit of the IRW and Heston models based on a sample of 131,638 put option prices. For each model, the structural parametes are set to their optimal values. Model fit is evaluated based on the percentage IVRMSE. In Panel A, we present the IVRMSE for several subsamples. Panel B reports the performance of each model by moneyness. In Panel C, we present the models' IVRMSE by maturity.

Table 11: Out-of-Sample Fit

Panel A: Subsample IVRMSE

Year	IRW Model	Heston Model	Difference
1997-1999	4.209	3.736	0.473
2000-2002	2.642	2.468	0.174
2003-2005	2.572	3.390	-0.819
2006-2008	2.928	3.775	-0.847
2009-2011	3.110	3.967	-0.856
Average	3.092	3.467	-0.375

Panel B: IVRMSE by Moneyness

Moneyness	IRW Model	Heston Model	Difference
S/K≤0.95	3.275	3.452	-0.176
0.95<S/K≤1.05	3.338	3.614	-0.277
S/K>1.05	3.791	3.979	-0.188
Average	3.468	3.682	-0.214

Panel C: IVRMSE by Maturity

Months to Maturity	IRW Model	Heston Model	Difference
2 months	3.531	3.694	-0.163
3 months	3.301	3.642	-0.341
6 months	3.327	3.898	-0.571
Average	3.386	3.745	-0.358

Notes to Table: We present the out-of-sample fit for the IRW and Heston models based on a sample of 131,638 put option prices. For each model, the structural parameters are set to their optimal values. For the IRW model, we set the spot variance and inventory risk on any given day to their one-day-ahead forecast, given their values on the previous day. Similarly, the spot variance used for the Heston model on any given day is set to the one-day-ahead predicted value, taking the spot variance on the previous day as given. We compute the model IVRSME based on these predicted values. In Panel A, we present the IVRMSE for several subsamples. Panel B reports the performance of each model by moneyness. In Panel C, we present the models' IVRMSE by maturity.

Table A.1: Estimation Results for the HAR-RV Model and Variance Risk Premia

Panel A: Parameter Estimates and Fit for the HAR-RV Model

	Dependent Variable: $\ln(RV_{t,T})$				
	T=30	*T=60*	*T=90*	*T=180*	*T=270*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Intercept	-4.10 **	-5.12 **	-5.29 **	-4.11 *	-3.75 **
	(*0.03*)	(*0.02*)	(*0.03*)	(*0.01*)	(*0.02*)
$\ln(RV_{t-1,1})$	0.11	0.06	0.04	0.00	0.01
	(*0.13*)	(*0.28*)	(*0.36*)	(*0.48*)	(*0.46*)
$\ln(RV_{t-6,6})$	0.18	0.10	0.06	0.00	0.00
	(*0.26*)	(*0.43*)	(*0.46*)	(*0.48*)	(*0.46*)
$\ln(RV_{t-30,30})$	-0.02	0.08	0.02	0.02	-0.01
	(*0.35*)	(*0.36*)	(*0.33*)	(*0.29*)	(*0.27*)
$\ln(RV_{t-60,60})$	*0.28*	*0.08*	*0.08*	*-0.13*	*-0.04*
	(*0.26*)	(*0.29*)	(*0.29*)	(*0.26*)	(*0.24*)
$\ln(RV_{t-90,90})$	-0.23	-0.07	-0.15	0.08	0.06
	(*0.24*)	(*0.27*)	(*0.28*)	(*0.23*)	(*0.25*)
$\ln(RV_{t-120,120})$	*-0.37*	*-0.58*	*-0.45* *	*-0.10* *	*-0.06* *
	(*0.20*)	(*0.11*)	(*0.10*)	(*0.09*)	(*0.08*)
Adj. R^2 (%)	51	56	60	71	73

Panel B: Min, Max, and Average of Variance Risk Premia Implied by the HAR-RV model

	$VRP_{t,T} \times 100$				
	T=30	*T=60*	*T=90*	*T=180*	*T=270*
Min	-35.48	-30.06	-22.01	-14.72	-18.22
Max	12.16	6.08	16.27	50.02	17.63
Average	-2.38	-2.46	-2.38	-2.14	-2.08

Notes to Table: Panel A reports the average of the parameter estimates, *p*-values, and adjusted R-squares from estimating the HAR-RV model on every day in the sample. The *p*-values in parentheses are computed using Newey-West with 8 lags. We estimate the model daily using a rolling window of 252 observations and forecast future realized variance for different forecast horizons. Based on the model prediction and the expected risk-neutral variance inferred from option prices, we construct measures of the variance risk premium. Panel B reports summary statistics for the variance risk premia. See Appendix A for additional information on the estimation methodology. The sample period is 1997-2011.

Table A.2: Time Variation in the Term Structure of Log-Variance Risk Premia Implied by the HAR-RV Model

	Dependent Variable: $\Delta LogVRP_{t,T} \times 100$				
	T=30	*T=60*	*T=90*	*T=180*	*T=270*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Intercept	-0.14	-0.07	-0.07	-0.12	-0.10
	(*0.60*)	(*0.73*)	(*0.72*)	(*0.34*)	(*0.24*)
InvRisk(*t-1*)	0.92 ***	0.70 ***	0.59 ***	0.31 ***	0.23 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.01*)	(*0.01*)
ΔW(*t*) × InvRisk(*t-1*)	-3.78 ***	-2.71 ***	-2.31 ***	-1.48 ***	-1.13 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
S&P500LogRet(*t*)	7.79 ***	6.39 ***	5.40 ***	4.19 ***	3.47 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
JumpFactor(*t*)	-3.78 ***	-2.24 ***	-1.74 ***	-1.05 ***	-0.84 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
NetBuyingPressure(*t*)	-0.59 ***	-0.50 ***	-0.38 ***	-0.09	-0.04
	(*0.01*)	(*0.00*)	(*0.01*)	(*0.27*)	(*0.49*)
Disagreement(*t*)	0.67 **	0.34 *	0.05	0.04	-0.18 *
	(*0.03*)	(*0.09*)	(*0.83*)	(*0.76*)	(*0.06*)
$\Delta LogVRP_{t-1,T}$	-5.06 ***	-2.75 ***	-2.06 ***	-0.65 ***	-0.27 **
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.03*)
Adj. R^2 (%)	40	42	42	46	49
N	3774	3774	3774	3774	3774

Notes to Table: We present the results from regressing daily changes in the log-variance risk premium, denoted $\Delta LogVRP_{t,T}$, on the explanatory variables. To measure expected physical variance, we fit the HAR-RV model on each day in the sample, as described in Appendix A, and we use it to forecast future physical variance. Based on the model forecast, we construct measures of the log-variance risk premium. We consider five horizons T to capture the term structure of variance risk premia. All regressors are standardized to have unit variance. The *p*-values are in parentheses and are computed using Newey-West with 8 lags. The sample period is 1997-2011.

Table A.3: Time Variation in the Term Structure of Log-Variance Risk Premia. Delta-Hedged Inventory Gains and Losses Calculated Using Ask Prices

	Dependent Variable: $\Delta LogVRP_{t,T} \times 100$				
	T=30	*T=60*	*T=90*	*T=180*	*T=270*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Intercept	-0.23	-0.16	-0.13	-0.10	0.03
	(*0.24*)	(*0.21*)	(*0.23*)	(*0.16*)	(*0.84*)
InvRisk(*t-1*)	0.94 ***	0.64 ***	0.39 ***	0.40 ***	0.25 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
ΔW(*t*) × InvRisk(*t-1*)	-3.40 ***	-2.53 ***	-2.06 ***	-1.46 ***	-1.29 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
S&P500LogRet(*t*)	7.43 ***	6.37 ***	5.56 ***	4.06 ***	3.50 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
JumpFactor(*t*)	-4.02 ***	-2.34 ***	-1.83 ***	-1.09 ***	-0.77 ***
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)
NetBuyingPressure(*t*)	-0.15	-0.17	-0.15	-0.07	0.06
	(*0.45*)	(*0.21*)	(*0.17*)	(*0.34*)	(*0.62*)
Disagreement(*t*)	-0.59 ***	-0.44 **	-0.52 ***	-0.32 ***	-0.41 ***
	(*0.02*)	(*0.03*)	(*0.00*)	(*0.00*)	(*0.00*)
$\Delta LogVRP_{t-1,T}$	-2.95 ***	-1.47 ***	-1.19 ***	-0.52 ***	-0.15
	(*0.00*)	(*0.00*)	(*0.00*)	(*0.00*)	(*0.23*)
Adj. R^2 (%)	43	47	49	53	20
N	3774	3774	3774	3774	3774

Notes to Table: We present the results from regressing daily changes in the log-variance risk premium, denoted $\Delta LogVRP_{t,T}$, on the explanatory variables. Each day, market makers' delta-hedged inventory profits and losses are calculated using option ask prices. We consider five horizons T to capture the term structure of variance risk premia. All regressors are standardized to have unit variance. The p-values are in parentheses and are computed using Newey-West with 8 lags. The sample period is 1997-2011.