A GARCH Option Model with Variance-Dependent Pricing Kernel[*]

Peter Christoffersen   Steven Heston

Rotman, CBS, and CREATES   University of Maryland

Kris Jacobs

University of Houston and Tilburg University

March 10, 2011

Abstract

We develop a GARCH option model with a variance premium by combining the Heston-Nandi (2000) dynamic with a new pricing kernel. While the pricing kernel is monotonic in the stock return and in variance, its projection onto the stock return is nonmonotonic. A negative variance premium makes it appear U-shaped. We present new semi-parametric evidence to confirm this U-shaped relationship between the risk-neutral and physical probability densities. The new pricing kernel substantially improves our ability to reconcile the time series properties of stock returns with the cross-section of option prices. It provides a unified explanation for the implied volatility puzzle, the overreaction of long-term options to changes in short-term variance, and the fat tails of the risk-neutral return distribution relative to the physical distribution.

JEL Classification: G12

Keywords: Pricing kernel; variance premium; stochastic volatility; GARCH; overreaction.

1 Introduction

Continuous-time models have become the workhorse of modern option pricing theory. They typically offer closed-form solutions for European option values, and have the flexibility of incorporating stochastic volatility (SV) with leverage effects and various types of risk premia (Heston, 1993, Bakshi, Cao and Chen, 1997). Stochastic jumps and jump risk premia can capture additional variation in the conditional distribution of returns (Broadie, Chernov and Johannes, 2007). However, the resulting models can be cumbersome to estimate with time-series data, due to the need to filter the unobserved stochastic volatility and jump intensity.

In contrast, discrete-time GARCH models (Engle, 1982, Bollerslev, 1986) dominate the time-series literature. They are easy to filter and estimate with multiple factors (Engle and Lee, 1999), long memory (Bollerslev and Mikkelsen, 1996), and non-Gaussian innovations (Bollerslev, 1987, and Nelson, 1991). The GARCH framework offers four advantages for the purpose of empirical option valuation. First, it may be considered an accurate numerical approximation to a continuous-time model (Nelson, 1992 and 1996), or an internally consistent framework that makes exact predictions at the frequency of available data, avoiding any discretization bias.[1] Second, its predictions are exactly compatible with the filter used to extract the variance. Third, the volatility prediction performance of GARCH is often found to be very similar to that of SV models (Fleming and Kirby, 2003). Fourth, estimation is computationally fast.

Several valuation results are available for European option pricing with a GARCH dynamic,[2] but all existing models are characterized by the same limitation. The filtering problem in these models is straightforward because the distribution of one-period returns has a known conditional variance. This does not severely restrict variance modeling, but it has implications for option pricing. Because the models do not contain an independent adjustment for variance risk, they do not offer much flexibility in the modeling of variance risk premia.

Instead, the specification of variance premia in GARCH models has to be carefully integrated with the specification of the equity risk premium, which requires an explicit specification of the pricing kernel. Existing models use the risk-neutral valuation relationship of Rubinstein (1976) and Brennan (1979). This produces the Black-Scholes (1973) formula for one-period options, and limits the models' ability to explain longer-term option prices. This paper instead specifies a variance-dependent pricing kernel and combines it with the Heston-Nandi (2000) dynamic. The variance-dependent pricing kernel implies a quadratic pricing kernel when viewed

---

[1] See Lamoureux and Paseka (2009) on the role of discretization bias in implementations of the Heston (1993) model.

[2] See Duan (1995) and Heston and Nandi (2000) for the case of conditionally normal innovations, and Barone-Adesi, Engle and Mancini (2008) and Christoffersen, Heston, and Jacobs (2006) for applications with non-normal innovations. See Duan and Simonato (2001) and Ritchken and Trevor (1999) for numerical methods for American option valuation.

as a function of the stock return. This generates an additional risk premium associated with variance, but uncorrelated with equity risk, making the model a discrete-time analog of Heston's (1993) stochastic volatility model. Using the new pricing kernel, we obtain closed-form solutions for option values based on a risk-neutral variance process that has the same functional form as the physical variance process, but that differs from the physical variance process along several dimensions. The model nests and substantially generalizes the Heston-Nandi (2000) framework.

We demonstrate that the new model can address some existing option pricing puzzles. Most importantly, it not only captures the fat tails of the option-implied return distribution, commonly referred to as the implied volatility smile, it also explains why these tails are fatter than those of the return-based distribution. Bates (1996a) emphasized the importance of this question: "the central empirical issue in option research is whether the distributions implicit in option prices are consistent with the time series properties of the underlying asset prices." While subsequent studies have addressed this issue, it has proved difficult to reconcile the empirical distributions of spot returns with the risk-neutral distributions underlying option prices.

The new model also addresses several other empirical puzzles that have emerged from the options literature. A well-known puzzle is that volatilities implied by option prices tend to exceed realized volatility. This puzzle is well-known and understood in terms of a negative price of variance risk. Another variance puzzle is the expectations puzzle: implied variances do not provide an unbiased forecast of subsequent variance.[3] Furthermore, Stein (1989) and Poteshman (2001) show that long-term implied variance overreacts to changes in short-term variance. This puzzle involves movements in the term structure of implied volatility and is related to the expectations puzzle. Taken together, these anomalies indicate misspecification in the dynamic relationship between option values and the time series of spot returns. Because they are usually not discussed in the context of a parametric framework, the literature has not explicitly linked them to Bates' statement, but they are intimately related. In addition to these longitudinal expectations puzzles, available models have difficulty explaining the cross-section of option prices, particularly the prices of out-of-the-money options. It has been recognized that this evidences a "pricing kernel puzzle," in the sense that available pricing kernels may not be general enough to explain option data.[4]

Together the puzzles pose a collective challenge to option models. We attempt to provide a unified explanation for these puzzles by formulating a more general pricing kernel. Our analytical results indicate that the suggested pricing kernel is able to qualitatively account for the puzzles. We also want to demonstrate that the new pricing kernel can quantitatively explain option prices

---

[3]See Day and Lewis (1992), Canina and Figlewski (1993), Lamoureux and Lastrapes (1993), Jorion (1995), Fleming (1998), Blair, Poon and Taylor (2001), and Chernov (2007) among others.

[4]See for instance Brown and Jackwerth (2001), Bates (2008), and Bakshi, Madan, and Panayotov (2010).

and resolve the discrepancy between option prices and the time series of underlying index returns pointed out by Bates (1996a). So we implement an empirical analysis using an objective function with a return component and an option component. The discrete-time GARCH structure of the model facilitates filtering, and its numerical efficiency makes it possible to maximize this objective function in a large-scale empirical exercise.

The empirical results are quite striking. Imposing the new variance dependent pricing kernel dramatically improves model fit compared to a traditional pricing kernel with equity risk only. The new model reduces valuation biases across strike price and maturity, and the resulting fit is reasonably close to that of an unrestricted ad-hoc model. The new pricing kernel adequately captures the premium of risk-neutral variance relative to physical variance, as well as the higher risk-neutral volatility of variance. While the estimated persistence of risk-neutral variance is larger than the persistence of physical variance, the difference is smaller than in the ad-hoc model, indicating that the new pricing kernel qualitatively but not quantitatively captures this stylized fact. Presumably this is due to the fact that capturing variance persistence is not heavily weighted in the likelihood.

A number of existing studies on option valuation and general equilibrium modeling are related to our findings. Several studies have argued that modifications to standard preferences are needed to explain option data.[5] Ait-Sahalia and Lo (2000) and Jackwerth (2000) have noted the surprising implications of option prices for risk-aversion, and Shive and Shumway (2006) suggest using non-monotonic pricing kernels. Rosenberg and Engle (2002) and Chernov (2003) document nonmonotonicities in pricing kernels using parametric assumptions on the underlying returns. Chabi-Yo (2009) documents nonmonotonicities after projecting on the market return. Brown and Jackwerth (2001) argue that in order to explain option prices, the pricing kernel needs a momentum factor. Bollerslev, Tauchen, and Zhou (2009) show that incorporating variance risk in the pricing kernel can explain why option volatilities predict market returns. Bakshi, Madan, and Panayotov (2010) show that the prices of S&P500 calls are inconsistent with monotonically declining kernels, and that the mimicking portfolio for the pricing kernel is U-shaped.

The remainder of the paper is organized as follows. Section 2 reviews the standard SV model and presents a new discrete-time GARCH model incorporating a quadratic pricing kernel. Section 3 discusses a number of stylized facts and also presents new evidence on the shape of the conditional pricing kernel. Section 4 estimates the new GARCH model with a quadratic pricing kernel jointly on returns and options, and Section 5 concludes. The Appendix collects proofs of propositions.

---

[5]See for instance Bates (2008), Pan (2002), Benzoni, Collin-Dufresne, and Goldstein (2006), and Liu, Pan and Wang (2004).

2 From Continuous-Time Stochastic Volatility to Discrete-Time GARCH

In order to value options, we need both a statistical description of the physical process and a pricing kernel. This section reviews the Heston (1993) model and then builds a GARCH option model with the same features.[6] The Heston (1993) model assumes the following dynamics for the spot price $S(t)$

$$
\begin{aligned}
dS(t) &= (r + \mu v(t))S(t)\,dt + \sqrt{v(t)}S(t)\,dz_1(t), \\
dv(t) &= \kappa(\theta - v(t))dt + \sigma\sqrt{v(t)}\left(\rho dz_1(t) + \sqrt{1 - \rho^2}dz_2(t)\right),
\end{aligned}
\tag{1}
$$

where r is the risk-free interest rate, the parameter μ governs the equity premium, and $z_1(t)$ and $z_2(t)$ are independent Wiener processes. The notation in (1) emphasizes the separate sources of equity risk, $z_1(t)$, and independent volatility risk, $z_2(t)$. An important aspect of our analysis is the separate premia for these risks.

In addition to the physical dynamics (1), we assume the pricing kernel takes the exponential-affine form

$$
M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^{\phi}\exp\left(\delta t + \eta\int_0^t v(s)ds + \xi(v(t) - v(0))\right),
\tag{2}
$$

where parameters δ and η govern the time-preference, while ϕ and ξ govern the respective aversion to equity and variance risk. When variance is constant, (2) amounts to the familiar power utility from Rubinstein's (1976) preference-based derivation of the Black-Scholes model. But with stochastic variance, it has distinctive implications for option valuation.[7] Appendix A shows the risk-neutral process takes the form

$$
\begin{aligned}
dS(t) &= rS(t)\,dt + \sqrt{v(t)}S(t)\,dz_1^*(t), \\
dv(t) &= (\kappa(\theta - v(t)) - \lambda v(t))dt + \sigma\sqrt{v(t)}(\rho dz_1^*(t) + \sqrt{1 - \rho^2}dz_2^*(t)),
\end{aligned}
\tag{3}
$$

where $z_1^*(t)$ and $z_2^*(t)$ are independent Wiener processes under the risk-neutral measure and the reduced-form parameter λ governs the variance risk premium. The pricing kernel in (2) is the unique arbitrage-free specification consistent with both the physical (1) and risk-neutral (3) dynamics. We can express the equity premium μ and variance premium λ parameters in terms

---

[6]See Hull and White (1987), Melino and Turnbull (1990), and Wiggins (1987) for other examples of option valuation with stochastic volatility.

[7]Stochastic variance in the pricing kernel could result, for instance, if $v(t)$ governs the variance of aggregate production in a Cox-Ingersoll-Ross (1985) model with non-logarithmic utility. It could also result from the model of Benzoni, Collin-Dufresne, and Goldstein (2006) where uncertainty directly affects preferences. See also Bakshi, Madan, and Panayotov (2010) who consider short-sale constraints.

of the underlying preference parameters ϕ and ξ

$$
\begin{aligned}
\mu &= -\phi - \xi\sigma\rho, \\
\lambda &= -\rho\sigma\phi - \sigma^2\xi = \rho\sigma\mu - (1-\rho^2)\sigma^2\xi.
\end{aligned}
\tag{4}
$$

This allows us to interpret both the equity risk premium μ and the variance risk premium λ in terms of two distinct components originating in preferences. One component is related to the risk-aversion parameter ϕ and the other one to the variance preference parameter ξ. We can therefore use economic intuition to sign the equity premium and the variance premium. If the pricing kernel is decreasing in the spot price, we have $\phi < 0$, because marginal utility is a decreasing function of stock index returns. If hedging needs increase in times of uncertainty then we anticipate the pricing kernel to be increasing in volatility, $\xi > 0$. Empirically the correlation between stock market returns and variance ρ is strongly negative. Therefore, from (4) the equity premium μ must be positive. The variance premium λ has a component based on covariance with equity risk, and a separate independent component based on the variance preference ξ. With a negative correlation ρ, we see that λ must be negative.

It is important to note that the reduced-form risk-neutral dynamics of variance in (3) do not distinguish whether the variance risk premium $\lambda v(t)$ emanates exclusively from ϕ (and therefore indirectly from the equity premium μ) or whether it has an independent component ξ. In other words, assuming $\xi = 0$ in (2) is consistent with a nonzero variance risk premium λ, as can be seen from (4). Therefore, when estimating option models with stochastic volatility using both return data and option data, it is important to explicitly write down the pricing kernel that provides the link between the physical dynamic (1) and the risk-neutral dynamic (3). It is not sufficient to simply state that (3) holds for arbitrary (negative) λ, because this assumption is consistent with the pricing kernel (2) but also with the special case with $\xi = 0$, and the economic implications of those sets of assumptions are very different. This paper explores the distinct implications of variance premium $\xi \neq 0$ for option prices.

The option pricing model in (1), (2), and (3) captures important stylized facts, and can be combined with stochastic jumps and jump risk premia.[8] However, the resulting models are often cumbersome to estimate because of the complexity of the resulting filtering problem. A discrete-time analog of the physical square-root return process (1) is the Heston-Nandi (2000) GARCH

---

[8] Andersen, Benzoni, and Lund (2002), Bakshi, Cao and Chen (1997), Bates (1996b, 2000, 2006), Broadie, Chernov, and Johannes (2007), Chernov and Ghysels (2000), Eraker (2004), Eraker, Johannes, and Polson (2003), and Pan (2002) investigate jumps in returns. Broadie, Chernov, and Johannes (2007), Eraker (2004), and Eraker, Johannes, and Polson (2003) estimate models with additional jumps in volatility. Bates (2010), Carr and Wu (2004) and Huang and Wu (2004) investigate infinite-activity Levy processes. Bates (2000) and Christoffersen, Heston, and Jacobs (2009) investigate multifactor volatility models.

process

$$\ln(S(t)) = \ln(S(t-1)) + r + (\mu - \frac{1}{2})h(t) + \sqrt{h(t)}z(t), \tag{5}$$

$$h(t) = \omega + \beta h(t-1) + \alpha(z(t-1) - \gamma\sqrt{h(t-1)})^2,$$

where r is the daily continuously compounded interest rate and $z(t)$ has a standard normal distribution. We will implement this model using daily data, and we are therefore interested in its predictions for a fixed daily interval. Paralleling properties of the diffusion model (1), the expected future variance is a linear function of current variance

$$E_{t-1}(h(t+1)) = (\beta + \alpha\gamma^2)h(t) + (1 - \beta - \alpha\gamma^2)E(h(t)), \tag{6}$$

where $E(h(t)) = (\omega + \alpha)/(1 - \beta - \alpha\gamma^2)$. In words, the variance reverts to its long-run mean with daily autocorrelation of $\beta + \alpha\gamma^2$. The conditional variance of the $h(t)$ process is also linear in past variance.

$$Var_{t-1}(h(t+1)) = 2\alpha^2 + 4\alpha^2\gamma^2 h(t). \tag{7}$$

The parameter γ determines the correlation of the variance $h(t+1)$ with stock returns $R(t) = \ln(S(t)/S(t-1))$, via

$$Cov_{t-1}(R(t), h(t+1)) = -2\alpha\gamma h(t) \tag{8}$$

The data robustly indicate sizeable negative correlation, which means that γ must be positive.

In existing GARCH models, Duan (1995) and Heston and Nandi (2000) use Rubinstein's (1976) power pricing kernel. In a lognormal context, this is equivalent to using the Black-Scholes formula for one-period options. Instead, we value securities using a discrete analogue of the continuous-time pricing kernel in (2), namely

$$M(t) = M(0) \left(\frac{S(t)}{S(0)}\right)^\phi \exp\left(\delta t + \eta \sum_{s=1}^{t} h(s) + \xi(h(t+1) - h(1))\right). \tag{9}$$

The discrete-time specification (9) is identical to the continuous pricing kernel (2) with the integral replaced by a summation. Recall that in the diffusion model, the variance process follows square-root dynamics with different parameters under the physical (1) and risk-neutral (3) measures. The following proposition shows an analogous result in the discrete model – the risk-neutral process remains in the same GARCH class.

Proposition 1 *The risk-neutral stock price process corresponding to the physical Heston-Nandi*

GARCH process (5) and the pricing kernel (9) is the GARCH process

$$\ln(S(t)) = \ln(S(t-1)) + r - \frac{1}{2}h^*(t) + \sqrt{h^*(t)}z^*(t), \tag{10}$$

$$h^*(t) = \omega^* + \beta h^*(t-1) + \alpha^*(z^*(t-1) - \gamma^*\sqrt{h^*(t-1)})^2,$$

where $z^*(t)$ has a standard normal distribution and

$$h^*(t) = h(t)/(1 - 2\alpha\xi), \tag{11}$$

$$\omega^* = \omega/(1 - 2\alpha\xi),$$

$$\alpha^* = \alpha/(1 - 2\alpha\xi)^2,$$

$$\gamma^* = \gamma - \phi.$$

Proof. See Appendix B. ∎

The risk-neutral dynamics differ from the physical dynamics through the effect of the equity premium parameter μ and scaling factor $(1 - 2\alpha\xi)$. Conditional on the parameters characterizing the physical dynamic, these risk-neutral dynamics are therefore implied by the values of the kernel parameters ϕ and ξ in equation (9).[9] The intuition is similar to the continuous-time case in (4), where the values of the equity premium and volatility risk premium parameters μ and λ are implied by the values of the kernel parameters ϕ and ξ.

It can be seen from (11) that a nonzero ξ parameter has important implications, because it influences the level, persistence, and volatility of the variance. In contrast to the Heston-Nandi (2000) model, the risk-neutral variance $h^*(t)$ differs from the physical variance $h(t)$.[10] When $\alpha\xi > 0$, the pricing kernel puts more weight on the tails of innovations and the risk-neutral variance $h^*(t)$ exceeds the physical variance $h(t)$. The Heston-Nandi (2000) model corresponds to the special case of $\xi = 0$. The variance premium also affects the risk-neutral drift of $h^*(t)$

$$E^*_{t-1}(h^*(t+1)) = (\beta + \alpha^*\gamma^{*2})h^*(t) + (1 - \beta - \alpha^*\gamma^{*2})E^*(h^*(t)), \tag{12}$$

where $E^*(h^*(t)) = (\omega^* + \alpha^*)/(1 - \beta - \alpha^*\gamma^{*2})$. The risk-neutral autocorrelation equals $\beta + \alpha^*\gamma^{*2}$, and a negative variance premium ($\xi > 0$) increases the risk-neutral persistence as well as the level of the future variance.

Comparison of physical parameters with risk-neutral parameters shows that if the correlation

---

[9]The mapping between μ and ϕ is contained in Appendix B. With an annual U.S. equity premium $\mu h(t)$ of around 8% and variance $h(t)$ of 20%^{2}, it can be inferred that the value of the equity premium parameter μ is small, around 2.

[10]In diffusion models, the instantaneous variance is identical under the physical and risk-neutral measures, but the risk-neutral variance will differ from the physical variance over a discrete interval such as one day.

between returns and variance is negative ($\gamma > 0$), if the equity premium is positive ($\mu > 0$, which corresponds to $\phi < 0$) and if the variance premium is negative ($\xi > 0$), then the risk-neutral mean reversion will be smaller than the actual mean reversion. Finally, note that the variance premium alters the conditional variance of the risk-neutral variance process

$$Var^*_{t-1}(h^*(t+1)) = 2\alpha^{*2} + 4\alpha^{*2}\gamma^{*2}h^*(t). \tag{13}$$

If the correlation between returns and variance is negative ($\gamma > 0$), the equity premium is positive ($\mu > 0$), and the variance premium is negative ($\xi > 0$), then substituting the risk-neutral parameters α^* and γ^* from (10) shows that the risk-neutral variance of variance is greater than the actual variance of variance. Furthermore we can define the risk-neutral conditional covariance

$$Cov^*_{t-1}(R(t), h^*(t+1)) = -2\alpha^*\gamma^*h^*(t). \tag{14}$$

The following corollary summarizes the results for this discrete-time GARCH model, which parallel those of the continuous-time model.

Corollary 1 *If the equity premium is positive ($\mu > 0$), the independent variance premium is negative ($\xi > 0$), and variance is negatively correlated with stock returns ($\gamma > 0$) then:*

1. *The risk-neutral variance $h^*(t)$ exceeds the physical variance $h(t)$,*

2. *The risk-neutral expected future variance exceeds the physical expected future variance,*

3. *The risk-neutral variance process is more persistent than the physical process, and*

4. *The risk-neutral variance of variance exceeds the physical variance of variance.*

Corollary 1 summarizes how a premium for volatility can explain a number of puzzles concerning the level and movement of implied option variance compared to observed time-series variance. The final puzzle concerns the stylized fact pointed out by Bates (1996b), and more recently by Broadie, Chernov, and Johannes (2007), that the physical and risk-neutral volatility smiles differ, which corresponds to risk-neutral negative skewness and kurtosis exceeding physical negative skewness and kurtosis. Our model captures this stylized fact through a U-shaped pricing kernel. Interestingly, even though the pricing kernel in (9) is a monotonic function of the stock price and variance, the projection of the pricing kernel onto the stock price alone can have a U-shape. The following corollary formalizes this relationship.

Corollary 2 *The logarithm of the pricing kernel is a quadratic function of the stock return.*

$$\ln\left(\frac{M(t)}{M(t-1)}\right) = \frac{\xi\alpha}{h(t)}(R(t)-r)^2 - $$
$$\mu(R(t)-r) + \left(\eta + \xi(\beta-1) + \xi\alpha\left(\mu - \tfrac{1}{2} + \gamma\right)^2\right)h(t) + \delta + \xi\omega + \phi r, \quad (15)$$

where $R(t) = \ln(S(t)/S(t-1))$.

Proof. See Appendix C. ∎

In words, the pricing kernel is a parabolic curve when plotted in log-log space. Note that whether this shape is a positive smile or a negative frown depends on the independent variance premium ξ, not on the total variance premium. Due to the component of variance premium that is correlated with equity risk, it is conceivable that the total variance premium could have a different sign than the independent negative component. A negative independent variance premium ($\xi > 0$) corresponds to a U-shaped pricing kernel and thus a strong option smile. The shape of the option smile therefore provides a revealing diagnostic on the underlying preferences.

Corollary 3 *When the independent variance premium is negative ($\xi > 0$), the pricing kernel has a U-shape.*

In summary, the model allows option values to display an implied variance process that is larger, more persistent, and more volatile than observed variance. The resulting risk-neutral distribution will have higher variance and fatter tails than the physical distribution. This increases the values of all options, particularly long-term options and out-of-the-money options. A negative premium for variance therefore potentially explains a number of puzzles regarding the cross-section of option prices, the relationship between physical volatility and option-implied volatility, and the relative thickness of the tails of the physical and risk-neutral distribution.

Note that option valuation with this model is straightforward. Following Heston and Nandi (2000), the value of a call option at time t with strike price X maturing at T is equal to

$$C(S(t), h(t+1), X, T) = S(t)\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \text{Re}\left[\frac{X^{-i\varphi}g_t^*(i\varphi+1)}{i\varphi}\right]d\varphi\right)$$
$$-X\exp^{-r(T-t)}\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \text{Re}\left[\frac{X^{-i\varphi}g_t^*(i\varphi)}{i\varphi}\right]d\varphi\right). \quad (16)$$

where $g_t^*(.)$ is the conditional generating function for the risk-neutral process in (10). Heston and Nandi (2000) provide a closed-form solution for $g_t(.)$. Using the risk-neutral parameter mapping (11) this yields $g_t^*(.)$. Put options can be valued using the put-call parity.

3 Stylized Facts in Index Option Markets

Our new model nests the existing Heston-Nandi model, and so when fitted to the data it will trivially perform better in sample. Bates (2003) has argued that even traditional out-of-sample evaluations tend to favor more heavily parameterized models because the empirical patterns in option prices are so persistent over time. Before fitting the model to the data, we therefore assess the model by providing some relatively model-free empirical evidence on the pricing kernel implied in index returns and index option prices. We then compare this evidence with the model properties outlined in the above corollaries. In Section 4 below, we will subsequently estimate the model parameters and in detail quantify the ability of the model to simultaneously fit the physical and risk-neutral distributions.

We first discuss the option and return data used in the empirical analysis, and then document and analyze a number of well-known and lesser-known stylized facts in option markets. We pay particular attention to the shape of the pricing kernel implied by option data. Subsequently we discuss how the new model addresses these stylized facts.

3.1 Data

Our empirical analysis uses out-of-the-money S&P500 call and put options for the 1996-2004 period from OptionMetrics. Rather than using a short time series of daily option data, we use an extended time period, but we select option contracts for one day per week only. This choice is motivated by two constraints. On the one hand, it is important to use as long a time period as possible, in order to be able to identify key aspects of the model. See for instance Broadie, Chernov, and Johannes (2007) for a discussion. On the other hand, despite the numerical efficiency of our model, the optimization problems we conduct are very time-intensive, because we use very large cross-sections of option contracts. Selecting one day per week over a long time period is therefore a useful compromise. We use Wednesday data, because it is the day of the week least likely to be a holiday. It is also less likely than other days such as Monday and Friday to be affected by day-of-the-week effects. Moreover, following the work of Dumas, Fleming and Whaley (1998) and Heston and Nandi (2000), several studies have used a long time series of Wednesday contracts.

Table 1 presents descriptive statistics for the option data by moneyness and maturity. Moneyness is defined as implied futures price F divided by strike price X. When F/X is smaller than one, the contract is an out-of-the-money (OTM) call, and when F/X is larger than one, the contract is an OTM put. The out-of-the-money put prices were converted into call prices using put-call parity. The sample includes a total of 21,391 option contracts with an average mid-price of $28.42 and average implied volatility of 21.47%. The implied volatility is largest for

11

the OTM put options, reflecting the well-known volatility smirk in index options. The average implied volatility term structure is roughly flat during the period.

Table 1 also presents descriptive statistics for the return sample. The return sample is from January 1, 1990 to December 31, 2005. It is longer than the option sample, in order to give returns more weight in the optimization, as explained in more detail below. The standard deviation of returns, at 16.08%, is substantially smaller than the average option-implied volatility, at 21.47%. The higher moments of the return sample are consistent with return data in most historical time periods, with a very small negative skewness and substantial excess kurtosis. Table 1 also presents descriptive statistics for the return sample from January 1, 1996 to December 31, 2004, which matches the option sample. In comparison to the 1990-2005 sample, the standard deviation is somewhat higher. Average returns, skewness and kurtosis in the subsample are very similar to the 1990-2005 sample.

3.2   Fat Tails and Fatter Tails

We now document the shape of the conditional pricing kernel using semiparametric methods. The literature does not contain a wealth of evidence on this issue. Much of what we know is either entirely (see for instance Bates, 1996b) or partly (Rosenberg and Engle, 2002) filtered through the lens of a parametric model.

Among the papers that study risk-neutral and physical densities, Jackwerth (2000) focuses on risk aversion instead of the (obviously related) shape of the pricing kernel. Ait-Sahalia and Lo (2000, p. 36) provide a picture of the pricing kernel as a by-product of their analysis of risk aversion, but because of their empirical technique, their estimate is most usefully interpreted as an unconditional pricing kernel. Our focus is on the conditional pricing kernel. Shive and Shumway (2006) and Bakshi, Madan, and Panayotov (2010) present the most closely related evidence on the conditional pricing kernel, but our conditioning approach is very different.

It is relatively straightforward to estimate the risk-neutral conditional density of returns using option data, harnessing the insights of Breeden and Litzenberger (1978) and Banz and Miller (1978), and there is an extensive empirical literature reporting on this. Ait-Sahalia and Lo (2000) obtain non-parametric estimates of the risk-neutral density or state-price density. This necessitates combining option data on different days, because non-parametric methods are very data intensive. Other papers, such as Jackwerth and Rubinstein (1996), Jackwerth (2000), Rubinstein (1994), Bliss and Panigirtzoglou (2004), Rosenberg and Engle (2002), and Rompolis and Tzavalis (2008) use option data on a single day to infer risk-neutral densities, using a variety of methods.

Our objective is to stay as nonparametric as possible, but to provide evidence on the condi-

12

tional density. We therefore need to impose a minimum of parametric assumptions. We proceed as follows. Using the entire cross-section of options on a given day, we first estimate a second-order polynomial function for implied Black-Scholes volatility as a function of moneyness and maturity. Using this estimated polynomial, we then generate a grid of at-the-money implied volatilities for a desired grid of strikes. Call these generated implied volatilities $\hat{\sigma}(S(t), X, \tau)$. Call prices can then be obtained using the Black-Scholes functional form.

$$\hat{C}(S(t), X, \tau, r) = C_{BS}(S(t), X, \tau, r; \hat{\sigma}(S(t), X, \tau)). \tag{17}$$

Following Breeden and Litzenberger (1978), the risk-neutral density for the spot price on the maturity date $T = t + \tau$ is calculated as a simple function of the second derivative of the semiparametric option price with respect to the strike price

$$\hat{f}_t^*(S(T)) = \exp(r)\left[\frac{\partial^2 \hat{C}(S(t), X, \tau, r)}{\partial X^2}\right]_{|X=S(T)}. \tag{18}$$

We calculate this derivative numerically across a grid of strike prices for each horizon, setting the current interest rate to its average sample value.

Finally, in order to plot the density against log returns rather than future spot prices, we use the transformation

$$\hat{f}_t^*(R(t, T)) = \frac{\partial}{\partial u}\Pr\left(\ln\left(\frac{S(T)}{S(t)}\right) \leq u\right) = S(t)\exp(u)\hat{f}_t^*(S(t)\exp(u)). \tag{19}$$

The resulting densities are truly conditional because they only reflect option information for that given day.

It is much more challenging to construct the conditional physical density of returns. Available studies walk a fine line between using short samples of daily returns, which makes the estimate truly conditional, and using longer samples, which improves the precision of the estimates. Ait-Sahalia and Lo (1998) use a relatively long series because they are less worried about the conditional nature of the estimates. Jackwerth (2000) uses one month worth of daily return data because he wants to illustrate the time-varying nature of the conditional density. We use a somewhat different approach. We discuss the case of monthly returns, because this is consistent with the maturity of the options used in the empirical work, but the method can easily be applied for shorter- or longer-maturity returns.

Because we want to estimate the tails of the distribution as reliably as possible, we use a long daily time series of the natural logarithm of one-month returns, from January 1, 1990 to December 30, 2005. A histogram based on this time series is effectively an estimate of the unconditional

physical density of one-month log returns. We obtain a conditional density estimate for a given day by first standardizing the monthly return series by the sample mean $\bar{R}$ and the conditional one-month variance on that day, $h(t,T)$, as implied by the daily GARCH model in (5). This provides a series of return shocks $Z(t,T) = \left(R(t,T) - \bar{R}\right)/\sqrt{h(t,T)}$. We then construct a conditional histogram for a given day t day using the conditional variance for that day, $h(t,T)$, and the historical series of monthly shocks, Z. We write this estimate of the conditional physical distribution as

$$\hat{f}_t\left(R(t,T)\right) = \hat{f}\left(\bar{R} + \sqrt{h(t,T)}Z\right)$$

A subset of the resulting estimates of physical and risk-neutral conditional densities are given in Figure 1. Recall that our sample consists of nine years worth of option data, for 1996-2004, and that we use Wednesday data only when we estimate the models. We conduct the estimation of the conditional densities for each of the Wednesdays in our sample, which is straightforward to execute. We cannot report all these results because of space constraints. In order to show the time variation in the conditional densities, and the appeal of our method, Figure 1 presents nine physical and nine risk-neutral conditional densities, one for the first Wednesday of each year in our sample. The sample year is indicated in the title to each graph. The horizontal axis indicates annualized log returns. The results in Figure 1 are interesting because they illustrate that the conditional densities significantly change through time. The shapes of both the physical and the risk-neutral densities vary substantially over the years.

Figure 1 clearly demonstrates the fat left tail of the estimate of the risk-neutral conditional density, compared to that of the physical density. This finding is robust despite the fact that the conditional densities look very different across the years. This stylized fact gives rise to risk-neutral model estimates that display excess kurtosis and excess negative skewness in comparison to physical estimates. Figure 1 indicates that it is difficult to draw any definitive conclusion regarding the relative thickness of the right tail of the risk-neutral and physical density. The estimate of the right tail of the physical density is somewhat noisy, but more critically the right tail of the risk-neutral density is much harder to estimate than the left tail, because of the relative scarceness of traded out-of-the-money call options. Moreover, those options are usually thinly traded. Fortunately, the relative thickness of the right tail is inconsequential for establishing the nonlinearity of the logarithm of the ratio of the densities.

Figure 2 depicts the natural logarithm of the ratio of the estimates of the weekly conditional 1-month risk-neutral and conditional physical density. We want to investigate the natural logarithm of the pricing kernel at different levels of return. As in Figure 1, we present nine sets of results, one for each year of the sample. Recall that in Figure 1 we only present results for the first week of each year, in order to illustrate the time-varying nature of the conditional density. Plotting the

densities for all 52 weeks in a given year would make the figure unwieldy. In Figure 2, because the densities move together, we are able to present more information and plot results for all weeks of the year on each picture. Specifically, we plot

$$\ln\left(\hat{f}_t^*\left(R\left(t,T\right)\right)/\hat{f}_t\left(R\left(t,T\right)\right)\right), \quad \text{for } t = 1, 2, .., 52$$

In each week we trimmed 5% of observations in the left and right tails, because these observations are sometimes very noisy.

Two very important conclusions obtain. First, the pricing kernel is clearly not a monotonic function of returns, rejecting a hypothesis implicit in the Black-Scholes model and much of the option pricing literature. Second, the shape of the pricing kernel is remarkably stable across time. It is evident that the shape of the pricing kernel varies somewhat across certain years. For instance, the 1998 kernel is different from the 1996 kernel, and by 2004 the kernel again looks similar to the 1996 kernel. But we are able to draw the fifty-two pricing kernels generated for a given year on one picture to clearly illustrate the nonlinear nature of the logarithm of the kernel. If the kernel varied more within the year, Figure 2 would contain nothing but a cloudy scatter without much structure. Whether the logarithm of the kernel is exactly a quadratic function of stock returns is perhaps less obvious, because there is some noise in the estimates of the densities' right tail. However, it is clear that the relationship is nonlinear.

In summary, Figure 2 illustrates that the logarithm of the pricing kernel is nonlinear and roughly quadratic as a function of the return, and that this pricing relationship is relatively stable over time.

3.3 Returns on Straddles

A successful model for index options also has to address a number of other stylized facts and anomalies. It is well-known that on average, risk-neutral volatility exceeds physical volatility.[11] Several authors have argued that the risk premium that explains this difference makes it interesting to short sell straddles.[12] Figures 3 and 4 illustrate these stylized facts using the 1996-2004 option sample from Table 1. Figure 3 illustrates that risk-neutral volatility exceeds physical volatility, when both are filtered by a GARCH process. This stylized fact is robust to a large number of variations in the empirical setup, such as for instance measuring the physical volatility using a different filter, using realized volatility instead of GARCH volatility, or measuring risk-neutral volatility using the VIX.

---

[11]See for instance Bates (2000, 2003), Broadie, Chenov, and Johannes (2007), Chernov and Ghysels (2000), Eraker (2004), Heston and Nandi (2000), Jones (2003), and Pan (2002).

[12]See among others Coval and Shumway (2001), Bondarenko (2003), and Driessen and Maenhout (2007).

Figure 4 illustrates the returns and cumulative returns of a short straddle strategy, which for simplicity are computed using the nearest to at-the-money nearest to 30-day maturity call and put option on the third Friday of every month. The options are held until maturity, the cash account earns the risk-free rate, and the index starts out with $100 in cash on January 1, 1996. The dashed line in Figure 4 plots the S&P500 monthly closing price normalized to 100 in January 1996 for comparison. It is obvious from Figure 4 that the short straddle strategy was very rewarding in the 1996-2004 period, especially in periods when the S&P500 performed well. In the Black-Scholes model, the average return on this strategy would be approximately zero, and the strategy's returns would not be correlated with market returns.

3.4   The Overreaction Hypothesis

Stein (1989) documents another stylized fact in option markets that is equally robust, but has attracted somewhat less attention. He demonstrates using a simple regression approach that longer-term implied volatility overreacts to changes in shorter-term implied volatility. Stein's most general empirical test, which is contained in Table V of his paper, is motivated by the restriction

$$E_t \left[(IV_{t+(LT-ST)}^{ST} - IV_t^{ST}) - 2(IV_t^{LT} - IV_t^{ST}) \right] = 0, \tag{20}$$

where IV_t^{LT} is the implied volatility of a long-term option and IV_t^{ST} is the implied volatility of a short-term option that has half the maturity of the long-term option. Intuitively, this says that the slope of the term structure of implied volatility is equal to one half of the expected change in implied volatility. This restriction can be tested by regressing the time series in brackets on the left hand side on current information. Stein (1989) regresses on IV_t^{ST} and finds a negative sign, which is consistent with his overreaction hypothesis, as well as with his other empirical results. When the term structure of implied volatility is steep, then future implied volatilities tend to be below the forward forecasts implied by the term structure of volatility. In other words, long-term options seem to overreact to changes in short-term volatility.

We follow Stein's implementation of (20), using weekly time series of one-month and two-month implied volatilities. The regression is

$$(IV_{t+4}^{1M} - IV_t^{1M}) - 2(IV_t^{2M} - IV_t^{1M}) = a_0 + a_1 IV_t^{1M} + e_{t+4},$$

where $2M$ and $1M$ denote 2-month and 1-month maturity, and we test the null hypothesis that $a_1 = 0$.

Table 2 presents the results for the Stein regression using the 1996-2004 option data. Remember that the frequency of the time series of implied volatilities is weekly, as in Stein (1989),

making our results directly comparable to his. We use options that are at-the-money, according to the definition used in Table 1. Rather than averaging the two contracts that are closest to at-the-money, we fit a polynomial in maturity and moneyness to all option contracts on a given day, and then interpolate in order to obtain at-the-money implied volatility for the desired maturities. This strategy eliminates some of the noise from the data.

Table 2 demonstrates convincingly just how robust Stein's results are. We run the regressions first for the full sample 1996-2004, and subsequently for nine sub-samples, one for each of the years in the sample. We find a highly significant negative sign in all ten cases.

Stein (1989) interprets this stylized fact as an anomaly. Long-term options overreact to short-term fluctuations in implied volatility, even though volatility shocks decay very quickly. Stein (1989) therefore argues that this is a violation of rational expectations. We argue next that this robust stylized fact does not signal an anomaly but is entirely consistent with the new model developed in Section 2.

3.5   Stylized Facts and the Variance Dependent Pricing Kernel

In summary, this section has documented three stylized facts: First, the (log) pricing kernel appears to be quite robustly U-shaped. Second, option implied volatility is almost always higher than the physical volatility from index returns so that selling straddles is profitable on average. Third, long-term options tend to overreact to changes in short-term volatility.

Qualitatively, these findings match the model predictions captured in the three corollaries in Section 2. Corollary 1 shows that if we assume that the equity premium is positive and the independent variance premium is negative, and that variance is negatively correlated with stock returns, then in the model the risk-neutral variance will exceed the physical variance. Under realistic assumptions, the model thus qualitatively captures the fact that profits from selling straddles tend to be positive.

Corollary 1 also shows that under the same assumptions the risk-neutral variance process will be more persistent than the physical variance process in the model. This will qualitatively produce the Stein finding of overreaction: High persistence in the risk neutral variance will generate large reactions ("overreaction") in the model prices of long-term options, when short-term volatility changes.

Corollary 2 above shows that the daily log pricing kernel will be quadratic in the model and Corollary 3 shows that when the variance premium is negative then the daily log pricing kernel will be U-shaped. Thus at least in a qualitative sense, and at the daily frequency, the new model matches this stylized fact. Assessing whether the model generates a quantitatively adequate pricing kernel at the horizons of interest for option valuation requires estimation of the model's

parameters. This is the topic to which we now turn.

4   Estimating the Model

We now present a detailed empirical investigation of the model outlined in Section 2. It is important to realize that the model's success in quantitatively capturing some of the stylized facts we discuss in Section 3 can only be evaluated in an appropriately designed empirical experiment. Specifically, the model's ability to capture the differences between the physical and risk-neutral distributions requires fitting both distributions using the same, internally consistent set of parameters. Perhaps somewhat surprisingly, in the stochastic volatility option pricing literature such an exercise has only been attempted by a very limited number of studies. In order to understand the implications of our empirical results, a brief summary of the existing empirical literature on index options is therefore warranted.

While the theoretical literature on option valuation is grounded in an explicit description of the link between the risk-neutral and physical distribution, much of the empirical literature on index options studies the valuation of options without contemporaneously fitting the underlying returns. In fact, it is possible to fit separate cross-section of options while side-stepping the issue of return fit completely by parameterizing the volatility state variable.[13] When estimating multiple cross-sections, one can parameterize the volatility state variable in the same way, at the cost of estimating a high number of parameters,[14] or one can filter the volatility from underlying returns, using a variety of filters. Some papers take into account returns through the filtering exercise, but do not explicitly take into account returns in the objective function.[15] Eraker (2004) and Jones (2003) conduct a Bayesian analysis based on options and return data. A few studies take a frequentist approach using an objective function which contains an option data component as well as a return data component. Chernov and Ghysels (2000) and Pan (2002) do this in a method-of-moments framework, while Santa-Clara and Yan (2010) estimate parameters using a likelihood which contains a returns component and an options component.

The literature also contains comparisons of the risk-neutral and physical distribution. Bates (1996b) observes that parameters for stochastic volatility models estimated from option data cannot fit returns. Eraker, Johannes, and Polson (2003) show the reverse. Broadie, Chernov, and Johannes (2007) use parameters estimated from returns data, and subsequently estimate the jump risk premia needed to price options.

Our empirical setup is most closely related to Santa-Clara and Yan (2010). We use a joint

---

[13]See for instance the seminal paper by Bakshi, Cao and Chen (1997)

[14]See for instance Bates (2000), Christoffersen, Heston, and Jacobs (2009), and Huang and Wu (2004).

[15]See for instance Christoffersen and Jacobs (2004).

likelihood consisting of an option-based component and a return-based component which is relatively easy in a discrete time GARCH setting. Note that the conditional density of the daily return is normal so that

$$f\left(R(t)|h(t)\right) = \frac{1}{\sqrt{2\pi h(t)}} \exp\left(-\frac{(R(t) - r - \mu h(t))^2}{2h(t)}\right).$$

The return log likelihood is therefore

$$\ln L^R \propto -\frac{1}{2} \sum_{t=1}^{T} \left\{\ln\left(h(t)\right) + \left(R(t) - r - \mu h(t)\right)^2 / h(t)\right\}. \tag{21}$$

Define the Black-Scholes Vega (BSV) weighted option valuation errors as

$$\varepsilon_i = \left(C_i^{Mkt} - C_i^{Mod}\right) / BSV_i^{Mkt},$$

where C_i^{Mkt} represents the market price of the ith option, C_i^{Mod} represents the model price, and BSV_i^{Mkt} represents the Black-Scholes vega of the option (the derivative with respect to volatility) at the market implied level of volatility. Assume these disturbances are i.i.d. normal so that the option log likelihood is

$$\ln L^O \propto -\frac{1}{2} \sum_{i=1}^{N} \left\{\ln\left(s_\varepsilon^2\right) + \varepsilon_i^2 / s_\varepsilon^2\right\}. \tag{22}$$

where we can concentrate out s_ε^2 using the sample analogue $\hat{s}_\varepsilon^2 = \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i^2$. These vega-weighted option errors are very useful because it can be shown that they are an approximation to implied volatility based errors, which have desirable statistical properties. Unlike implied volatility errors, they do not require Black-Scholes inversion of model prices at every step in the optimization, which is very costly in large scale empirical estimation exercises such as ours. See for instance Carr and Wu (2007) and Trolle and Schwartz (2009) for applications of BSV^{Mkt} weighted option errors.

We can now solve the following optimization problem

$$\max_{\Theta,\Theta^*} \ln L^R + \ln L^O, \tag{23}$$

where $\Theta = \{\omega, \alpha, \beta, \gamma, \mu\}$ denotes the physical parameters and Θ^* denotes the risk-neutral parameters which are mapped from Θ using (11). The riskless rate r in (21) set to 5 percent, and we use the term structure of interest rates from OptionMetrics when pricing options in (22).

To demonstrate the usefulness and implications of the pricing kernel (2), we conduct four

19

different empirical exercises. The first exercise is intended as a benchmark. We maximize the joint log likelihood in (23) with respect to nine parameters: the five physical parameters $\omega, \alpha, \beta, \gamma, \mu$ as well as the four risk-neutral parameters $\omega^*, \alpha^*, \beta^*, \gamma^*$. This exercise has no value from an economic perspective, it is merely a fitting exercise. We refer to it as the ad-hoc model. Because there are no built-in restrictions between the physical and risk-neutral parameters, we effectively fit both the risk-neutral and physical distribution as well as possible, which will serve as a benchmark for three other models which impose economic restrictions between the physical and the risk-neutral distribution.

The first of the three models we consider contains no premia. This amounts to setting $\mu = 0$ and $\xi = 0$, so that we have $h^*(t+1) = h(t+1)$. The resulting model has only four free parameters: ω, α, β, and γ. The second model is the Heston-Nandi (2000) model that allows for equity risk. This effectively leaves the equity premium μ (or equivalently the risk-aversion ϕ) as a free parameter, yielding five parameters in total. The most general economic model contains an equity premium and an independent volatility premium by allowing ξ to be a free parameter.

Table 3 presents the empirical results, which are quite striking.[16] First, consider the likelihood functions at the optimum for the four different models. Notice that at the optimum, the contribution from the options part of the likelihood is equal in all four cases. This indicates that the contribution of the option data to the likelihood is so important that the parameters adjust to fit the option data as in the benchmark ad-hoc case, and by implication sacrificing some goodness of fit in the return component of the likelihood, which is in all cases smaller than in the ad-hoc specification. This occurs despite using a return sample, 1990-2005, which is longer than the option sample, 1996-2004.

The most important feature of the new GARCH model is that it captures the U-shape in the log ratio of the risk-neutral and physical densities found in Figure 2. In order to show this we plot the model-implied log density ratio in Figure 5 for various maturities. The solid lines plot the log density ratio for the new model and the dashed lines show the special Rubinstein (1976) and Brennan (1979) case of a variance premium of zero. We set the conditional variance to its unconditional level and we fix the risk-free rate at 5% per year. We use the parameters from the third column of Table 3. Note that the one-month maturity used in Figure 2 corresponds to the upper-right panel of Figure 5 and note that the empirical U-shapes in Figure 2 are matched remarkably well by Figure 5. The new model thus captures the first stylized fact discussed in Section 3.

To explore other features of the new GARCH model, it is instructive to inspect some of the properties of the ad-hoc model, which does not impose restrictions across the physical and

---

[16]We impose $\omega^* = \omega = 0$ in estimation because the nonnegativity constraint is binding and it is a necessary condition for positive variances.

risk-neutral parameters. The resulting properties are therefore entirely determined by the data, and serve as a useful benchmark for the three other models. We observe four very important features.

First, the rightmost column in Table 3 shows that the average risk-neutral volatility, 0.234, is much higher than the average physical volatility, 0.151. The Second, the risk-neutral variance persistence, 0.985, is higher than the physical variance persistence, 0.962, in Table 3. This is the stylized fact underlying Stein's overreaction regressions.

Third, the average annualized risk-neutral volatility of risk-neutral variance in (13) is 0.087 which is much larger than the physical volatility of physical variance, 0.046. Fourth, (the absolute value of) the average of the risk-neutral leverage correlation defined as

$$\frac{Cov^*_{t-1}(R\left(t\right),h^*(t+1))}{\sqrt{h^*(t)Var^*_{t-1}(h^*(t+1))}} = \frac{-2\alpha^*\gamma^*h^*(t)}{\sqrt{h^*(t)\left(2\alpha^{*2}+4\alpha^{*2}\gamma^{*2}h^*(t)\right)}}$$

annualized and evaluated at $h^*(t) = E\left[h^*(t)\right]$ is -0.977 which exceeds in magnitude its physical counterpart, -0.905. The last two empirical features are of course critical in ensuring that the risk-neutral distribution has fatter tails than the physical distribution.

The model without premia serves as another benchmark. The model's physical properties are identical to its risk-neutral properties. Allowing a nonzero equity premium results in a rather small increase in the total likelihood (from $55, 366.1$ to $55, 368.4$). Asymptotically, twice the difference in likelihood values has a chi-square distribution with one degree of freedom. Therefore this improvement is statistically insignificant, it can be seen that in economic terms the improvements are modest, as the risk-neutral average volatility, variance persistence, and volatility of variance are not very different from their physical counterparts. While these quantitative effects are small, it is reassuring that qualitatively all effects go in the expected direction, as emphasized in Section 2. The risk-neutral volatility exceeds the physical volatility, the risk-neutral variance persistence is higher, and the risk-neutral tails are fatter.

When adding an independent volatility premium to the specification, the likelihood function improves spectacularly (from $55, 368.4$ to $55, 485.5$). Perhaps even more pertinently, some of the model properties for the model with equity and volatility premia are very similar to those of the ad-hoc model, which provides benchmark properties that are completely data-driven. The reported parameter combination $(1 - 2\alpha\xi)^{-1/2}$ shows the ratio of the risk-neutral volatility $\sqrt{h^*(t)}$ to the physical volatility $\sqrt{h(t)}$. The difference between the long-run physical, 0.153, and risk-neutral volatility, 0.232, is almost exactly the same as in the ad-hoc model. The physical volatility of variance, 0.039, is a good approximation of the ad-hoc benchmark as is its risk-neutral counterpart, 0.086. However, while the physical leverage correlation, -0.974, is smaller (in absolute value) than the risk-neutral leverage correlation, -0.977, and even though the physical variance

persistence, 0.983, is smaller than the risk-neutral variance persistence, 0.985, the difference between the two persistence measures is not nearly large enough. The model therefore explains only part of the Stein puzzle. To understand why, note that the model's parsimonious specification forces the same parameters to capture a large number of stylized facts. In maximizing the likelihood, the relative fit of these stylized facts is traded off, and it is well-known that it is difficult to precisely estimate the mean-reversion of highly persistent processes. The procedure will put weight on capturing the second, third and fourth moments, and Nelson and Foster (1994) show that the covariance between stock returns and variance is particularly important in this regard. Adding components to the GARCH model would have increased flexibility to fit these dynamics as in Christoffersen, Jacobs, Ornthanalai and Wang (2008).

Figure 6 further explores the models' performance in matching the physical and risk-neutral volatility of variance. It plots the square root of the risk-neutral $Var^*_{t-1}(h^*(t+1))$ and the physical counterpart $Var_{t-1}(h(t+1))$ over time for the four specifications in Table 3. Note that the internally consistent model with equity and volatility premia in the bottom left panel tracks quite closely the ad-hoc model in the bottom right panel. The risk-neutral volatility of variance exceeds its physical counterpart in both cases. In contrast, in the top two panels the models without volatility premia have virtually identical physical and risk-neutral volatility of variance.

Table 4 explores further the option valuation performance of the new model that allows for equity and volatility premia by providing measures of the implied volatility root mean squared error (IV RMSE) and IV Bias by moneyness and maturity. The first row in each panel of Table 4 is marked "Equity and Volatility Premia" and reports the IV RMSE and Bias corresponding to the parameters estimated in the column labelled "Equity and Volatility Premia" in Table 3. The second row in Table 4 labelled "Equity Premium only" uses the same physical parameters as the first row but forces $\xi = 0$ so that only the equity premium remains. The third row in each panel labelled "No Premia" further forces $\mu = \xi = 0$ thus eliminating both premia but again keeping all other physical parameter values from the first two rows. Note that we are not using the different sets of estimates from Table 3 because they give very similar option fits due to the fact that we have many more option than return observations. We are instead taking the estimates for the new model and then shutting down the premium parameters one by one.

Comparing the first and second rows in Table 4 shows that the volatility premium offers a dramatic improvement in option fit–both in terms of IV RMSE and Bias. When comparing the second and third rows we see that the equity premium plays a much smaller role in improving option fit when comparing with a model with no premia. The parameterizations with no volatility premia imply a strong positive bias implying that on average the models underprice options when the volatility premium is excluded. This bias is virtually eliminated–and the RMSE is radically improved–when the volatility premium is incorporated into the model.

5   Conclusion

We develop a new GARCH option model for the purpose of index option valuation by specifying a more general pricing kernel. Unlike the traditional Black-Scholes (1973) and Rubinstein (1976) pricing kernel, which is a function of the index return only, we specify that the pricing kernel is also a function of the return variance. Although the pricing kernel is specified as monotonic in the index return, the projection of the pricing kernel onto returns is U-shaped. With a negative variance premium, this model feature is consistent with semi-parametric evidence from returns and options which reveals that the conditional pricing kernel is U-shaped in returns, and is relatively stable over time.

The new model generalizes the Heston-Nandi (2000) model by allowing the risk-neutral variance, persistence, and volatility of variance to differ from their physical counterparts. We demonstrate that the model can qualitatively account for a number of important puzzles in the option pricing literature. In order to demonstrate that the more general pricing kernel can reconcile the return distributions implicit in the time series of returns and option prices, we implement the model by maximizing the sum of the return likelihood and a likelihood based on successive cross-sections of option prices. We benchmark the model's performance to an ad-hoc model that does not impose restrictions across the physical and risk-neutral parameters. We find that the fit of the model with the new pricing kernel is dramatically better than the fit resulting from the traditional pricing kernel, and that two important differences between physical and risk-neutral moments are very similar to the differences obtained for the ad-hoc model.

The results and empirical exercises in this paper can be extended and generalized in a number of ways. First, while different models may be needed, alternative loss functions may emphasize different moments and therefore yield different results. Alternatively, it may prove interesting to investigate the implications of more general pricing kernels for option valuation in the presence of richer return dynamics, for instance with jumps in returns and/or volatility (Christoffersen, Heston, and Jacobs, 2006), multiple volatility components (Christoffersen, Jacobs, Ornthanalai and Wang, 2008), or economic drivers of uncertainty (David and Veronesi, 2011). Our results indicate that it is critical to evaluate such models using a loss function similar to the one in this paper, with a return component as well as an option component. Two questions arise in this regard: on the one hand, whether these models can improve option fit; on the other hand, whether they can describe a richer link between the physical and risk-neutral distributions with intuitively plausible prices of risk. Models that can reconcile the differences between physical and risk-neutral persistence, as well as physical and risk-neutral leverage correlation, would be of particular interest.

Appendix

Appendix A. Risk Neutralization of the Heston (1993) SV Model

We show that the physical SV process in (1) and the risk neutral SV process in (3) are linked by the pricing kernel in (2) by imposing the condition that the product of any traded asset and the pricing operator is a martingale under the physical probability measure.

Let $B(t)$ be the risk-free bond $\left(\frac{dB(t)}{B(t)} = rdt\right)$ and let $U(t)$ be an asset that depends on the spot price, $S(t)$, and the volatility, $v(t)$. From (2) we get the dynamic of $M(t)$

$$\log(M(t)) = \log(M(0)) + \phi \log(S(t)) - \phi \log(S(0)) + \delta t + \eta \int_0^t v(s)ds + \xi \left[v(t) - v(0)\right],$$

This gives

$$
\begin{aligned}
d\log(M) &= \phi d\log(S) + \delta dt + \eta v dt + \xi dv \\
&= \left[\phi\left(r + \mu v - \frac{1}{2}v\right) + \delta + \eta v + \xi\kappa(\theta - v)\right] dt + \\
&\quad \left[\phi\sqrt{v} + \xi\sigma\rho\sqrt{v}\right] dz_1 + \left[\xi\sigma\sqrt{1-\rho^2}\sqrt{v}\right] dz_2,
\end{aligned}
$$

where we use the fact that $d\log(S) = \left(r + \mu v - \frac{1}{2}v\right) dt + \sqrt{v}dz_1$, which is obtained by applying Itô's lemma to equation (1). Note that we have suppressed dependence on t in the notation. Again by Itô's lemma, we have

$$
\begin{aligned}
\frac{dM}{M} &= \left[\phi\left(r + \mu v - \frac{1}{2}v\right) + \delta + \eta v + \xi\kappa(\theta - v) + \frac{1}{2}\phi^2 v + \phi\xi\sigma\rho v + \frac{1}{2}\xi^2\sigma^2 v\right] dt \\
&\quad + \left[\phi\sqrt{v} + \xi\sigma\rho\sqrt{v}\right] dz_1 + \left[\xi\sigma\sqrt{1-\rho^2}\sqrt{v}\right] dz_2.
\end{aligned}
$$

From the condition that $B(t)M(t)$ is a martingale, i.e. its drift is equal to zero, we deduce the restrictions on δ and η.

$$\phi\left(r + \mu v - \frac{1}{2}v\right) + \delta + \eta v + \xi\kappa(\theta - v) + \frac{1}{2}\phi^2 v + \phi\xi\sigma\rho v + \frac{1}{2}\xi^2\sigma^2 v = -r.$$

This must hold for $v = 0$ and for $v = +\infty$, giving the restrictions on δ and η

$$
\begin{aligned}
\delta &= -(1 + \phi)r - \xi\kappa\theta. \\
\eta &= -\phi\mu + \frac{1}{2}\phi + \xi\kappa - \frac{1}{2}\left(\phi^2 + 2\phi\xi\sigma\rho + \xi^2\sigma^2\right).
\end{aligned}
$$

Similarly, from the condition that $S(t)M(t)$ is a martingale, i.e. its drift is equal to zero, we

deduce a restriction on ϕ. Using the fact that the drift of M is equal to $-rMdt$, we have

$$\mu v + (\phi + \xi\sigma\rho)\, v = 0,$$

or equivalently $\phi = -\mu - \xi\sigma\rho$.

Finally, by equating the drift of $U(t)M(t)$ to zero, we deduce the restriction on ξ

$$0 = \left[M\left(U_t + rSU_S + \mu vSU_S + \kappa(\theta - v)U_v + \frac{1}{2}vS^2U_{SS} + vS\sigma\rho U_{Sv} \right.\right. \tag{A1}$$
$$\left.\left. + \frac{1}{2}v\sigma^2 U_{vv}\right) - rUM + Mv(\phi + \xi\sigma\rho)(SU_S + \sigma\rho U_v) + M\xi v\sigma^2(1 - \rho^2)U_v\right]dt.$$

Since $\frac{U(t)}{B(t)}$ is a martingale under the risk-neutral measure, we can also show that

$$\lambda vU_v + rU = U_t + rSU_S + \kappa(\theta - v)U_v + \frac{1}{2}vS^2U_{SS} + vS\sigma\rho U_{Sv} + \frac{1}{2}v\sigma^2 U_{vv}. \tag{A2}$$

Substituting (A2) into (A1) we get for the drift D^{UM}

$$D^{UM} = \left[M\left(\lambda vU_v + rU + \mu vSU_S\right) - rUM\right.$$
$$\left. + Mv(\phi + \xi\sigma\rho)(SU_S + \sigma\rho U_v) + M\xi v\sigma^2(1 - \rho^2)U_v\right]dt$$
$$= 0.$$

After simplification and using the fact that $\phi + \xi\sigma\rho = -\mu$, we get

$$\lambda vU_v - \mu v\sigma\rho U_v + \xi v\sigma^2(1 - \rho^2)U_v = 0,$$

and therefore

$$\xi = \frac{\mu\sigma\rho - \lambda}{\sigma^2(1 - \rho^2)}. \tag{A3}$$

Using (A3) we also obtain

$$\phi = \frac{-\mu + \lambda\sigma^{-1}\rho}{(1 - \rho^2)}.$$

Appendix B. Proof of Proposition 1

The discrete-time form of the pricing operator (2) is

$$M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^\phi \exp\left(\delta t + \eta \sum_{s=1}^{t} h(s) + \xi\left(h(t+1) - h(1)\right)\right), \tag{A4}$$

and therefore

$$\frac{M(t)}{M(t-1)} = \left(\frac{S(t)}{S(t-1)}\right)^{\phi} \exp\left(\delta + \eta h(t) + \xi\left(h(t+1) - h(t)\right)\right). \tag{A5}$$

The summations in (A4) are equivalent to the integrals in the continuous-time form (2) under the standard GARCH convention that variance is constant throughout the day, and changes discretely overnight. We shall show that the pricing kernel in (A4) is consistent with the Heston-Nandi GARCH dynamic (5) for the following parameter mapping

$$\begin{aligned}
\delta &= -\left(\phi + 1\right) r - \xi\omega + \tfrac{1}{2}\ln\left(1 - 2\xi\alpha\right), \\
\eta &= -\left(\mu - \tfrac{1}{2}\right)\phi - \xi\alpha\gamma^2 + (1 - \beta)\xi - \frac{\left(\phi - 2\xi\alpha\gamma\right)^2}{2\left(1 - 2\xi\alpha\right)}, \\
\phi &= -(\mu - \tfrac{1}{2} + \gamma)(1 - 2\alpha\xi) + \gamma - \tfrac{1}{2}, \\
\xi &= \frac{\mu\sigma\rho - \lambda}{\sigma^2(1 - \rho^2)}.
\end{aligned}$$

From the GARCH dynamic (5) we can write

$$\begin{aligned}
\frac{S(t)}{S(t-1)} &= \exp\left(r + \left(\mu - \tfrac{1}{2}\right) h(t) + \sqrt{h(t)}z(t)\right), \\
h(t+1) - h(t) &= \omega + (\beta - 1) h(t) + \alpha\left(z(t) - \gamma\sqrt{h(t)}\right)^2.
\end{aligned}$$

Substituting these into (A5) gives

$$\frac{M(t)}{M(t-1)} = \exp\left(\phi r + \phi\left(\mu - \tfrac{1}{2}\right)h + \phi\sqrt{h}z + \delta + \eta h + \xi\omega + \xi(\beta - 1)h + \xi\alpha\left(z - \gamma\sqrt{h}\right)^2\right)$$

where we have dropped the time subscripts for z and h. Expanding the square and collecting terms gives

$$\frac{M(t)}{M(t-1)} = \exp\left(\phi r + \delta + \xi\omega + \left[\phi\left(\mu - \tfrac{1}{2}\right) + \eta + \xi(\beta - 1) + \xi\alpha\gamma^2\right]h + \left[\phi - 2\xi\alpha\gamma\right]\sqrt{h}z + \left[\xi\alpha\right]z^2\right).$$

First, we use the fact that for any initial value $h(t)$, the parameters must be consistent with the Euler equation for the riskless asset.

$$E_{t-1}\left[\frac{M(t)}{M(t-1)}\right] = \exp\left(-r\right). \tag{A6}$$

Note that

$$E_{t-1}\left[\frac{M\left(t\right)}{M\left(t-1\right)}\right] = \exp\left(\phi r + \delta + \xi\omega + \left[\phi\left(\mu - \tfrac{1}{2}\right) + \eta + \xi\left(\beta - 1\right) + \xi\alpha\gamma^2\right]h\right)$$
$$*E\left(\exp\left(\left[\phi - 2\xi\alpha\gamma\right]\sqrt{h}z + \left[\xi\alpha\right]z^2\right)\right).$$

We need the following result

$$E\left[\exp\left(az^2 + 2abz\right)\right] = \exp\left(-\tfrac{1}{2}\ln\left(1 - 2a\right) + \frac{2a^2b^2}{1 - 2a}\right).$$

For our application we have

$$a = \xi\alpha,$$
$$b = \left(\frac{\phi - 2\xi\alpha\gamma}{2\xi\alpha}\right)\sqrt{h},$$

and thus

$$2a^2b^2 = 2\xi^2\alpha^2\left(\frac{\phi - 2\xi\alpha\gamma}{2\xi\alpha}\right)^2 h = \tfrac{1}{2}\left(\phi - 2\xi\alpha\gamma\right)^2 h.$$

Therefore

$$E\left(\exp\left(\left[\phi - 2\xi\alpha\gamma\right]\sqrt{h}z + \xi\alpha z^2\right)\right) = \exp\left(-\tfrac{1}{2}\ln\left(1 - 2\xi\alpha\right) + \frac{\left(\phi - 2\xi\alpha\gamma\right)^2}{2\left(1 - 2\xi\alpha\right)}h\right),$$

and

$$E_{t-1}\left[\frac{M\left(t\right)}{M\left(t-1\right)}\right] = \exp\left(\begin{array}{c}\phi r + \delta + \xi\omega + \left[\phi\left(\mu - \tfrac{1}{2}\right) + \eta + \xi\left(\beta - 1\right) + \xi\alpha\gamma^2\right]h \\ -\tfrac{1}{2}\ln\left(1 - 2\xi\alpha\right) + \frac{\left(\phi - 2\xi\alpha\gamma\right)^2}{2\left(1 - 2\xi\alpha\right)}h\end{array}\right).$$

Rearranging and using (A6) we get

$$\left(\phi + 1\right)r + \delta + \xi\omega - \tfrac{1}{2}\ln\left(1 - 2\xi\alpha\right) + \left[\phi\left(\mu - \tfrac{1}{2}\right) + \eta + \xi\left(\beta - 1\right) + \xi\alpha\gamma^2 + \frac{\left(\phi - 2\xi\alpha\gamma\right)^2}{2\left(1 - 2\xi\alpha\right)}\right]h = 0.$$

Therefore we must have

$$\delta = -\left(\phi + 1\right)r - \xi\omega + \tfrac{1}{2}\ln\left(1 - 2\xi\alpha\right),$$
$$\eta = -\left(\mu - \tfrac{1}{2}\right)\phi - \xi\alpha\gamma^2 + \left(1 - \beta\right)\xi - \frac{\left(\phi - 2\xi\alpha\gamma\right)^2}{2\left(1 - 2\xi\alpha\right)}.$$

Now we use the Euler equation for the underlying index

$$E_{t-1}\left[\frac{S(t)}{S(t-1)}\frac{M(t)}{M(t-1)}\right]=1.$$

First, note that $\frac{S(t)}{S(t-1)}\frac{M(t)}{M(t-1)}$ is equal to $\frac{M(t)}{M(t-1)}$ in (A5) with ϕ is replaced by $\phi+1$, thus we can use the expression for $E_{t-1}\left[\frac{M(t)}{M(t-1)}\right]$ to write

$$E_{t-1}\left[\frac{S(t)}{S(t-1)}\frac{M(t)}{M(t-1)}\right]=\exp\left(\begin{array}{c}(\phi+1)\,r+\delta+\xi\omega+\left[(\phi+1)\left(\mu-\frac{1}{2}\right)+\eta+\xi\,(\beta-1)+\xi\alpha\gamma^2\right]h\\-\frac{1}{2}\ln\left(1-2\xi\alpha\right)+\frac{(\phi+1-2\xi\alpha\gamma)^2}{2(1-2\xi\alpha)}h\end{array}\right).$$

Taking logs, setting equal to zero and using the above solutions for δ and η gives

$$\left[\mu-\frac{1}{2}+\frac{1+2\phi-4\xi\alpha\gamma}{2\left(1-2\xi\alpha\right)}\right]h=0.$$

Solving this for ϕ yields

$$\phi=-(\mu-\frac{1}{2}+\gamma)(1-2\alpha\xi)+\gamma-\frac{1}{2}.$$

To find the risk-neutral dynamic, note that the risk-neutral density is proportional to the physical density times the pricing kernel

$$f_{t-1}^*\left(S(t)\right)=\frac{f_{t-1}(S(t))M(t)}{E_{t-1}(M(t))}$$

Tedious integration shows that $z(t)$ is normally distributed under the risk-neutral measure, but with a different mean and variance. This is a direct implication of the form of the pricing kernel. It is therefore convenient to define a standardized risk-neutral innovation

$$z^*(t)=\sqrt{1-2\alpha\xi}\left(z(t)+\left(\mu+\frac{\alpha\xi}{1-2\alpha\xi}\right)\sqrt{h(t)}\right). \tag{A7}$$

The risk-neutral dynamics in equation (10) can then be derived by substituting the risk-neutral innovation (A7) into the physical GARCH process in equation (5).

Appendix C. Proof of Corollary 2

We want to show that the logarithm of the pricing kernel takes the quadratic form

$$
\ln\left(\frac{M\left(t\right)}{M\left(t-1\right)}\right) = \frac{\xi\alpha}{h\left(t\right)}\left(R(t)-r\right)^2 - \mu\left(R(t)-r\right) +
$$
$$
\left(\eta + \xi\left(\beta-1\right) + \xi\alpha\left(\mu - \tfrac{1}{2} + \gamma\right)^2\right)h\left(t\right) + \delta + \xi\omega + \phi r
$$

where $R(t) = \ln(S\left(t\right)/S(t-1))$. First recall that

$$
\frac{M\left(t\right)}{M\left(t-1\right)} = \left(\frac{S\left(t\right)}{S\left(t-1\right)}\right)^\phi \exp\left(\delta + \eta h\left(t\right) + \xi\left(h\left(t+1\right) - h\left(t\right)\right)\right), \qquad\text{(A8)}
$$

and that

$$
h\left(t+1\right) - h\left(t\right) = \omega + \left(\beta-1\right)h\left(t\right) + \alpha\left(z\left(t\right) - \gamma\sqrt{h\left(t\right)}\right)^2.
$$

We also have

$$
R(t) = r + \left(\mu - \tfrac{1}{2}\right)h\left(t\right) + \sqrt{h\left(t\right)}z\left(t\right)
$$

so that

$$
z\left(t\right) = \frac{R(t) - r - \left(\mu - \tfrac{1}{2}\right)h\left(t\right)}{\sqrt{h\left(t\right)}}.
$$

From this we get

$$
\begin{aligned}
\ln\left(\frac{M\left(t\right)}{M\left(t-1\right)}\right) &= \phi R(t) + \delta + \eta h\left(t\right) + \xi\left(h\left(t+1\right) - h\left(t\right)\right)\\
&= \phi R(t) + \delta + \xi\omega + \left(\eta + \xi\left(\beta-1\right)\right)h\left(t\right) + \xi\alpha\left(z\left(t\right) - \gamma\sqrt{h\left(t\right)}\right)^2\\
&= \phi R(t) + \delta + \xi\omega + \left(\eta + \xi\left(\beta-1\right)\right)h\left(t\right) + \frac{\xi\alpha}{h\left(t\right)}\left(R(t) - r - \left(\mu - \tfrac{1}{2} + \gamma\right)h\left(t\right)\right)^2
\end{aligned}
$$

expanding the square and collecting terms yields

$$
\ln\left(\frac{M\left(t\right)}{M\left(t-1\right)}\right) = \frac{\xi\alpha}{h\left(t\right)}\left(R(t)-r\right)^2 + \left(\phi - 2\xi\alpha\left(\mu - \tfrac{1}{2} + \gamma\right)\right)\left(R(t)-r\right)
$$
$$
+ \left(\eta + \xi\left(\beta-1\right) + \xi\alpha\left(\mu - \tfrac{1}{2} + \gamma\right)^2\right)h\left(t\right) + \delta + \xi\omega + \phi r
$$

From the equation for ϕ we have

$$
\phi - 2\alpha\xi(\mu - \tfrac{1}{2} + \gamma) = -\mu
$$

and the result obtains.

References

[1] Ait-Sahalia, Y., and A. Lo (1998), Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, Journal of Finance, 53, 499-547.

[2] Ait-Sahalia, Y., and A. Lo (2000), Nonparametric Risk Management and Implied Risk Aversion, Journal of Econometrics, 94, 9-51.

[3] Andersen, T.G., L. Benzoni, and J. Lund (2002), Estimating Jump-Diffusions for Equity Returns, Journal of Finance, 57, 1239-1284.

[4] Bakshi, G., C. Cao, and Z. Chen (1997), Empirical Performance of Alternative Option Pricing Models, Journal of Finance, 52, 2003-2049.

[5] Bakshi, G., and N. Kapadia (2003), Delta Hedged Gains and the Negative Market Volatility Risk Premium, Review of Financial Studies, 16, 527-566.

[6] Bakshi, G. D. Madan, and G. Panayotov (2010), Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels, Journal of Financial Economics, 97, 130-154.

[7] Banz, R. and M. Miller (1978), Prices for State-Contingent Claims: Some Estimates and Applications, Journal of Business, 51, 653-672.

[8] Barone-Adesi, G., R. Engle, and L. Mancini (2008), A GARCH Option Pricing Model with Filtered Historical Simulation, The Review of Financial Studies, 21, 1223-1258.

[9] Bates, D. (1996a), Testing Option Pricing Models, in Handbook of Statistics, Statistical Methods in Finance, G.S. Maddala and C.R. Rao (eds.), 567-611. Amsterdam: Elsevier.

[10] Bates, D. (1996b), Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options, Review of Financial Studies, 9, 69-107.

[11] Bates, D. (2000), Post-'87 Crash Fears in the S&P 500 Futures Option Market, Journal of Econometrics, 94, 181-238.

[12] Bates, D. (2003), Empirical Option Pricing: A Retrospection, Journal of Econometrics, 116, 387-404.

[13] Bates, D. (2006), Maximum Likelihood Estimation of Latent Affine Processes, Review of Financial Studies, 19, 909-965.

[14] Bates, D. (2008), The Market for Crash Risk, Journal of Economic Dynamics and Control, 32, 2291-2321.

[15] Bates, D. (2010), U.S. Stock Market Crash Risk. 1926-2009, Working Paper, University of Iowa.

[16] Benzoni, L., P. Collin-Dufresne, and R. Goldstein (2006), Can Standard Preferences Explain the Prices of Out-of-the-Money S&P 500 Put Options? Working Paper, Columbia University.

[17] Black, F., and M. Scholes (1973), The Pricing of Options and Corporate Liabilities, Journal of Political Economy, 81, 637-659.

[18] Blair, B., S.-H. Poon, and S. Taylor (2001), Forecasting S&P 500 Volatility: The Incremental Information Content of Implied Volatilities and High-Frequency Index Returns, Journal of Econometrics, 105, 5–26.

[19] Bliss, R. and N. Panigirtzoglou (2004), Option-Implied Risk Aversion Estimates, Journal of Finance, 59, 407-446.

[20] Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.

[21] Bollerslev, T. (1987), A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, Review of Economics and Statistics, 69, 542-547.

[22] Bollerslev, T., and H. Mikkelsen (1996), Modeling and Pricing Long Memory in Stock Market Volatility, Journal of Econometrics, 73, 151-184.

[23] Bollerslev, T., G. Tauchen, and H. Zhou (2009), Expected Stock Return and Variance Risk Premia, Review of Financial Studies, 22, 4463-4492.

[24] Bondarenko, O. (2003), Why Are Put Options So Expensive? Working Paper, University of Illinois at Chicago.

[25] Breeden, D. and R. Litzenberger (1978), Prices of State-Contingent Claims Implicit in Option Prices, Journal of Business, 51, 621-651.

[26] Brennan, M. (1979), The Pricing of Contingent Claims in Discrete-Time Models, Journal of Finance, 34, 53-68.

[27] Broadie, M., M. Chernov, and M. Johannes (2007), Model Specification and Risk Premiums: Evidence from Futures Options, Journal of Finance, 62, 1453-1490.

[28] Brown, D., and J. Jackwerth (2001), The Pricing Kernel Puzzle: Reconciling Index Option Data and Economic Theory, Working Paper, University of Wisconsin.

[29] Canina, L., and S. Figlewski (1993), The Informational Content of Implied Volatility, Review of Financial Studies, 6, 659-681.

[30] Carr, P. and L. Wu (2004), Time-Changed Levy Processes and Option Pricing, Journal of Financial Economics, 17, 113-141.

[31] Carr, P., and L. Wu (2007), Stochastic Skew in Currency Options, Journal of Financial Economics, 86, 213-247.

[32] Chabi-Yo, F. (2009), Pricing Kernels with Stochastic Skewness and Volatility Risk, Working Paper, Ohio State University.

[33] Chernov, M. (2003), Empirical Reverse Engineering of the Pricing Kernel, Journal of Econometrics, 116, 329-364.

[34] Chernov, M. (2007), On the Role of Risk Premia in Volatility Forecasting, Journal of Business and Economic Statistics, 25, 411-426.

[35] Chernov, M. and E. Ghysels (2000), A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Option Valuation, Journal of Financial Economics, 56, 407-458.

[36] Christoffersen, P., S. Heston, and K. Jacobs (2006), Option Valuation with Conditional Skewness, Journal of Econometrics, 131, 253-284.

[37] Christoffersen, P., S. Heston, and K. Jacobs (2009), The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well, Management Science, 55, 1914-1932.

[38] Christoffersen, P., and K. Jacobs (2004), Which GARCH Model for Option Valuation? Management Science, 50, 1204-1221.

[39] Christoffersen, P., K. Jacobs, C. Ornthanalai, and Y. Wang (2008), Option Valuation with Long-Run and Short-Run Volatility Components, Journal of Financial Economics, 90, 272-297.

[40] Coval, J. and T. Shumway (2001), Expected Option Returns, Journal of Finance, 56, 983-1009.

[41] Cox, J., J. Ingersoll, and S. Ross (1985), An Intertemporal General Equilibrium Model of Asset Prices, Econometrica, 53, 363-384.

[42] David, A., and P. Veronesi (2011), Investor and Central Bank Uncertainty and Fear Measures Embedded in Index Options, Working Paper, University of Chicago.

[43] Day, T. and C. Lewis (1992), Stock Market Volatility and the Information Content of Stock Index Options, Journal of Econometrics, 52, 267-287.

[44] Driessen, J. and P. Maenhout (2007), An Empirical Portfolio Perspective on Option Pricing Anomalies, Review of Finance, 11, 561-603.

[45] Duan, J. (1995), The GARCH option pricing model. Mathematical Finance 5, 13-32.

[46] Duan, J., and J.G. Simonato (2001), American Option Pricing under GARCH by a Markov Chain Approximation, Journal of Economic Dynamics and Control, 25, 1689-1718.

[47] Dumas, B., J. Fleming, and R. Whaley (1998), Implied Volatility Functions: Empirical Tests, Journal of Finance, 53, 2059-2106.

[48] Engle, R. (1982), Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation, Econometrica, 50, 987-1008.

[49] Engle, R., and G. Lee (1999), A Permanent and Transitory Component Model of Stock Return Volatility. In: Engle, R., White, H. (Eds.), Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger. Oxford University Press, New York, pp. 475-497.

[50] Eraker, B. (2004), Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, Journal of Finance, 59, 1367-1403.

[51] Eraker, B., M. Johannes, and N. Polson (2003), The Role of Jumps in Returns and Volatility, Journal of Finance, 58, 1269-1300.

[52] Fleming, J. (1998), The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices, Journal of Empirical Finance, 5, 317-345.

[53] Fleming, J., and C. Kirby (2003), A Closer Look at the Relation Between GARCH and Stochastic Autoregressive Volatility, Journal of Financial Econometrics, 1, 365-419.

[54] Heston, S. (1993), A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, Review of Financial Studies, 6, 327-343.

[55] Heston, S. and S. Nandi (2000), A Closed-Form GARCH Option Pricing Model, Review of Financial Studies, 13, 585-626.

[56] Huang, J.-Z. and L. Wu (2004), Specification Analysis of Option Pricing Models Based on Time-Changed Levy Processes, Journal of Finance, 59, 1405–1439.

[57] Hull, J., and A. White (1987), The Pricing of Options with Stochastic Volatilities, Journal of Finance, 42, 281-300.

[58] Jackwerth, J. (2000), Recovering Risk Aversion from Option Prices and Realized Returns, Review of Financial Studies, 13, 433-451.

[59] Jones, C. (2003), The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, Journal of Econometrics, 116, 181-224.

[60] Jorion, P. (1995), Predicting Volatility in the Foreign Exchange Market, Journal of Finance, 50, 507–528.

[61] Lamoureux, C., and W. Lastrapes (1993), Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities, Review of Financial Studies, 6, 293-326.

[62] Lamoureux, C., and A. Paseka (2009), Information in Option Prices and the Underlying Asset Dynamics, Working Paper, University of Arizona.

[63] Liu, J., J. Pan, and T. Wang (2004), An Equilibrium Model of Rare Event Premia, Review of Financial Studies, 18, 131-164.

[64] Melino, A., and S. Turnbull (1990), Pricing Foreign Currency Options with Stochastic Volatility, Journal of Econometrics, 45, 239-265.

[65] Nelson, D. (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, Econometrica, 59, 347-370.

[66] Nelson, D. (1992), Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model, Journal of Econometrics, 52, 61-90.

[67] Nelson, D., (1996), Asymptotically Optimal Smoothing with ARCH Models, Econometrica, 64, 561-573.

[68] Nelson, D., and D. Foster (1994), Asymptotic Filtering Theory for Univariate ARCH Models, Econometrica, 62, 1-41.

[69] Pan, J. (2002), The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study, Journal of Financial Economics, 63, 3-50.

[70] Poteshman, A. (2001), Underreaction, Overreaction, and Increasing Misreaction to Information in the Options Market, Journal of Finance, 56, 851-876.

[71] Ritchken, P., R. Trevor (1999), Pricing Options under Generalized GARCH and Stochastic Volatility Processes, Journal of Finance, 54, 377-402.

[72] Rompolis, L. and E. Tzavalis (2008), Recovering Risk Neutral Densities from Option Prices: A New Approach, Journal of Financial and Quantitative Analysis, 43, 1037-1054.

[73] Rosenberg, J. and R. Engle (2002), Empirical Pricing Kernels, Journal of Financial Economics 64, 341-372.

[74] Rubinstein, M. (1976), The Valuation of Uncertain Income Streams and the Pricing of Options, Bell Journal of Economics, 7, 407-425.

[75] Rubinstein, M. (1994), Implied Binomial Trees, Journal of Finance, 49, 771-818.

[76] Santa-Clara, P., and S. Yan (2010), Crashes, Volatility, and the Equity Premium, Review of Economics and Statistics, 92, 435-451.

[77] Shive, S., and T. Shumway (2006), Is the Pricing Kernel Monotonic? Working Paper, Notre Dame University.

[78] Stein, J. (1989), Overreactions in the Options Markets, Journal of Finance, 44, 1011-1023.

[79] Trolle, A., and E. Schwartz (2009), Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives, Review of Financial Studies, 22, 4423-4461.

[80] Wiggins, J. (1987), Option Values Under Stochastic Volatility: Theory and Empirical Evidence, Journal of Financial Economics, 19, 351-372.

Figure 1: One-Month Physical Conditional Histograms and Risk-Neutral Conditional Densities
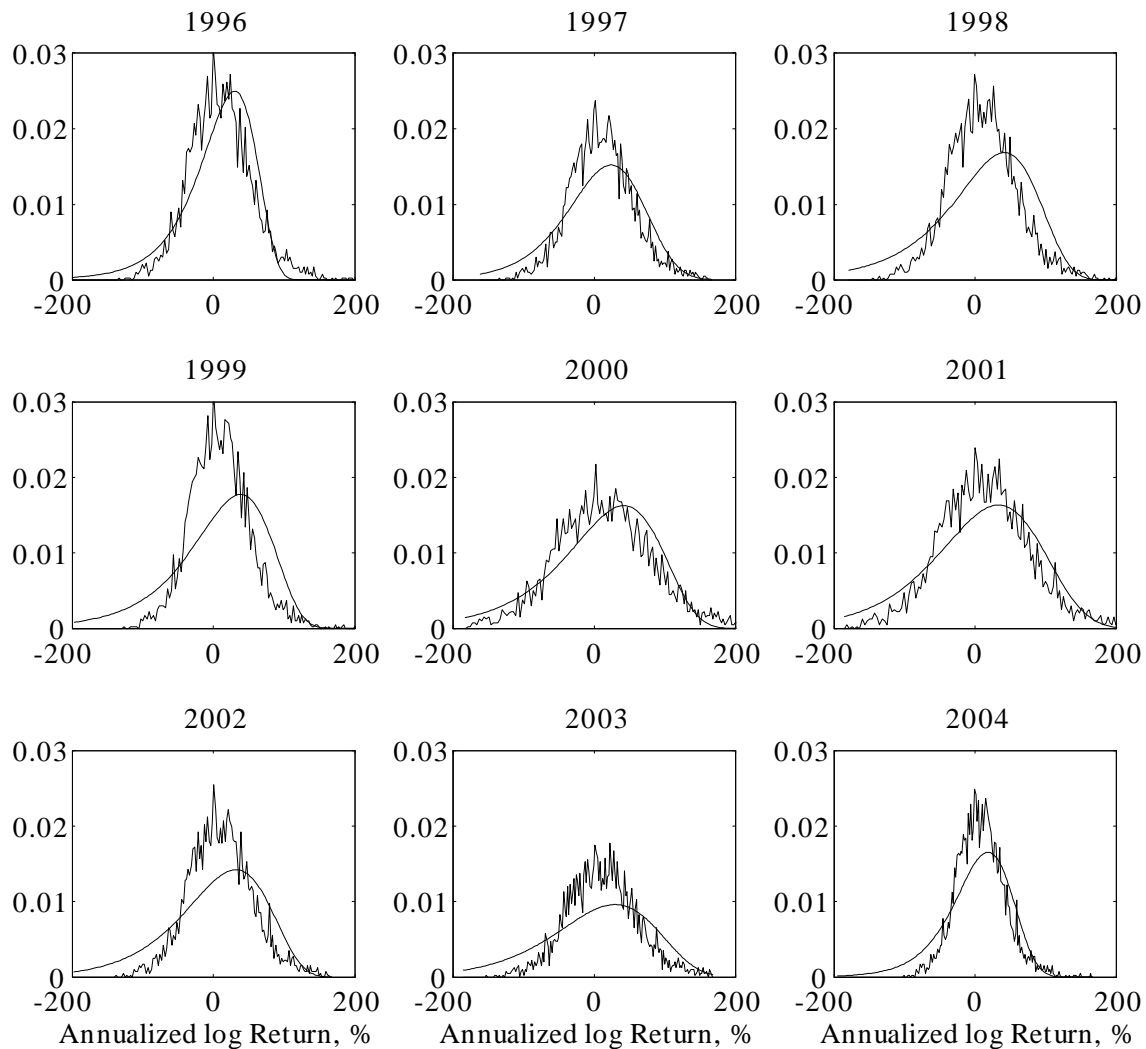

Notes to Figure: We plot physical conditional histograms and risk-neutral conditional densities. The risk-neutral conditional density is obtained using the first Wednesday of option data for each of the nine years in our sample. For each Wednesday, a polynomial is fitted to the data, and the density is obtained from the polynomial fit using the method of Breeden and Litzenberger (1978). To construct the physical conditional histogram, we use return innovations for 1990-2004, scale the time series by the Wednesday's conditional volatility, and draw the histogram. On the horizontal axis are annualized log returns in percent.

Figure 2: Log Ratios of Risk-Neutral One-Month Densities and Physical One-Month Histogram
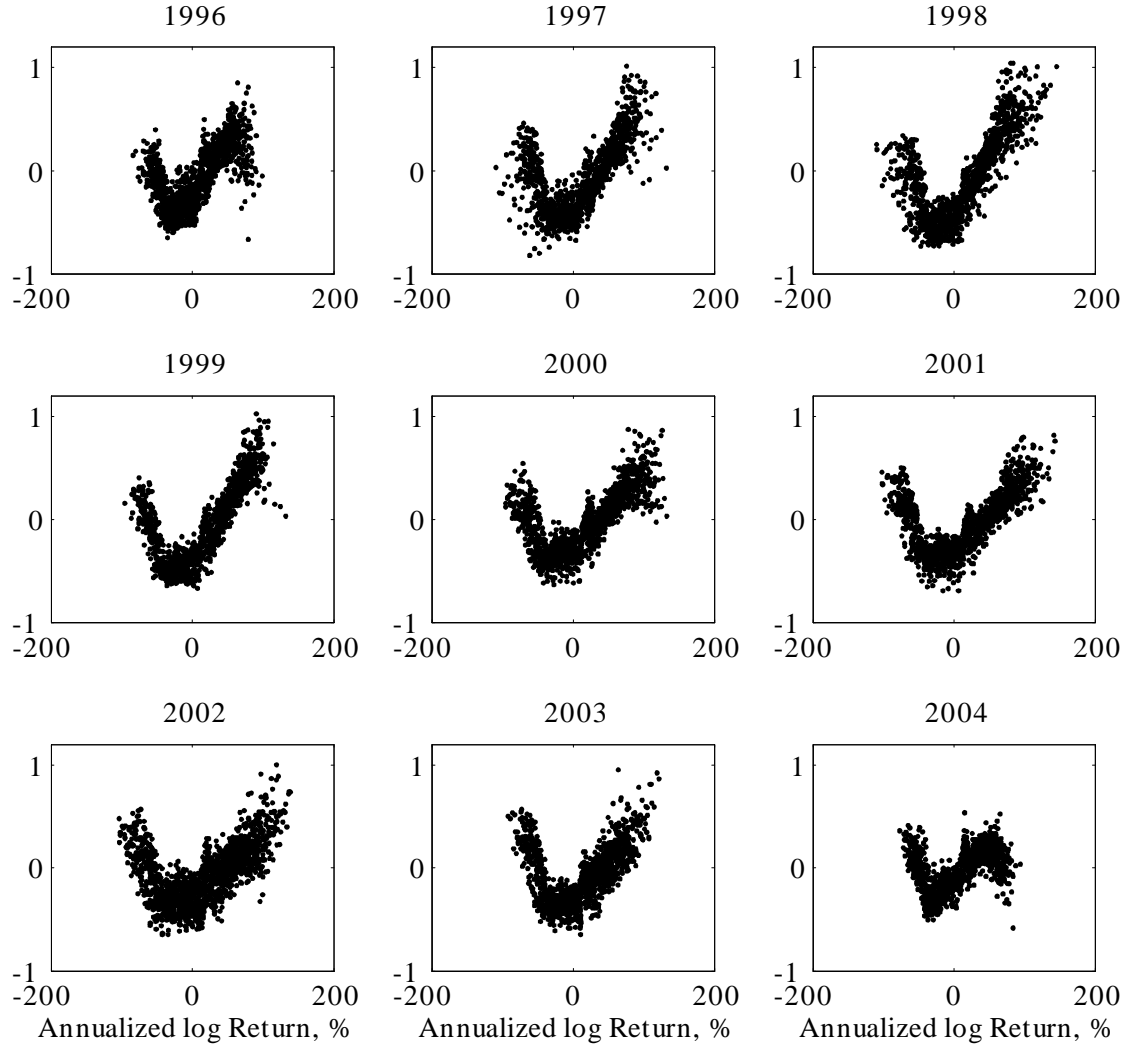

Notes to Figure: We plot the natural logarithm of the ratio of the risk-neutral conditional densities and the physical conditional histogram. For each year in the option sample, we plot the ratios for each of the Wednesdays in that year. On the horizontal axis are annualized log returns in percent.

Figure 3: Physical and Risk-Neutral Volatility on the S&P500. 1996-2004.
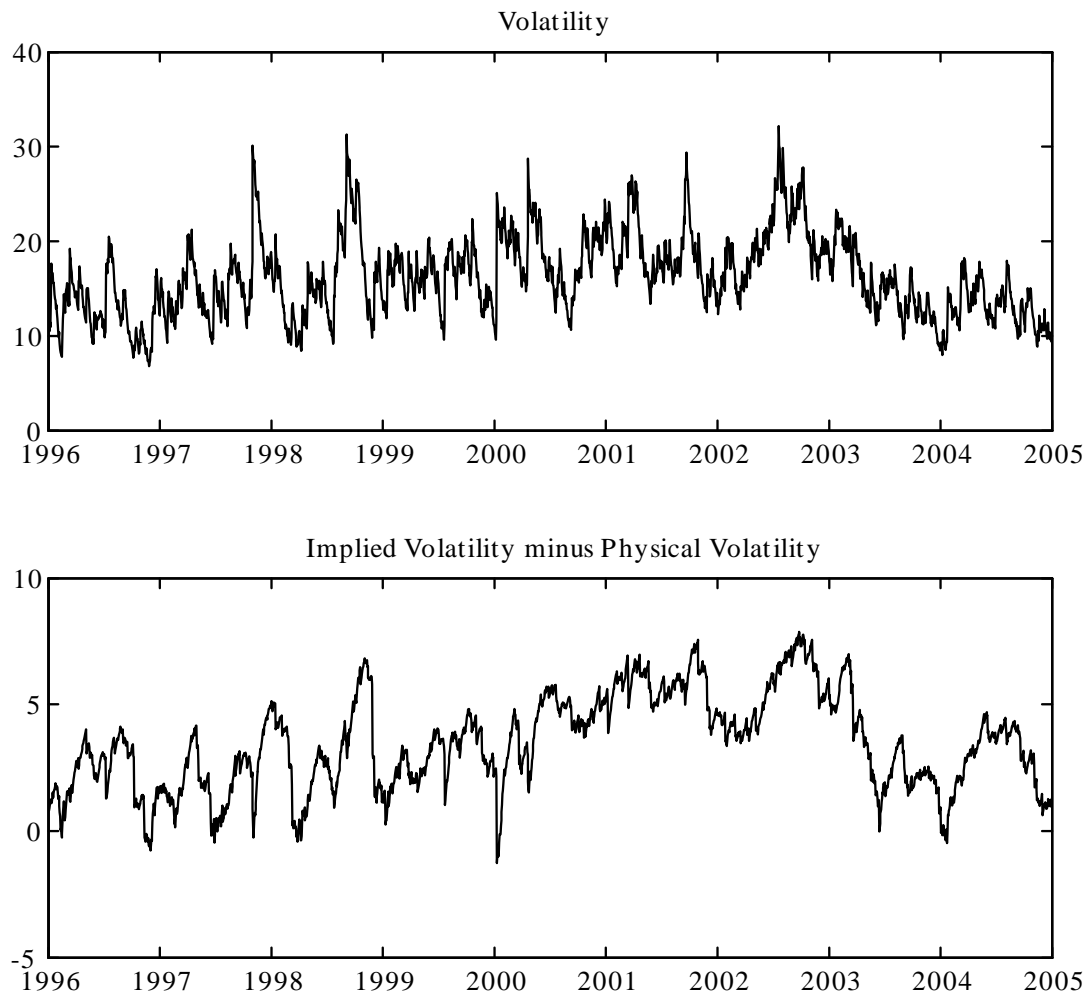
Notes to Figure: The top panel plots the annualized standard deviation of the daily S&P500 return from 1996 through 2004. This volatility is filtered through a GARCH process and is a physical volatility. The bottom panel plots the daily difference between the option-implied (risk-neutral) and physical volatilities.

Figure 4: Returns on Short Straddles. 1996-2004.

Monthly Index of Short Straddle Strategy and S&P500

Monthly Percent Return on Short Straddle Strategy
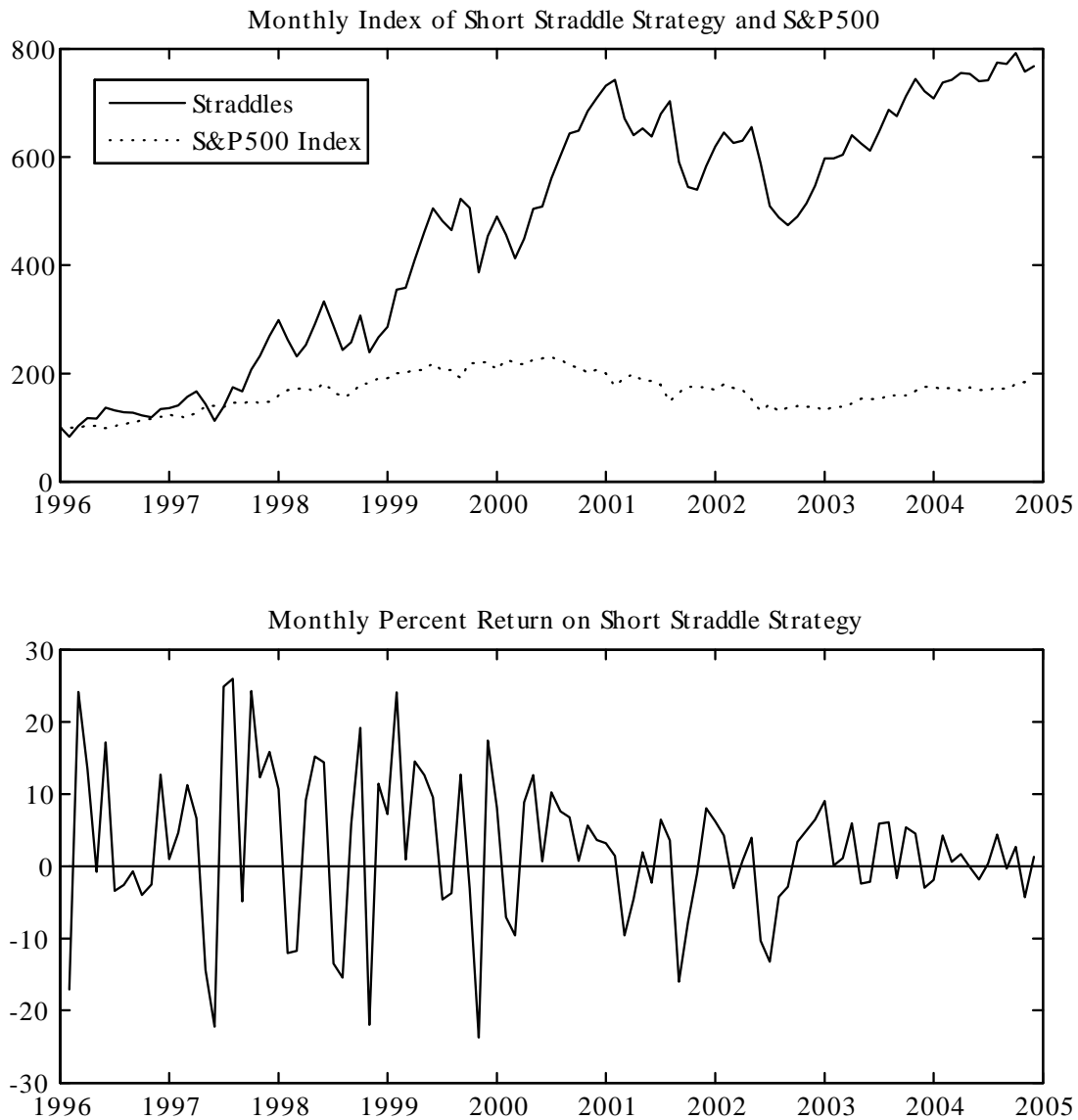
Notes to Figure: The top panel plots the monthly index of a short straddle strategy (solid line) and the S&P500 index (dashes) both normalized to 100 on January 1, 1996. The bottom panel plots the monthly returns on the straddle strategy. Straddle returns are computed using the nearest to at-the-money nearest to 30-day maturity call and put option on the third Friday of every month, with payoffs computed at maturity. The monthly index is computed starting with $100 in cash, and keeping track of the cash account, with the cash account earning the risk-free rate.

Figure 5: Model Implied Log Ratio of the Risk-Neutral and Physical Density.
Various Maturities.
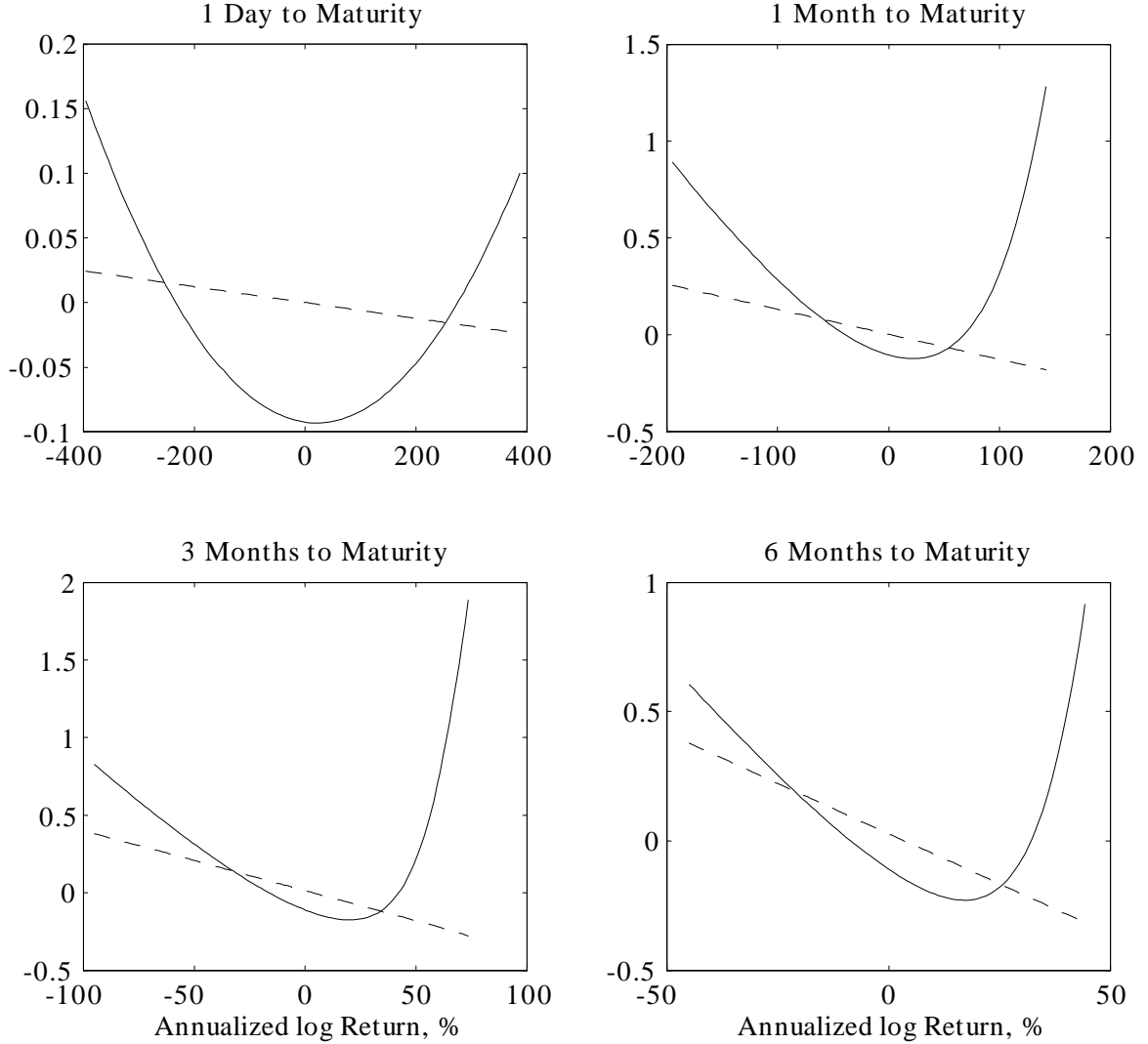
Notes to Figure: We plot the natural logarithm of the ratio of the risk-neutral and physical conditional densities implied by the new GARCH model using solid lines and the special case of a zero variance premium using dashed lines. Each panel corresponds to a different maturity. We use the parameters in the third column of Table 3. The conditional variance is set to its unconditional level. The risk-free rate is set to 5% per year.

Figure 6: Volatility of Variance, Annualized, 1996-2004.
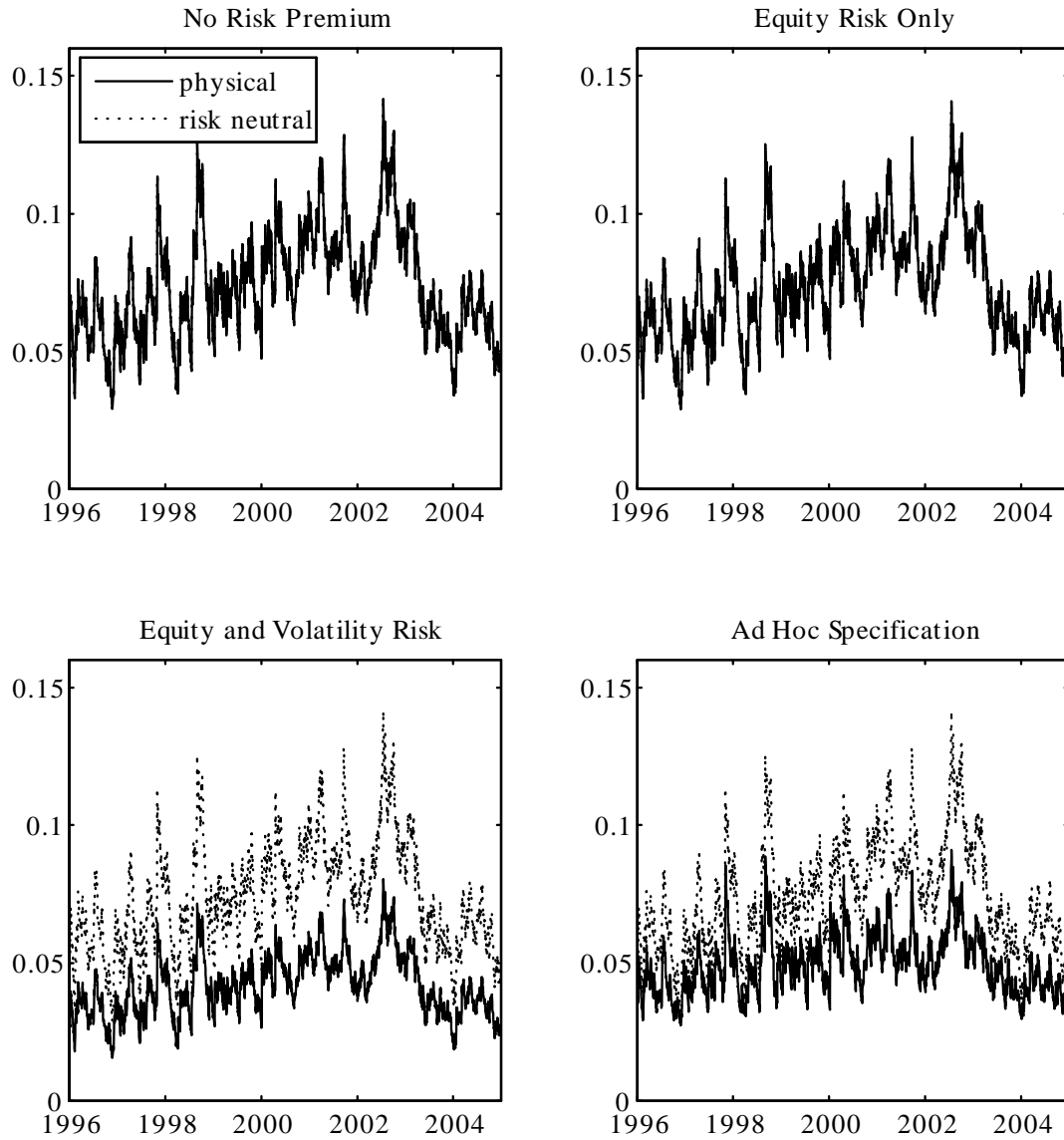
Notes to Figure: We plot the annualized square root of the conditional variance of variance under the physical measure in solid lines and under the risk-neutral measure in dashed lines. The parameters used are reported in Table 3.

Table 1: Returns and Options Data

Panel A: Return Characteristics (Annualized)

	1990-2005	1996-2004
Mean	7.77	7.52
St. Deviation	16.08	19.01
Skewness	-0.102	-0.094
Kurtosis	6.786	5.696

Panel B. Option Data by Moneyness

	F/X<0.96	.96<F/X<.98	.98<F/X<1.02	1.02<F/X<1.04	1.04<F/X<1.06	F/X>1.06	All
Number of Contracts	3,118	1,925	5,688	2,498	1,979	6,183	21,391
Average IV	19.31	18.88	19.71	21.21	22.04	24.91	21.47
Average Price	24.02	27.58	35.28	31.34	28.42	23.40	28.42
Average Spread	1.39	1.40	1.57	1.44	1.35	1.29	1.41

Panel C. Option Data by Maturity

	DTM<30	30<DTM<60	60<DTM<90	90<DTM<120	120<DTM<180	DTM>180	All
Number of Contracts	1,069	5,717	3,996	1,726	3,253	5,630	21,391
Average IV	21.04	21.17	21.69	21.78	21.68	21.49	21.47
Average Price	12.58	17.49	23.61	26.34	31.81	44.62	28.42
Average Spread	0.87	1.14	1.36	1.44	1.53	1.76	1.41

Notes: We present descriptive statistics for daily return data from January 1, 1990 to December 31, 2005, as well as for daily return data from January 1, 1996 to December 31, 2004. We use Wednesday closing OTM options contracts from January 1, 1996 to December 31, 2004.

Table 2: Stein Regression of Overreaction

Sample Period	Coefficient	Standard Error	t-Statistic
Full Sample	-0.241	0.0091	-26.42
1996	-0.358	0.0284	-12.58
1997	-0.195	0.0269	-7.24
1998	-0.298	0.0410	-7.26
1999	-0.311	0.0169	-18.41
2000	-0.211	0.0141	-15.00
2001	-0.211	0.0241	-8.76
2002	-0.167	0.0274	-6.08
2003	-0.231	0.0203	-11.34
2004	-0.364	0.0207	-17.60

Notes: Using 1996-2004 option data, we run the forecasting regressions from Stein (1989, p. 1021). We use at-the-money, fixed-maturity options obtained by fitting a polynomial in maturity and moneyness on every day in the option sample. As in Stein (1989) we use 1-month maturity for short-term options and 2-month maturity for long-term options. We run the regressions for the full sample as well as for each year in the sample separately.

Table 3: Parameter Estimation and Model Fit. Joint Estimation using Returns and Options.

		No Premia	Equity Premium Only	Equity and Volatility Premia	Ad Hoc Specification
Physical Parameters	ω	0	0	0	0
	α	3.248E-06	3.249E-06	1.546E-06	3.452E-06
	β	0.826	0.826	0.826	0.875
	γ	220.9	219.4	318.5	158.7
	μ	0	1.501	1.543	1.954
RN Parameters	$(1-2\alpha\xi)^{-1/2}$	1	1	1.2039	
	ω^*	0	0	0	0
	α^*	3.248E-06	3.249E-06	3.248E-06	3.248E-06
	β^*	0.826	0.826	0.826	0.826
	γ^*	220.93	220.93	220.78	220.93
Likelihood	Total	55,366.1	55,368.4	55,485.5	55,512.2
	From returns	13,188.3	13,190.6	13,307.7	13,334.4
	From options	42,177.8	42,177.8	42,177.8	42,177.8
Physical Properties	Long run volatility	0.234	0.219	0.153	0.151
	Daily autocorrelation, h(t)	0.985	0.983	0.983	0.962
	Annualized volatility of h(t)	0.087	0.081	0.039	0.046
	Corrrelation(R(t),h(t+1))	-0.977	-0.974	-0.974	-0.905
RN Properties	Long run volatility	0.234	0.234	0.232	0.234
	Daily autocorrelation, h*(t)	0.985	0.985	0.985	0.985
	Annualized volatility of h*(t)	0.087	0.087	0.086	0.087
	Corrrelation(R(t),h*(t+1))	-0.977	-0.977	-0.977	-0.977

Notes: Parameter estimates are obtained by optimizing an joint likelihood on returns and options. Parameters as well as autocorrelations are daily. The returns and option samples are described in Table 1. For each model, we report the total likelihood value at the optimum as well as the value of the returns component at the optimum and the option component at the optimum. We estimate four models. In the "Ad Hoc Specification" the physical and risk-neutral parameters are not linked. This model has nine parameters. The "No Premia" has four parameters, with $\mu=0$ and $\xi=0$. The model with "Equity Premium Only" has five parameters. It imposes $\xi=0$. The model with "Equity and Volatility Premia" has six parameters, and estimates μ and ξ. All volatility parameters are constrained to be positive which is a sufficient condition for positive variance.

Table 4: IV RMSE and Bias (%) by Moneyness and Maturity. 1996-2004.
Using Physical Parameters from the Model with Equity and Volatility Premia

Panel A. IV RMSE by Moneyness

Model	F/X<0.96	.96<F/X<.98	.98<F/X<1.02	1.02<F/X<1.04	1.04<F/X<1.06	F/X>1.06	All
Equity and Volatility Premia	2.9741	3.1391	3.1399	3.2364	3.2766	3.8717	3.3685
Equity Premium Only, ξ=0	4.8723	4.7336	5.1241	5.6927	5.9226	6.7394	5.7111
No Premia, ξ=μ=0	5.2882	5.1320	5.5185	6.0926	6.3163	7.1125	6.0986

Panel B. IV Bias by Moneyness (Data Less Model)

Model	F/X<0.96	.96<F/X<.98	.98<F/X<1.02	1.02<F/X<1.04	1.04<F/X<1.06	F/X>1.06	All
Equity and Volatility Premia	-0.6011	-0.7463	-0.3734	0.1899	0.4802	1.3824	0.2121
Equity Premium Only, ξ=0	4.1721	3.7105	4.0581	4.7634	5.0777	5.9468	4.7661
No Premia, ξ=μ=0	4.6481	4.1875	4.5256	5.2185	5.5214	6.3695	5.2190

Panel C. IV RMSE by Maturity

Model	DTM<30	30<DTM<60	60<DTM<90	90<DTM<120	120<DTM<180	DTM>180	All
Equity and Volatility Premia	3.3513	3.4181	3.4614	2.9799	3.4059	3.3447	3.3685
Equity Premium Only, ξ=0	4.7976	5.4619	5.8224	5.6526	5.8944	5.9440	5.7111
No Premia, ξ=μ=0	5.0606	5.7668	6.1823	6.0429	6.3122	6.4292	6.0986

Panel D. IV Bias by Maturity (Data Less Model)

Model	DTM<30	30<DTM<60	60<DTM<90	90<DTM<120	120<DTM<180	DTM>180	All
Equity and Volatility Premia	0.1206	0.7148	0.5890	0.4085	0.0273	-0.5019	0.2121
Equity Premium Only, ξ=0	3.7334	4.4842	4.8342	4.9162	4.9451	5.0505	4.7661
No Premia, ξ=μ=0	4.0889	4.8651	5.2597	5.3583	5.4299	5.5997	5.2190

Notes: We report option implied volatility based RMSE and bias by moneyness and maturity using the parameter estimates from the column labelled "Equity and Volatility Premia" in Table 3. In the row labelled "Equity Premium Only" we force the volatility premium to be zero and in the row labelled "No Premia" we force both the volatility and the equity risk premium to be zero. The physical parameters are the same in all three rows and are taken from the column labeled "Equity and Volatility Premia" in Table 3.