A New Closed-Form Discrete-Time Option Pricing Model with Stochastic Volatility*

Steven Heston
University of Maryland

Kris Jacobs
University of Houston

Hyung Joo Kim
Federal Reserve Board

November 30, 2024

Abstract

In the option pricing literature, closed-form pricing formulas offer many advantages, but very few solutions are available. Among models that can incorporate the critically important stylized fact of stochastic volatility, many are related to the square root model of Heston (1993). Heston and Nandi (2000) offer a discrete-time alternative, but this is a GARCH-type model which does not feature stochastic volatility. We propose a new closed-form discrete-time option pricing model with stochastic volatility. The model is straightforward to implement. We estimate it using (jointly) a long historical time series of index returns and large option panels with various moneyness and maturities. The model vastly outperforms the existing discrete-time Heston-Nandi benchmark and slightly improves on the continuous-time benchmark. The model-implied pricing kernel and risk premiums are very plausible. The newly proposed pricing formula can be used to implement various extensions of the model.

*Heston: sheston@umd.edu; Jacobs: kjacobs@bauer.uh.edu; Kim: hyungjoo.kim@frb.gov. We would like to thank Tim Bollerslev, Christian Dorion, Nicola Fusari, Michael Gordy, Yang-Ho Park, Olivier Scaillet, and seminar participants at the 2024 Cancun Derivatives and Asset Pricing Conference, the 2024 Conference on Derivatives and Volatility, the 2024 NYU Mathematical Modeling in Finance Workshop, the 2024 SoFiE Conference, Federal Reserve Board, Seoul National University, and the University Paris 1 for helpful conversations and comments. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by the Federal Reserve Board or other members of its staff.

1 Introduction

Arguably the most important innovation to the seminal Black and Scholes (1973) model is to allow for time-varying volatility of the underlying asset. In stochastic volatility models, this time-varying volatility contains an innovation to volatility that is independent of the return innovation. Heston (1993) proposes a stochastic volatility dynamic that has two appealing properties: 1) It allows for nonzero correlation between the innovations to the return and the variance; and 2) It results in a (quasi) closed-form price for a European call option. This model is usually referred to as the square-root model. Despite the obvious appeal of closed-form solutions for option prices, it remains the only stochastic volatility dynamic that allows for a closed-form solution and that is relatively straightforward to implement.[1] Partly because of the lack of alternatives, the Heston (1993) square-root dynamic has been used as a building block in richer models with multiple square-root factors and/or jumps in returns and volatility.[2] These models provide a better fit to the cross-section and time-series of option prices.

GARCH option pricing models (Duan, 1995) provide an interesting alternative to stochastic volatility option pricing models. Heston and Nandi (2000) adapt the dynamic of the GARCH model in order to allow for a (quasi) closed-form price for a European call option, similar to the solution in the Heston (1993) model. These GARCH models are formulated in discrete time and therefore do not require discretization. They are also easier to implement and estimate compared to the Heston (1993) model. However, it is well understood that these advantages in implementation directly result from the simpler structure of the GARCH model, specifically from the assumption that the variance can be modeled as a deterministic function of (squared) lagged return innovations. Consequently, while an extensive literature estimates GARCH models using the underlying returns,

[1] See Lewis (2000) and Heston (1997) on the so-called 3/2 model, which also allows for a quasi closed-form solution, but which is much harder to implement.

[2] For examples of such studies, see Bakshi, Cao, and Chen (1997), Bates (2000, 2006, 2019), Chernov and Ghysels (2000), Duffie, Pan, and Singleton (2000), Pan (2002), Chernov (2003), Eraker, Johannes, and Polson (2003), Eraker (2004), Broadie, Chernov, and Johannes (2007), Christoffersen, Jacobs, and Mimouni (2010), Andersen, Fusari, and Todorov (2015a,b, 2017), Hurn, Lindsay, and McClelland (2015), Bardgett, Gourier, and Leippold (2019), and Aït-Sahalia, Karaman, and Mancini (2020). See Bates (2022) for an overview of this literature.

it is often argued that the GARCH functional form may pose problems for option pricing, because it prevents the implied volatility surface from evolving independently from the underlying returns. However, to the best of our knowledge there is no evidence in the literature that quantifies the importance of (relaxing) this assumption.

This paper bridges these different literatures by proposing a new discrete-time stochastic volatility model. The model allows for quasi closed form prices for European call options, like the Heston (1993) and Heston and Nandi (2000) models. It is closely related to the existing discrete-time option pricing literature, because it nests the Heston and Nandi (2000) model, which can be obtained by restricting the parameter that scales the independent volatility innovation. It is also closely related to the continuous-time Heston (1993) model, because it allows the innovations to returns and volatility to be less than perfectly correlated. The model is straightforward to implement. We estimate it using (jointly) a long historical time series of index returns and large option panels with various moneyness and maturities. We first compare its performance to that of the nested Heston and Nandi (2000) model, as well as that of the model of Christoffersen, Heston, and Jacobs (2013), which allows for a more general pricing kernel that is more closely related to the kernel in our new discrete-time stochastic volatility model. The new model vastly outperforms the existing discrete-time benchmarks in terms of option fit. These findings confirm the intuition that the independent volatility shock in stochastic volatility models ought to lead to improved option pricing performance. Moreover, the properties of the pricing kernel implied by the new model are more plausible than the kernels implied by the existing models.

We next compare the empirical performance of the new discrete-time stochastic volatility option pricing model with that of the continuous-time Heston (1993) model. We find that the new model outperforms the continuous-time benchmark, but the mechanics of both models are similar, and the difference in performance mainly results from the different constraints that need to be imposed in both models to keep the variance positive. The level, term structure, and time series properties of the equity and variance risk premiums in the discrete-time model are similar to those in the continuous-time model, and more plausible than the risk premiums implied by the GARCH models.

It is important to note that throughout, we focus on the simplest possible volatility model, in the sense that we implement a model with a single volatility factor. As discussed above, it is well known that adding additional (volatility) factors will improve empirical performance, but our focus on a single factor allows for the clearest and most insightful comparison of the different modeling choices. While we do not pursue extensions in this paper, it is straightforward to use the newly proposed pricing formula as a building block for models with multiple volatility factors, similar to what the literature does with the Heston (1993) pricing formula. We show that, similar to a GARCH model, the new model can also be extended by incorporating additional lags; this type of extension is not available in a continuous-time setup.

The paper proceeds as follows. Section 2 presents the new discrete-time stochastic volatility model and discusses competing models and benchmarks. Section 3 discusses the data and the estimation approach. Section 4 compares the empirical performance of the new model with that of existing GARCH models. Section 5 provides a comparison with the continuous-time Heston (1993) model. Section 6 concludes. Technical material is collected in an Appendix.

2 A New Discrete-Time Stochastic Volatility Model

This section presents the new option pricing model. We first discuss the simplest version of the SV dynamic, which by analogy to the GARCH literature we refer to as the SV(1,1) dynamic. We show that it yields a closed-form solution for the conditional moment-generating function of the log stock price, which in turn leads to a closed-form pricing formula for European options. We characterize the pricing kernel that is consistent with the existence of this functional form for the dynamic under the physical and risk-neutral measures. We show that our results can be extended to more general models with additional lags. We show that the stochastic volatility (SV) dynamic nests the Heston and Nandi (2000) GARCH option pricing model and discuss the relation between the SV(1,1) version of the new dynamic and the (Euler discretization of the) Heston (1993) model.

2.1 The Discrete-Time SV(1,1) Dynamics

We start by characterizing the simplest and most parsimonious model in the new class of models we consider. By analogy to the GARCH nomenclature of models, we refer to this dynamic as the SV(1,1) model. This initial focus on the simplest model is primarily motivated by ease of exposition. It makes it easier to illustrate the relation to the GARCH(1,1) class of option pricing models, which is popular in the empirical literature. The use of the SV(1,1) model also clarifies the relation to continuous-time stochastic volatility models, which we discuss below, and it is straightforward to generalize the model in this section by adding additional volatility factors.

The new model can be thought of as extending the Heston and Nandi (2000) discrete-time GARCH model by allowing for stochastic volatility, inspired by the mechanics of the Heston (1993) stochastic volatility model. For option pricing, we require the risk-neutral stock price dynamic. The SV(1,1) version of the new model specifies the risk-neutral dynamics of the spot index $S(t)$ and its stochastic variance $v^*(t)$ as follows:

$$\begin{aligned} \ln(S(t+1)) &= \ln(S(t)) + \left(r - \frac{1}{2}v^*(t)\right) + \sqrt{v^*(t)}z_1^*(t+1), \\ v^*(t+1) &= \omega^* + \beta^* v^*(t) + (\alpha_1^* z_1^*(t+1) + \alpha_2^* z_2^*(t+1) - \lambda^*\sqrt{v^*(t)})^2, \end{aligned} \tag{1}$$

where each time interval is one day, r is the risk-free rate, and $z_1^*(t)$ and $z_2^*(t)$ follow independent standard normal distributions under the risk-neutral measure. Note that the notation is a bit different from the one typically used in the GARCH literature; specifically, $v^*(t)$, the conditional daily variance of the subsequent compound stock return, is known at time t. We motivate this notational convention below. From a modeling perspective, the model's essential difference with GARCH models is that due to the existence of the independent innovation $z_2^*(t)$, the variance cannot be filtered from the path of stock returns only.

Consistent with most of the existing option pricing literature, we focus on a physical dynamic

that has the same functional form as the risk-neutral dynamic:

$$\begin{aligned} \ln(S(t+1)) &= \ln(S(t)) + \left(r + \left(\mu - \frac{1}{2}\right) v(t)\right) + \sqrt{v(t)} z_1(t+1), \\ v(t+1) &= \omega + \beta v(t) + (\alpha_1 z_1(t+1) + \alpha_2 z_2(t+1) - \lambda\sqrt{v(t)})^2, \end{aligned} \tag{2}$$

where μ is the parameter that identifies the equity premium, and where $z_1(t)$ and $z_2(t)$ follow independent standard normal distributions under the physical measure. We analyze the mapping between the physical and risk-neutral dynamics below in our discussion of (the) pricing kernel(s).

The discrete-time dynamics in equation (2) implies the following moments of the return and the stochastic variance:

$$E(\log R) = r + \left(\mu - \frac{1}{2}\right) E(v), \tag{3}$$

$$E(v) = \frac{\omega + \alpha_1^2 + \alpha_2^2}{1 - (\beta + \lambda^2)}, \tag{4}$$

$$\begin{aligned} Var(v) &= E(Var_t(v(t+1))) + Var(E_t(v(t+1))) \\ &= \frac{2(\alpha_1^2 + \alpha_2^2)(\alpha_1^2 + \alpha_2^2 + 2\lambda^2 E(v))}{1 - (\beta + \lambda^2)^2}, \end{aligned} \tag{5}$$

$$AR1(v) = \beta + \lambda^2, \tag{6}$$

$$Corr_t(\log R(t+1), v(t+1)) = \frac{-2\alpha_1 \lambda v(t)}{\sqrt{v(t)}\sqrt{2(\alpha_1^4 + \alpha_2^4) + 4\lambda^2(\alpha_1^2 + \alpha_2^2)v(t) + 4\alpha_1^2\alpha_2^2}}, \tag{7}$$

The moments in equations (3)-(7) are for the physical dynamic, but they are of course the same for the risk-neutral dynamic subject to the required change of parameters, including $\mu = 0$.

2.2 Option Pricing

Given the risk-neutral dynamic in equation (1), the price of a call option with strike price K and time to maturity τ can be obtained in (quasi) closed form (up to a numerical integration). Our implementation uses the fast Fourier technique of Carr and Madan (1999):

$$C_t(S(t), v^*(t), k, \tau) = \frac{e^{-\alpha k}}{\pi} \int_0^\infty Re\left[e^{-iuk}\psi(u)\right] du, \tag{8}$$

where k is the natural log of K. The function $\psi(u)$ is the Fourier transform of the modified call price, which is the call price multiplied by $e^{\alpha k}$ for $\alpha > 0$. We found that $\alpha = 4$ works well. The function $\psi(u)$ is given by:

$$\psi(u) = \frac{e^{-r\tau} f_\tau^{MG}((\alpha+1) + iu | S(t), v^*(t))}{(\alpha + iu)(\alpha + 1 + iu)},$$

where i is the imaginary unit, and $f_\tau^{MG}(\phi | S(t), v^*(t)) = E_t^*\left[S(t+\tau)^\phi\right]$ is the risk-neutral conditional moment generating function (MGF) of $\log S(t+\tau)$. The closed-form expression of $f_\tau^{MG}(\phi | S(t), v^*(t))$ is given by the following recursive formula:

$$\begin{aligned}
E_t\left[S(T)^\phi\right] &= S(t)^\phi \exp\left(A(t) + B(t)v(t)\right), \qquad (9)\\
A(t) &= A(t+1) + \phi r + \omega B(t+1) - \frac{1}{2}\log\left(1 - 2\left(\alpha_1^2 + \alpha_2^2\right)B(t+1)\right),\\
B(t) &= \phi\left(\mu - \frac{1}{2}\right) + \beta B(t+1) + \frac{2\lambda^2 B(t+1) - 4\alpha_1\lambda\phi B(t+1) + \phi^2(1 - 2\alpha_2^2 B(t+1))}{2 - 4(\alpha_1^2 + \alpha_2^2)B(t+1)},\\
A(T) &= B(T) = 0,
\end{aligned}$$

where $T = t + \tau$. See Appendix A.1 for a detailed derivation.[3] The price of a put option with the same strike price and maturity can be obtained through put-call parity. The option pricing formula in equation (8) does not account for dividends. We follow the existing literature and adjust the index price for the future dividend. Specifically, we use $S(t)e^{-q\tau}$, where q is the dividend yield at time t.

[3]Equation (9) provides the MGF under the physical measure. Note that $\mu = 0$ under the risk-neutral dynamics. The MGF formula also applies to the GARCH dynamics when $\alpha_2 = 0$. The MGF for the continuous-time SV dynamics is given in Heston (1993).

2.3 The Pricing Kernel

To link the risk-neutral and physical dynamics in equations (1) and (2), we characterize a class of exponential-affine pricing kernels that are a function of $S(t)$ and the historical path of $v(t)$. Appendix B shows that the following class of variance-dependent pricing kernels is consistent with these dynamics:

$$M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^{-\gamma}\exp\left(\delta t + \eta\sum_{s=0}^{t-1} v(s) + \xi(v(t) - v(0))\right), \tag{10}$$

where γ and ξ are the index level and variance preference parameters. Note that the logarithm of the pricing kernel (10) is linear in $\log S(t)$ and (the path of) $v(t)$. Henceforth, we therefore refer to equation (10) as the exponential-affine pricing kernel.

The γ and ξ parameters in the pricing kernel are economic parameters for which we have very clear priors. Economies with higher index returns and lower variance indicate good times, in which marginal utility decreases. We therefore expect $\gamma > 0$ and $\xi > 0$. Appendix B shows that given the pricing kernel (10), the risk-neutral and physical dynamics in equations (1) and (2) are linked as follows:

$$\begin{aligned}
\omega^* &= \frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}\omega, && (11)\\
\beta^* &= \beta, && (12)\\
\alpha_1^* &= \frac{\alpha_1}{1-2(\alpha_1^2+\alpha_2^2)\xi}, && (13)\\
\alpha_2^* &= \frac{\alpha_2}{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}}, && (14)\\
\lambda^* &= \left(\frac{\alpha_1\mu+\lambda}{1-2\xi\alpha_2^2} + \frac{\alpha_1^3\xi}{(1-2\xi\alpha_2^2)(1-2(\alpha_1^2+\alpha_2^2)\xi)}\right), && (15)\\
v^*(t) &= \zeta v(t), && (16)
\end{aligned}$$

where $\zeta = \frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}$. The index level and variance preference parameters γ and ξ can be expressed

as functions of the SV(1,1) parameters:

$$\gamma = \mu - \frac{\alpha_1 \xi (2\alpha_1 \mu - \alpha_1 + 2\lambda)}{1 - 2\alpha_2^2 \xi}, \tag{17}$$

$$\xi = \frac{\zeta - 1}{2\zeta(\alpha_1^2 + \alpha_2^2) - 2\alpha_2^2}. \tag{18}$$

Equation (10) contains two other parameters, δ and η. We refer to δ as the time-preference parameter and to η as the path-dependence parameter, respectively. Using the fact that $e^{rt}M(t)$ is a martingale and that the drift term of a martingale process has to be zero, Appendix B shows that δ and η are given by:

$$\delta = (\gamma - 1)r - \xi\omega + \frac{1}{2}\log\left(1 - 2\left(\alpha_1^2 + \alpha_2^2\right)\xi\right), \tag{19}$$

$$\eta = \left(\mu - \frac{1}{2}\right)\gamma + \xi(1 - \beta) - \frac{-2\lambda^2\xi - 4\alpha_1\lambda\gamma\xi + \gamma^2(-1 + 2\alpha_2^2\xi)}{-2 + 4(\alpha_1^2 + \alpha_2^2)\xi}. \tag{20}$$

2.4 The Benchmark GARCH Model

Our empirical strategy demonstrates the benefits of the new model by comparing its performance with that of existing models. We first compare the model's performance with that of discrete-time models. The new discrete-time SV model nests the well-known GARCH option pricing models by Heston and Nandi (2000) and Christoffersen, Heston, and Jacobs (2013) (henceforth HN and CHJ respectively). In particular, the risk-neutral GARCH dynamics in both these models correspond to imposing the restriction $\alpha_2 = 0$ in equation (1):

$$\begin{aligned} \ln(S(t+1)) &= \ln(S(t)) + \left(r - \frac{1}{2}v^*(t)\right) + \sqrt{v^*(t)}z_1^*(t+1), \\ v^*(t+1) &= \omega^* + \beta^* v^*(t) + (\alpha_1^* z_1^*(t+1) - \lambda^*\sqrt{v^*(t)})^2. \end{aligned} \tag{21}$$

As mentioned above, the notation in equation (21) slightly differs from the notation in Heston and

Nandi (2000) and the GARCH literature (Engle, 1982; Bollerslev, 1986), which is as follows:

$$\begin{aligned} \ln(S(t+1)) &= \ln(S(t)) + \left(r - \frac{1}{2}v^*(t+1)\right) + \sqrt{v^*(t+1)}z_1^*(t+1), \\ v^*(t+1) &= \omega^* + \beta^* v^*(t) + \tilde{\alpha}_1^*(z_1^*(t) - \tilde{\lambda}^*\sqrt{v^*(t)})^2. \end{aligned} \tag{22}$$

Our notation in equation (21) facilitates the comparison with the continuous-time benchmark in the next section. Despite the difference in notation, the mechanics in equations (21) and (22) are the same. The variance $v^*(t)$ is known at time t in the former, whereas $v^*(t+1)$ is known at time t in the latter. The one-to-one correspondence between the two sets of parameters is defined by $\alpha_1^* = \sqrt{\tilde{\alpha}_1^*}$ and $\lambda^* = \sqrt{\tilde{\alpha}_1^*(\tilde{\lambda}^*)^2}$. The essence in both notations is that volatility is no longer stochastic in equations (21) and (22), and the volatility path is entirely determined by (lagged) return shocks.

We investigate two GARCH models. When $\alpha_2 = 0$ is the only restriction, the model corresponds to the CHJ model, together with the variance-dependent pricing kernel. The HN model is obtained by additionally imposing $\xi = 0$, or equivalently $\zeta = 1$. This clarifies that the CHJ and HN model only differ in their specification of the pricing kernel.

2.5 The SV(p,q) Dynamics

One of the advantages of a discrete-time approach is that it does not require additional assumptions on discretization. Another advantage is that the model can be extended in a very straightforward way by generalizing the autoregressive or moving average structure of the variance dynamic. Appendix A.2 shows that the following extension leads to a closed-form option price:

$$\begin{aligned} \ln(S(t+1)) &= \ln(S(t)) + \left(r + \left(\mu - \frac{1}{2}\right)v(t)\right) + \sqrt{v(t)}z_1(t+1), \\ v(t+1) &= \omega + \sum_{i=1}^{p}\beta_i v(t+1-i) \\ &\quad + \sum_{i=1}^{q} k_i\left(\alpha_1 z_1(t+2-i) + \alpha_2 z_2(t+2-i) - \lambda_i\sqrt{v(t+1-i)}\right)^2, \end{aligned} \tag{23}$$

where $k_1 = 1$. This SV(p,q) dynamics nests the GARCH(p,q) dynamics in Heston and Nandi (2000) and Christoffersen, Jacobs, and Ornthanalai (2013), again with the slight difference in notation as discussed in Section 2.4; specifically, when $\alpha_2 = 0$, the SV(p,q) model corresponds to the GARCH(p,q) dynamic. We do not pursue an empirical investigation of this more general model here, but leave it for later work.

2.6 The Benchmark Continuous-Time Dynamics

Our main empirical contribution is the comparison between the new SV model and existing GARCH models. However, a natural question is how the new discrete-time SV model compares to continuous-time SV models. We therefore also provide results for the Heston (1993) SV model, which specifies the risk-neutral dynamics of the spot price $S(t)$ and its stochastic variance $v(t)$ as follows:

$$\begin{aligned} d\ln S(t) &= \left(r - \frac{1}{2}v(t)\right)dt + \sqrt{v(t)}dz_1^*(t), \\ dv(t) &= \kappa^*(\theta^* - v(t))dt + \sigma\sqrt{v(t)}dz_2^*(t), \end{aligned} \tag{24}$$

where dz_1^* and dz_2^* are Wiener processes with correlation coefficient ρ. We again consider a physical dynamic that has the same functional form as the risk-neutral dynamic.

$$\begin{aligned} d\ln S(t) &= \left(r + \left(\mu - \frac{1}{2}\right)v(t)\right)dt + \sqrt{v(t)}dz_1(t), \\ dv(t) &= \kappa(\theta - v(t))dt + \sigma\sqrt{v(t)}dz_2(t). \end{aligned} \tag{25}$$

The physical and risk-neutral dynamics are linked by the following variance-dependent pricing kernel:

$$M(t) = M(0)\left(\frac{S(t)}{S(0)}\right)^{-\gamma}\exp\left(\delta t + \eta\int_0^t v(s)ds + \xi(v(t) - v(0))\right). \tag{26}$$

It is straightforward to see that this pricing kernel is a continuous-time counterpart of the discrete-time pricing kernel in equation (10).

To see the relation with the discrete-time SV model in equations (1) and (2), it is useful to consider the implementation of these continuous-time dynamics using the Euler discretization. Under the physical measure, we have:

$$\begin{aligned} \ln(S(t+1)) &= \ln(S(t)) + \left(r + \left(\mu - \frac{1}{2}\right) v(t)\right) + \sqrt{v(t)} z_1(t+1), \\ v(t+1) &= \kappa\theta + (1-\kappa) v(t) + \sigma \sqrt{v(t)} z_2(t+1). \end{aligned} \tag{27}$$

The similarities with equation (2) are clear, but there are some subtle differences. Due to technical restrictions, the continuous-time risk-neutral and physical dynamics rely on the same variance path $v(t)$; moreover, the parameters σ and ρ are identical under both measures.

3 Data and Estimation

We first discuss the return and option data. Then we present the likelihood functions and discuss our estimation approach.

3.1 Data

Our empirical analysis uses out-of-the-money (OTM) S&P 500 call and put options with maturities between 14 and 365 days. The sample period is from January 1996 to December 2021. We obtain option data from OptionMetrics. We apply the following filters:

1. Discard options with implied volatility smaller than 5% or greater than 150%.
2. Discard options with volume or open interest less than ten contracts.
3. Discard options with mid price less than \$0.50 or bid price less than \$0.375 to avoid low-valued options.
4. Discard options with data errors – when bid price exceeds offer price, or when a negative price is implied through put-call parity.
5. Discard options with moneyness less than 0.75 or greater than 1.25.

We include options that expire on the third Friday or the Saturday following the third Friday of the current month, the next month, and the following March, June, September, and December (consistent with standard option contracts).[4] We keep the six most actively traded strike prices for each available maturity. It is important to use as long a time period as possible, in order to be able to identify key aspects of the model including volatility persistence (Broadie, Chernov, and Johannes, 2007). On the other hand, estimation using large option panels and long time series is very time-intensive. Rather than using a short(er) time series of daily option data, we use an extended time period, but we select option contracts for one day per week only. Following several existing studies (see, e.g., Heston and Nandi, 2000; Christoffersen, Heston, and Jacobs, 2013), 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 to be affected by day-of-the-week effects. These steps result in a dataset with 40,226 option contracts. Table 1 presents descriptive statistics.

We obtain S&P 500 index returns from CRSP. We use data for January 1990 to December 2021. This sample period is longer than the option sample to help with the identification of the return parameters under the physical measure, as in Christoffersen, Heston, and Jacobs (2013). We also use data on the VIX from January 1990 to December 2021, which we obtain from the Federal Reserve Bank of St. Louis Economic Database. The time series for the risk-free rate is proxied by the one-month Treasury Bill rates obtained from CRSP. Following the standard implementation in the literature, options are valued using a maturity-specific risk-free rate. We apply a cubic spline interpolation to the risk-free rates obtained from OptionMetrics.

3.2 The Stochastic Model Variance and the VIX

In the new model as well as in the Heston (1993) model and its many generalizations studied in the literature, the stochastic variance is unknown. This latency is typically addressed in estimation by using filtering- or simulation-based techniques (see, e.g., Eraker, Johannes, and Polson, 2003;

[4] Until January 2015, standard option contracts expired on the Saturday following the third Friday of the expiration month. Standard options listed after August 2013 and expiring after January 2015 expire on the third Friday of the month.

Eraker, 2004; Bates, 2006; Christoffersen, Jacobs, and Mimouni, 2010). It is well known that the implementation of such techniques is computationally very demanding, especially when using long time series and large cross-sections of option prices in estimation.

To alleviate the computational burden, we use the fact that the daily stochastic variance $v^*(t)$ can be represented as a linear function of $\text{VIX}_d^2(t)$, where VIX_d represents the daily VIX.[5] This directly follows from the model specification. When $v(t)$ follows the dynamic in equation (1), $\text{VIX}_d^2(t)$ is a linear function of $v^*(t)$. Specifically, the model-implied $\text{VIX}_d^2(t)$ can be written as:

$$\text{VIX}_d^2(t) = \psi_0 + \psi_1 v^*(t) \tag{28}$$

where $\psi_0 = \frac{\theta_v(1-\psi_1)}{1-\rho_v}$, $\psi_1 = \frac{\rho_v(1-\rho_v^n)}{n(1-\rho_v)}$, $\theta_v = \omega^* + (\alpha_1^*)^2 + (\alpha_2^*)^2$, $\rho_v = \beta^* + (\lambda^*)^2$, and $n = 21$ is the number of trading days per month. Appendix C provides a detailed derivation.

Using equation (28) and an assumption on the measurement error, we obtain a measurement equation which can be used to filter the latent state variable. Jones (2003), Cheung (2008), and Chernov, Graveline, and Zviadadze (2018) use this measurement equation and a Bayesian framework with Markov Chain Monte Carlo methods to estimate option pricing models. We further simplify the setup and instead use:

$$v^*(t) = \eta_0 + \eta_1 \text{VIX}_d^2(t). \tag{29}$$

This implementation follows Aït-Sahalia and Kimmel (2007), who use it in a sample that contains a single short-maturity at-the-money option at each time t. However, we experienced some identification problems with equation (29) and therefore assume:

$$v^*(t) = \eta_1 \text{VIX}_d^2(t). \tag{30}$$

We then use equation (30) in the valuation formula for all options in the sample.

[5] See Bates (2000) and Andersen, Fusari, and Todorov (2015a) for other approaches.

3.3 The Return-Based Likelihood

In the previous subsection, we use the fact that the risk-neutral stochastic variance is an affine function of VIX to alleviate the high computational burden when the model is implemented on option data. This assumption also has implications for the return-based estimation. In most existing studies, the variance is filtered from the underlying returns, and the VIX is not used in the return-based estimation. However, if the physical stochastic variance is proportional to the risk-neutral stochastic variance, which is the case for the class of pricing kernels in Section 2.3, we can formulate the joint likelihood function of the return and the VIX_d^2 data to estimate the physical parameters, since we observe the total return on the stock index and the VIX at each time t. We obtain the likelihood function by rearranging equation (2) as follows:

$$\begin{aligned}
\log R(t+1) &= r + \left(\mu - \frac{1}{2}\right) v(t) + \sqrt{v(t)} z_1(t+1), \\
v(t+1) &= \omega + \beta v(t) + (\alpha_1 z_1(t+1) + \alpha_2 z_2(t+1) - \lambda\sqrt{v(t)})^2,
\end{aligned}$$

where $R(t+1) = S(t+1)/S(t)$ represents the gross return. Bayes' rule implies the following log-likelihood function:

$$\begin{aligned}
\log L^R &= \sum_{t=1}^{T-1} \log f\left(\log R(t+1), \text{VIX}_d^2(t+1) | \text{VIX}_d^2(t)\right) \\
&= \sum_{t=1}^{T-1} \log\left[f(\log R(t+1)|v(t)) \times f(v(t+1)|\log R(t+1), v(t)) \times J_v(t+1)\right], \qquad (31)
\end{aligned}$$

where $f(\log R(t+1)|v(t))$ is the probability density of $\log R(t+1)$, $f(v(t+1)|\log R(t+1), v(t))$ is the probability density of $v(t+1)$ conditional on $\log R(t+1)$, and $J_v(t+1)$ is the Jacobian between $\text{VIX}_d^2(t+1)$ and $v(t+1)$.

Note that $f(R(t+1)|v(t))$ follows a normal distribution governed by z_1. Because $\left(\frac{\alpha_1}{\alpha_2} z_1 + z_2 - \frac{\lambda}{\alpha_2}\sqrt{v}\right)$ follows a normal distribution with a nonzero mean, conditional on z_1 and v, $f(v(t+1)|R(t+1), v(t))$ follows a non-central chi-squared distribution. For expositional purposes, we denote

$y(t+1) \equiv \frac{v(t+1)-\omega-\beta v(t)}{\alpha_2^2} = \left(\frac{\alpha_1}{\alpha_2} z_1(t+1) + z_2(t+1) - \frac{\lambda}{\alpha_2}\sqrt{v(t)}\right)^2$. The log-likelihood function in equation (31) can therefore be expressed as:

$$\log L^R = \log L_r^R + \log L_v^R, \tag{32}$$

where

$$\begin{aligned}
\log L_r^R &= \sum_{t=1}^{T-1} -\frac{1}{2}\log(2\pi v(t)) - \frac{z_1(t+1)^2}{2v(t)}, \\
\log L_v^R &= \sum_{t=1}^{T-1} \log\frac{1}{2} - \frac{y(t+1)+\chi(t+1)}{2} - \frac{1}{4}\log\frac{y(t+1)}{\chi(t+1)} + \log I_{(-\frac{1}{2})}\left(\sqrt{\chi(t+1)y(t+1)}\right) \\
&\quad + \log\eta_1 - 2\log\alpha_2,
\end{aligned}$$

and where $\chi(t+1) = \left(\frac{\alpha_1}{\alpha_2} z_1(t+1) - \frac{\lambda}{\alpha_2}\sqrt{v(t)}\right)^2$ and $I_{(p)}$ is a p-th order modified Bessel function of the first kind.[6] The log likelihood in the continuous-time SV model also has these two components, but the functional form is slightly different. See Heston, Jacobs, and Kim (2023) for details. GARCH models do not contain separate shocks to variance, and thus the log-likelihood function consists of only $\log L_r^R$. See Christoffersen and Jacobs (2004) and Christoffersen, Heston, and Jacobs (2013) for details.

3.4 The Option-Based Likelihood

We use vega-weighted option pricing errors. Let O_i^{Mkt} denote the market price of the i^{th} option and O_i^{Mod} the model price computed using the method in Section 2.2. O_i^{Mkt} and O_i^{Mod} represent call option prices if $F/K < 1$ and put option prices if $F/K > 1$, where $F = Se^{(r-q)\tau}$ denotes the

[6]The log-likelihood function consists of two components: one based on the return dynamics and the other based on the variance dynamics. The model has two state variables, $R(t)$ and $v(t)$. Although $v(t)$ is unobservable, we assume it is a function of the observable $\text{VIX}^2(t)$. We can therefore express the joint likelihood of the two state variables with the aid of a Jacobian. If we instead assume that the stochastic variance is latent, we cannot directly obtain its transition likelihood and we have to rely on filtering to obtain the next-period variance, using Markov Chain Monte Carlo or particle filtering techniques for example. For additional discussion, see Aït-Sahalia and Kimmel (2007).

implied futures price. The vega-weighted option pricing errors are defined as

$$\epsilon_{o,i} = \frac{O_i^{Mkt} - O_i^{Mod}}{\nu_i^{Mkt}},$$

where ν_i^{Mkt} is the Black-Scholes vega of option i. Vega-weighted errors have advantages that are similar to implied volatilities, i.e. the magnitudes are more similar across options compared to the option prices, but they also offer a computational advantage. See for example Carr and Wu (2007), Trolle and Schwartz (2009), and Christoffersen, Heston, and Jacobs (2013) for a discussion.

Maximum likelihood estimation requires a distributional assumption on the errors. Following most of the existing literature, we assume that $\epsilon_{o,i}$ follows a normal distribution, *i.e.* $\epsilon_{o,i} \sim N(0, s_o^2)$, where s_o^2 is the sample variance of the errors. With the additional assumption that the option valuation errors are independently and identically distributed (i.i.d.), the log-likelihood function $\log L^O$ is given by:

$$\log L^O = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log s_o^2 - \frac{1}{2s_o^2}\sum_{i=1}^{N}\epsilon_{o,i}^2, \tag{33}$$

where N is the number of option contracts in the sample.

3.5 Joint Estimation Using Index Returns and Options

We jointly estimate the risk-neutral and physical dynamics based on a joint likelihood for options and the underlying returns, imposing the relation between the physical and the risk-neutral parameters in equations (11)-(16). This joint estimation identifies the pricing kernel and risk premiums, and is helpful to detect model misspecification (Bates, 2003). The number of observations in the option and return datasets is very different. To assign equal weights to each datapoint, we impose an ad-hoc assumption and standardize the return-based and option-based log-likelihoods by $T-1$ and N, respectively. Hence, we solve the following optimization problem:

$$\max_{\Theta} \ \frac{(T-1)+N}{2}\left(\frac{1}{T-1}\log L^R + \frac{1}{N}\log L^O\right), \tag{34}$$

where $\Theta = \{\beta,\ \alpha_1,\ \alpha_2,\ \lambda,\ \eta_1,\ \zeta\}$ for the discrete-time SV(1,1) model. For the HN GARCH model, $\Theta = \{\beta,\ \alpha_1,\ \lambda\}$ because $\alpha_2 = 0$ and $\zeta = 1$ (and $\eta_1 = 0$). For the CHJ GARCH model, we have $\Theta = \{\beta,\ \alpha_1,\ \lambda,\ \zeta\}$, and for the continuous-time SV model, we have $\Theta = \{\kappa,\ \theta,\ \sigma,\ \rho,\ \eta_1,\ \xi\}$. Note that in the discrete-time models, we impose $\omega = 0$ to ensure positive variances. We discuss this further below. In the continuous-time model, we impose the Feller (1951) condition. We use mean targeting to capture the level of the unconditional equity premium in all models. That is, we fix the value of μ at $(\hat{E}(R) - \hat{E}(r))/\hat{E}(v)$, where $\hat{E}(R)$ and $\hat{E}(v)$ are the sample mean and variance of the return, respectively, and $\hat{E}(r)$ is the sample mean of the risk-free rate.

4 Discrete-Time Stochastic Volatility vs GARCH Option Pricing

We discuss parameter estimates for the HN GARCH, CHJ GARCH and discrete-time SV option pricing models, based on maximization of the joint likelihood of returns and options in Section 3.5. We compare option fit in these three models, document and compare the economic properties of the models, and discuss the implied estimates of the pricing kernel and risk premiums.

4.1 Parameter Estimates

Table 2 presents the parameter estimates for the three discrete-time models. Columns (1) and (2) report on the HN and CHJ GARCH models, respectively. Recall that the difference between the two models is the existence of variance risk aversion in the CHJ GARCH, which is absent in the HN model. The parameter estimates are reasonable. The mean μ is set to 2.879 by mean targeting, as discussed in Section 3.5. The estimates of λ and λ^* are positive in both columns, consistent with negative skewness under the physical and risk-neutral measures. The ζ parameter identifies the wedge between the physical and risk-neutral variances, see equation (16). For the HN model, zero variance risk aversion results in the absence of a wedge, which means that ζ is fixed at one. The estimates indicate a sizable wedge (1.084) for the CHJ model. This is smaller than the estimate in Christoffersen, Heston, and Jacobs (2013). We verified that this difference is due to the sample period. Another implication of (the absence of) variance risk aversion is the relation between the

α_1 and α_1^* parameters. In the HN model, $\alpha_1 = \alpha_1^*$, whereas the CHJ model has different values for α_1 and α_1^*.

The parameter estimates for the discrete-time SV model in column (3) are also reasonable. Note that α_1 and α_2 scale the shocks that drive the (conditional) volatility of variance. In the GARCH model, there is a single variance shock; in the SV model, the variance is driven by the return shock and an independent variance shock. The estimates of α_1 and α_2 are therefore smaller than the level of α_1 in the two GARCH models. Interestingly, the estimate of the wedge (ζ) parameter is much smaller than in the CHJ GARCH model, but it exceeds one and is statistically different from one, implying the existence of variance risk aversion.

We impose $\omega = \omega^* = 0$ in implementation for the two GARCH models and the SV model. A nonnegativity restriction on these parameters is necessary to ensure positivity of the conditional variance, and we verified that these restrictions are binding. This finding is consistent with existing estimates for GARCH models from options in Christoffersen, Heston, and Jacobs (2013), and it is well known that this is due to the affine structure of the dynamic.

4.2 Model-Implied Moments

Panel A of Table 3 presents the statistical moments of the log return and variance based on equations (3)-(7). These expressions highlight the role of the various parameters in the model and clarify how the parameters are identified. The variance autocorrelation is determined by β and λ. Given this autocorrelation, the long-run variance ($E(v)$) is determined by α_1 and α_2 (recall that we impose $\omega = \omega^* = 0$). Furthermore, α_1 and α_2 are important in anchoring the variance of variance and the conditional correlation between the log return and variance; α_1 is directly linked to the return shock. This ensures that α_1 and α_2 can be separately identified. Finally, λ affects the variance of variance and the conditional correlation. This identifies λ, and in turn, β can be determined from the variance autocorrelation.

Because we use mean-targeting, we match the expected log return with its data counterpart in all models. Panel C of Table 1 shows the annualized sample mean is equal to 0.102. The

annualized physical long-run variance for the models in columns (1)-(3) of Table 3 ranges from 0.0306 to 0.0349, implying a long-run annualized volatility between 0.175 and 0.187, close to the sample volatility of 0.181. The risk-neutral long-run variance exceeds the physical variance due to the presence of the variance risk premium. The annualized physical long-run volatility of variance is 0.449 for the HN model, 0.389 for the CHJ model, and 0.746 for the SV model. For comparison, the sample volatility of the annualized VIX_d^2, which is 0.76. Based on this, the SV model captures the volatility of variance more precisely than the two GARCH models, but this comparison is a bit tenuous since VIX itself represents a risk-neutral volatility. The autocorrelation of daily variance is approximately 0.99 in all cases; consistent with the existing literature, the risk-neutral autocorrelation slightly exceeds the physical autocorrelation because $\lambda^* > \lambda$.

The most critical difference between the models is the conditional correlation between the log return and the variance.[7] The signs of λ and α_1 imply that $Corr_t(\log R(t+1), v(t+1))$ is negative, indicating negative physical conditional skewness, and a similar remark holds for risk-neutral conditional skewness. However, in the GARCH models, both the physical and risk-neutral correlations are very close to minus one, while the correlation is -0.785 (-0.786) under the physical (risk-neutral) measure in the SV model. In other words, in the GARCH models, returns and variance almost always move in opposite directions. This finding suggests that it is challenging to separately identify return and variance risk in the GARCH models based on the dynamics of the stock return and variance. Because returns and variances sometimes move in the same direction, imperfect correlation is required to match the data.

4.3 Option Pricing

We optimize the joint likelihood of options and returns in all models, but we cannot directly compare the total likelihoods of the discrete-time SV and GARCH models, because the return likelihood of the SV model contains an additional component that originates in the independent volatility shock (see equation (32)). However, we can compare the option component $\log L^O$, and Table 2

[7] We compute this correlation assuming that the conditional variance $v(t)$ is equal to its long-run mean.

indicates that the SV model outperforms both GARCH models. The improvement in overall option fit is also evident from the vega-weighted root mean squared errors (RMSE) of option prices. The vega-weighted RMSE for the SV model is 0.0231, a 37.9% improvement over the 0.0372 RMSE for the GARCH models.[8]

Table 4 reports option fit by moneyness and maturity. We present the vega-weighted RMSEs as well as the implied volatility (IV) RMSEs to provide more intuition. The SV model outperforms the GARCH model across all moneyness and maturity buckets, with the most significant improvements observed for near-the-money options and deep out-of-the-money put options in moneyness dimension, and for short- and mid-term maturity options (less than 120 days to expiration). The RMSE patterns are mostly similar across models, except for the maturity pattern in the RMSEs of the GARCH model.

4.4 Pricing Kernels

We start our investigation of the pricing kernels with the model-implied risk preference parameters, based on equations (17)-(18). The definition of ζ indicates that $\zeta > 1$ is equivalent to a positive variance risk aversion parameter ξ. From equation (17), the index level risk aversion parameter γ is a function of the equity premium parameter μ, but it also depends on ξ. The HN model in column (1) assumes $\xi = 0$, and therefore γ is exclusively determined by μ and is positive. The CHJ model in column (2) does not impose restrictions on the signs of the risk preference parameters. The model estimates in Panel A of Table 3 show that ξ is very large and positive while γ is very large and negative, which is inconsistent with economic intuition. A possible explanation is that ξ does not exclusively capture variance risk aversion due to identification problems resulting from the correlation between the index return and variance. This intuition is confirmed by the HN GARCH estimates, where γ is positive. The parameter estimates for the discrete-time SV model in column (3) imply positive estimates of both γ and ξ, which is economically plausible.

[8]The HN and CHJ GARCH models have the same risk-neutral dynamic. However, their option fit can differ when maximizing the joint likelihood due to the different mapping between the two measures. The RMSEs indicate that these differences in option fit are on average negligible.

Next, we go beyond the sign of pricing kernel parameters, and we investigate if the estimates of the variance risk aversion parameter ξ are plausible. Consider a one-day increase in (annualized) volatility from 20% to 22%.[9] For simplicity, assume that the one-day return is 0%. Under this assumption, the pricing kernel is primarily influenced by the variance aversion component, $\exp(\xi(v(t+1)-v(t)))$, because the sum of the time-preference and the path-dependence components is very small in both the CHJ and the SV models. In the CHJ model ($\xi = 26,625.97$), this results in a daily pricing kernel of 2.429 for this parameterization. For the SV model ($\xi = 347.95$), the kernel value is 1.012. The former seems excessively large, while the latter value of the kernel is plausible. We conclude that identifying variance risk is difficult in the CHJ GARCH model, which is related to our findings on the conditional correlation between the log return and the variance in Section 4.2.

We now further explore the economic implications of these parameter estimates. Panels A, B, and C of Figure 1 plot daily log pricing kernels as a function of return and variance. The HN GARCH model does not consider variance risk aversion, and the pricing kernel is therefore a simple decreasing function of the return. The kernel with variance risk aversion is economically plausible for the SV model in Panel C, but in the CHJ GARCH model in Panel B, the pricing kernel increases with the return (conditional on the variance level), which reflects the implausible sign of the index level risk aversion parameter γ. The problems with the kernel in Panel B are also evident from inspecting its range. The log pricing kernel ranges between -5 to 15, which means that the pricing kernel ranges from 0.01 to 3,269,017.37, unrealistic values for a daily discount factor.

Panels A, B, and C of Figure 2 plot the time series of the realized log pricing kernel over the sample period.[10] The time series of the HN kernel in Panel A and that of the SV kernel in Panel C appear plausible, exhibiting large fluctuations in crisis periods such as the 2008-2009 financial crisis and the Covid-19 pandemic period. Once again, the time series of the CHJ GARCH kernel in Panel B seems implausible. It contains many outliers during relatively calm periods, and these

[9] In daily units, from $v(t) = 1.587E-04$ to $v(t+1) = 1.921E-04$.

[10] See Chernov (2003) and Ghosh, Julliard, and Taylor (2017) for other estimates of the time series of the pricing kernel.

outliers are excessively large.[11] Moreover, it exclusively contains large positive outliers. These positive spikes in the CHJ GARCH kernel are directly related to the combination of negative γ and positive ξ estimates. We frequently observe a joint occurrence of (large) positive index returns and changes in variance. In these instances, the negative γ and positive ξ estimates mechanically imply large positive spikes in the pricing kernel. To illustrate this more clearly, consider the back-of-the-envelope calculation above, but now assume a daily return of 1%. For the CHJ model, the value of the resulting pricing kernel is 3.379 (1.218 for the log pricing kernel), which is clearly extreme. However, Panel B in Figures 1 and 2 shows that such positive spikes are often attained.

4.5 Conditional Risk Premiums

4.5.1 The Risk Premium Level

Based on the closed-form solution for the moment generating function in equation (9) and the quadratic variation (in Appendix C), we can compute conditional moment conditions for the τ-horizon index return by inserting $S(t) = 1$. We define the τ-horizon integrated equity and variance risk premiums as follows:

$$\begin{aligned} \mathrm{ERP}_t(\tau) &= E_t\left[R(t+\tau)\right] - E_t^*\left[R(t+\tau)\right] \\ \mathrm{VRP}_t(\tau) &= E_t\left[\sum_{s=1}^{\tau} v(t+s)\right] - E_t^*\left[\sum_{s=1}^{\tau} v^*(t+s)\right]. \end{aligned} \tag{35}$$

To annualize both risk premiums, we multiply the above expressions by $252/\tau$.

Panel B of Table 3 reports the one-month conditional first and second moments of returns under both the physical and risk-neutral measures, along with the corresponding risk premiums. Note that the conditional equity and variance risk premiums in equation (35) are functions of the time-t variance, which we obtain from our estimation. Using the variance data, we compute the time series of the integrated conditional moments and compute the time-series averages. In columns (1)-(3), the (annualized) average equity premium closely matches the sample average (0.0942).

[11] The scale in Panels A and C of Figure 2 is the same as in Panel B to facilitate comparisons.

The average risk-neutral second moments exceed the average physical second moments, resulting in negative variance risk premiums for all three models.

Figure 3 plots the time series of the (annualized) integrated one-month conditional equity and variance risk premiums. Panels A and B of Figure 4 plot the corresponding distributions computed using kernel density estimation.[12] The time series of the one-month equity premium in the two GARCH models and the discrete-time SV model display a lot of similarities. As expected, it is consistently positive over time and spikes up during economic downturns such as the 2008-2009 financial crisis and the 2020 Covid-19 crisis. One interesting difference is that the equity risk premium in the SV model (Panel E of Figure 3) displays greater volatility and more pronounced spikes during crisis periods compared to the variance in the GARCH models (Panels A and C). This is due to the independent volatility shock in the SV model, which induces higher volatility of return variance, as confirmed by Panel A of Table 3.[13]

A comparison of the variance risk premiums in Panels B, D, and F of Figure 3 shows that the time series of the VRP in the discrete-time SV model in Panel F is highly correlated with the CHJ VRP in Panel D. However, the CHJ VRP is more volatile and more negative compared to the SV model during low-volatility periods (1999-2002), whereas the latter is slightly more negative during economically adverse periods. The VRP time series in the HN GARCH model is much less volatile than the VRPs in the two other models, due to the absence of a variance risk aversion parameter. Panel B of Figure 4 clearly highlights these differences between the models. Comparing Panels A and B shows that the VRP levels differ much more between models than the ERP levels.

4.5.2 The Risk Premium Term Structure

Panels A, C, and E of Figure 5 plot the difference between the 1-month and 6-month equity risk premiums. Panel C of Figure 4 plots the corresponding distributions computed using kernel

[12] We use a Gaussian kernel with Scott's (1992) rule-of-thumb bandwidth, $1.06\hat{\sigma}_d n_d^{-1/5}$, where $\hat{\sigma}_d$ is the sample standard deviation of observations and n_d is the number of observations.

[13] This finding may also be due to our estimation method, because we assume that the stochastic variance is proportional to the VIX in the implementation of the SV model, and the VIX fluctuates more than GARCH-filtered volatility.

density estimation. The distributions in Figure 4 suggest that, like the 1-month ERP levels, the average ERP term structure is also rather similar across the three models. The models also have similar implications for the relation between the sign of the term structure and the economic regime. In bad times, the 1-month risk premium exceeds the 6-month premium. However, in calm period, the sign reverses. The models capture the important stylized fact that when volatility is high, investors demand a higher compensation for equity risk over short horizons than over long horizons.[14] However, the time series in Figure 5 highlight some interesting differences between the models. Notably, when volatility spikes, the difference between the short- and long maturity premiums (the inverse of the term structure slope) is larger in the discrete-time SV model than in the GARCH models.

Panels B, D, and F of Figure 5 plot the difference between the 1-month and 6-month VRPs, and Panel D of Figure 4 plots the corresponding distributions. Note that since the VRP is negative, a positive value indicates that the longer maturity VRP is more negative than the shorter-maturity VRP. As with the levels, the VRP term structure differs more across models than the ERP term structure. Consistent with the findings of Li and Zinna (2018) and Aït-Sahalia, Karaman, and Mancini (2020), the (absolute value of the) longer-maturity VRPs exceeds the short-maturity VRP in all models. However, the slope of the term structure during downturns is much steeper in the SV model in Panel F compared to the GARCH models in Panels B and D. In the CHJ GARCH model, the term structure slope in downturns is also flatter compared to the HN model. This is due to the combination of mean-reverting variance and the large variance risk aversion parameter ξ.

Figure 6 provides additional evidence by plotting the average (unconditional) term structure of annualized equity and variance risk premiums in Panels A and B, respectively, for maturities up to one year. In Panel A, the annualized equity premium is relatively flat across horizons, and the estimates are very plausible.[15] In Panel B, the variance risk premium is downward sloping

[14]The same finding obtains when using longer horizons, up to 12 months. Panel E of Figure 4 plots the distributions of the 1-month minus 12-month ERPs.

[15]A related literature computes the equity premium term structure based on the information in short-term and

for all models. At short horizons, the VRPs are quite similar across models, but the discrete-time SV model has a steeper term structure. Regarding the magnitudes of the VRPs, Aït-Sahalia, Karaman, and Mancini (2020) report average premiums of -0.006 for the 2-month horizon, -0.012 for the 6-month horizon, and -0.016 for the 12-month horizon. Even though our sample period is different and we consider simple one-factor models, our estimates are similar. Panel C of Figure 6 highlights the most important stylized fact of the equity premium term structure. In all models, the 1-month risk premium is (much) more volatile than the longer-horizon risk premiums. This is due to the changes in the conditional equity term structure across economic regimes discussed above. In contrast, Panel D of Figure 6 shows that the VRP is more volatile at longer maturities for all models.

5 SV Option Pricing in Discrete and Continuous Time

5.1 Comparison Between the Two SV Models

An often-mentioned advantage of continuous-time models is their analytical tractability. Indeed, while our discrete-time SV model allows for a closed-form option price, it requires solving a difference equation, while the affine square-root model in Heston (1993) does not face a similar requirement. As discussed above, the discrete-time approach also provides distinct advantages. It does not require additional assumptions on discretization. Moreover, we have relatively few tools available to extend or generalize the Heston (1993) model. The most often used approach is to add a second and possibly a third volatility factor, and/or an additional factor with a different structure, such as a return or variance jump.[16] There are natural counterparts to these extensions that can be used in a discrete-time framework, but the discrete-time model can also be extended in other straightforward ways, for instance, by generalizing the autoregressive or moving average

long-term dividend strips. See for instance Binsbergen, Brandt, and Koijen (2012), Binsbergen and Koijen (2017), and Gormsen (2021).

[16] Another alternative is a "rough" volatility factor with long memory, see for instance Bayer, Friz, and Gatheral (2016), Euch and Rosenbaum (2018), and Garnier and Sølna (2018).

structure of the variance dynamic as in Section 2.5.

We now compare the option pricing performance and economic implications of the one-factor discrete-time and continuous-time option pricing models. While both models have an affine structure, we found the comparison is more complex than anticipated. Column (4) in Table 2 presents physical and risk-neutral parameter estimates for the Heston (1993) square root model, as well as the log-likelihood and vega-weighted RMSE. The model parameter estimates are consistent with the literature. The log likelihood and RMSE can be directly compared to the discrete-time SV results in column (3). Both indicate that the continuous-time model slightly underperforms the discrete-time SV model. We found that most of this difference results from imposing (binding) positivity restrictions on the variance processes. These restrictions are by design very different in these two models. In the discrete-time model, we impose positivity of the variance by setting $\omega = \omega^* = 0$. In the continuous-time model, we impose the Feller condition under both the P and Q measures, i.e., $2\kappa\theta = 2\kappa^*\theta^* > \sigma^2$. If we do not impose these positivity restrictions, the RMSEs improve modestly in both models and are roughly equal. We therefore conclude that from the perspective of option fit, the performance of both modeling approaches is rather similar, but that the positivity restrictions are critical.[17]

The similarities between the discrete- and continuous-time SV models are confirmed by the model-implied physical and risk-neutral properties in Panel A of Table 3, as well as the kernels in Figures 1 and 2. However, it is worth mentioning that both the bivariate log pricing kernel and the time path of the realized log pricing kernel for the continuous-time model in Panel D of both figures exhibit more variability than the ones in Panel C. This difference can be interpreted through the lens of the implied risk preference parameters. In Column (4) of Table 3, the equity level aversion parameter γ is 0.704, and the variance aversion parameter ξ is 1,544.21. The former (latter) is

[17]In addition to comparing the option fit, we also examine the model-implied variance and return distributions, as shown in Figures 7 and 8. To see the impact of the positivity restrictions, we calculate non-optimal model parameters by matching moment conditions between the discrete-time and continuous-time models and relaxing the positivity restrictions. For example, "CT SV (Moments Matched)" represents the continuous-time SV model with parameters determined by matching the mean, variance, and autocorrelation of variance, as well as the correlation between return and variance, calculated from the optimal discrete-time SV model. We confirm that the dynamics of the moments-matched discrete-time and continuous-time models closely resemble those of their optimal counterparts.

smaller (larger) than the one in column (3) for the discrete-time SV model. The relatively greater variability of the continuous-time pricing kernel is therefore due to the higher ξ.

As for risk premiums, the time paths of the conditional 1-month equity and variance risk premiums in Figure 3 are similar for the two SV models, but the 1-month ERP is higher in the discrete-time model during crises. Figure 5 indicates that the inverse slope of the equity premium term structure during crises is higher in the continuous-time model, but the time series of the slopes in the two models are very highly correlated. Although the correlation between the time series of the VRP slopes for the two models is also high, when volatility is high the slopes are significantly larger in the discrete-time model. The CHJ GARCH model implies a level of the VRP premium (in Figure 3) that is similar to the SV model, but it generates a very different VRP term structure.

5.2 Where Does the Performance Difference Come From?

As briefly mentioned in Section 5.1, the different positivity restrictions in the discrete time and continuous time models are the primary cause of the variations in pricing performance. We now further investigate the intrinsic source of these performance differences by analyzing subsets of options data alongside the model-implied forward variance and return distributions.

Table 5 reports the vega-weighted RMSEs and the IV RMSEs, categorized into buckets based on the level of conditional variance and moneyness (or maturity) for the two types of SV models. Across all panels, we observe that the discrete-time SV model consistently and substantially outperforms the continuous-time SV model in low conditional variance buckets (below the 10th percentile of its sample distribution). Figure 7 illustrates one-month and one-year forward physical (P) and risk-neutral (Q) variance distributions derived from simulating the model dynamics. When the current conditional variance is low, the discrete-time model produces forward variances that are more right-skewed and concentrated at a lower range than those of the continuous-time model.[18] In other words, low variance in the discrete-time model reverts to its mean more slowly, as further supported by a higher variance autocorrelation coefficient in Table 3. This finding suggests that

[18] In this figure, the low conditional variance corresponds to the 2.5th percentile of its time series distribution, while the high conditional variance corresponds to the 97.5th percentile.

the improvement in option fit arises from the discrete-time model's better capture of variance's mean-reversion speed when the current variance is at low levels.

Furthermore, examining pricing errors in more granular buckets reveals that the smaller RMSEs of the discrete-time model are particularly pronounced for long-term options and OTM call options (moneyness less than one). In Panel D of Figure 8, which depicts the one-year forward risk-neutral return distribution at low conditional variance, the continuous-time model exhibits a much ticker tail of positive returns compared to the discrete-time model. This thicker tail increases the risk-neutral likelihood of positive returns and thus inflates OTM call prices, as shown in Figure 9. This property is the main cause of the larger errors in the continuous-time model. Conversely, as seen in Panels A and B of Table 5, the discrete-time model exhibits slightly higher RMSEs for OTM call options when the conditional variance is high. Together with Panel H of Figure 8, this result implies that long-horizon positive returns are marginally overstated in the discrete-time model under high conditional variance.

Finally, the discrete-time model significantly outperforms the continuous-time model in pricing OTM puts (moneyness greater than one), regardless of the conditional variance level, as shown in Panels A and B of Table 5. When conditional variance is low, this result is primarily due to the continuous-time model's relative overstatement of negative returns (Panel D of Figure 8). When conditional variance is high, the result arises from the continuous-time model's understatement of extreme negative returns (Panel H). Figure 9 corroborates these findings, presenting the outcomes in terms of implied volatilities under both low and high conditional variance scenarios.

6 Conclusion

We contribute to the option pricing literature by proposing a new discrete-time stochastic volatility model. The model allows for quasi closed form prices for European call options, like the Heston (1993) and Heston and Nandi (2000) models. It is closely related to the existing discrete-time option pricing literature, because it nests the Heston and Nandi (2000) model. It is also closely related to the continuous-time Heston (1993) model, because it contains a second innovation to

volatility that can be correlated with the return innovation.

We implement the model based on a joint likelihood consisting of index option data and the underlying index returns. The new model vastly outperforms the existing discrete-time benchmark in terms of option fit. Moreover, the pricing kernel implied by the new model is more plausible than the kernels implied by existing discrete-time models, and the new model generates reasonable risk premiums. The new model's improved performance is due to the fact that when incorporating variance aversion into the pricing kernel in a GARCH model, it is difficult to separately identify equity and variance risk. The new model also outperforms the continuous-time benchmark, but the mechanics of both models are similar and the difference in performance mainly results from the different constraints that need to be imposed in both models to keep the variance positive.

In future work, we plan to investigate how option fit can be improved by including additional volatility factors, and to study how the performance of the resulting model compares with that of a model that includes additional lags in the volatility dynamic.

References

Aït-Sahalia, Yacine, Mustafa Karaman, and Loriano Mancini, 2020, The Term Structure of Equity and Variance Risk Premia, *Journal of Econometrics*.

Aït-Sahalia, Yacine, and Robert Kimmel, 2007, Maximum Likelihood Estimation of Stochastic Volatility Models, *Journal of Financial Economics* 83, 413–452.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015a, Parametric Inference and Dynamic State Recovery from Option Panels, *Econometrica* 83, 1081–1145.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2015b, The Risk Premia Embedded in Index Options, *Journal of Financial Economics* 117, 558–584.

Andersen, Torben G., Nicola Fusari, and Viktor Todorov, 2017, Short-Term Market Risks Implied by Weekly Options, *Journal of Finance* 72, 1335–1386.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical Performance of Alternative Option Pricing Models, *Journal of Finance* 52, 2003–2049.

Bardgett, Chris, Elise Gourier, and Markus Leippold, 2019, Inferring Volatility Dynamics and Risk Premia from the S&P 500 and VIX Markets, *Journal of Financial Economics* 131, 593–618.

Bates, David S., 2000, Post-'87 Crash Fears in the S&P 500 Futures Option Market, *Journal of Econometrics* 94, 181–238.

Bates, David S., 2003, Empirical Option Pricing: A Retrospection, *Journal of Econometrics* 116, 387–404.

Bates, David S., 2006, Maximum Likelihood Estimation of Latent Affine Processes, *Review of Financial Studies* 19, 909–965.

Bates, David S., 2019, How Crashes Develop: Intradaily Volatility and Crash Evolution, *Journal of Finance* 74, 193–238.

Bates, David S., 2022, Empirical Option Pricing Models, *Annual Review of Financial Economics* 14, 369–389.

Bayer, Christian, Peter Friz, and Jim Gatheral, 2016, Pricing Under Rough Volatility, *Quantitative Finance* 16, 887–904.

Binsbergen, Jules H. van, Michael W. Brandt, and Ralph S. Koijen, 2012, On the Timing and Pricing of Dividend, *American Economic Review* 102, 1596–1618.

Binsbergen, Jules H. van, and Ralph S. J. Koijen, 2017, The Term Structure of Returns: Facts and Theory, *Journal of Financial Economics* 124, 1–21.

Black, Fischer, and Myron Scholes, 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637–654.

Bollerslev, Tim, 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31, 307–327.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model Specification and Risk Premia: Evidence from Futures Options, *Journal of Finance* 62, 1453–1490.

Carr, Peter, and Dilip Madan, 1999, Option Valuation Using the Fast Fourier Transform, *Journal of Computational Finance* 2, 61–73.

Carr, Peter, and Liuren Wu, 2007, Stochastic Skew in Currency Options, *Journal of Financial Economics* 86, 213–247.

Chernov, Mikhail, 2003, Empirical Reverse Engineering of the Pricing Kernel, *Journal of Econometrics* 116, 329–364.

Chernov, Mikhail, and Eric Ghysels, 2000, A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation, *Journal of Financial Economics* 56, 407–458.

Chernov, Mikhail, Jeremy Graveline, and Irina Zviadadze, 2018, Crash Risk in Currency Returns, *Journal of Financial and Quantitative Analysis* 53, 137–170.

Cheung, Sam, 2008, An Empirical Analysis of Joint Time-Series of Returns and the Term-Structure of Option Prices, *Working Paper, Columbia University.*

Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2013, Capturing Option Anomalies with a Variance-Dependent Pricing Kernel, *Review of Financial Studies* 26, 1963–2006.

Christoffersen, Peter, and Kris Jacobs, 2004, Which GARCH Model for Option Valuation?, *Management Science* 50, 1204–1221.

Christoffersen, Peter, Kris Jacobs, and Karim Mimouni, 2010, Volatility Dynamics for the S&P500: Evidence from Realized Volatility, Daily Returns, and Option Prices, *Review of Financial Studies* 23, 3141–3189.

Christoffersen, Peter, Kris Jacobs, and Chayawat Ornthanalai, 2013, GARCH Option Valuation: Theory and Evidence, *Journal of Derivatives* 21, 8–41.

Duan, Jin-Chuan, 1995, The GARCH Option Pricing Model, *Mathematical Finance* 5, 13–32.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* 68, 1343–1376.

Engle, Robert F., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50, 987–1007.

Eraker, Bjørn, 2004, Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, *Journal of Finance* 59, 1367–1404.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The Impact of Jumps in Volatility and Returns, *Journal of Finance* 58, 1269–1300.

Euch, Omar El, and Mathieu Rosenbaum, 2018, Perfect Hedging in Rough Heston Models, *The Annals of Applied Probability* 28, 3813–3856.

Feller, William, 1951, Two Singular Diffusion Problems, *The Annals of Mathematics* 54, 173–182.

Garnier, Josselin, and Knut Sølna, 2018, Option Pricing Under Fast-Varying and Rough Stochastic Volatility, *Annals of Finance* 14, 489–516.

Ghosh, Anisha, Christian Julliard, and Alex P. Taylor, 2017, What is the Consumption-CAPM Missing? An Information-Theoretic Framework for the Analysis of Asset Pricing Models, *Review of Financial Studies* 30, 442–504.

Gormsen, Niels Joachim, 2021, Time Variation of the Equity Term Structure, *Journal of Finance* 76, 1959–1999.

Heston, Steven, Kris Jacobs, and Hyung Joo Kim, 2023, The Pricing Kernel in Options, *Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System* 2023-053.

Heston, Steven L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 327–343.

Heston, Steven L., 1997, A Simple New Formula for Options with Stochastic Volatility, *Working Paper, University of Maryland.*

Heston, Steven L., and Saikat Nandi, 2000, A Closed-Form GARCH Option Valuation Model, *Review of Financial Studies* 13, 585–625.

Hurn, A. Stan, Kenneth A. Lindsay, and Andrew J. McClelland, 2015, Estimating the Parameters of Stochastic Volatility Models Using Option Price Data, *Journal of Business & Economic Statistics* 33, 579–594.

Jones, Christopher S., 2003, The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, *Journal of Econometrics* 116, 181–224.

Lewis, A. L., 2000, *Option Valuation Under Stochastic Volatility*. (Finance Press).

Li, Junye, and Gabriele Zinna, 2018, The Variance Risk Premium: Components, Term Structures, and Stock Return Predictability, *Journal of Business & Economic Statistics* 36, 411–425.

Pan, Jun, 2002, The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study, *Journal of Financial Economics* 63, 3–50.

Scott, D. W., 1992, *Multivariate Density Estimation: Theory, Practice, and Visualization*. (Wiley).

Trolle, Anders B., and Eduardo S. Schwartz, 2009, Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives, *Review of Financial Studies* 22, 4423–4461.

Table 1: Descriptive Statistics

Panel A: Option Data by Moneyness							
	$F/K \leq 0.94$	$0.94 < F/K \leq 0.98$	$0.98 < F/K \leq 1.02$	$1.02 < F/K \leq 1.06$	$1.06 < F/K \leq 1.10$	$F/K > 1.10$	All
Number of contracts	5,640	5,127	9,488	6,676	4,914	8,381	40,226
Average IV (%)	17.65	16.35	17.71	20.05	22.33	26.02	20.21
Average price	21.21	34.27	55.13	39.48	32.97	27.58	36.67
Average spread	1.64	1.63	1.84	1.64	1.53	1.49	1.64
Panel B: Option Data by Maturity							
	DTM ≤ 30	$30 <$ DTM ≤ 60	$60 <$ DTM ≤ 90	$90 <$ DTM ≤ 120	$120 <$ DTM ≤ 180	DTM > 180	All
Number of contracts	4,158	8,774	5,385	2,490	5,491	13,928	40,226
Average IV (%)	20.43	20.46	20.35	20.64	20.28	19.83	20.21
Average price	11.46	17.56	26.73	33.03	37.07	60.57	36.67
Average spread	0.76	1.13	1.40	1.47	1.65	2.34	1.64
Panel C: Stock Returns (Annualized)							
	1990-2021	1996-2021					
Mean (%)	10.22	9.76					
Standard deviation (%)	18.11	19.33					
Skewness	-0.41	-0.41					
Kurtosis	14.37	13.49					

Notes: Panels A and B present descriptive statistics for Wednesday closing OTM option contracts for the January 1996 to June 2021 period. Moneyness is defined as the implied futures price $F = Se^{(r-q)\tau}$ divided by the strike price K. Panel C reports on the log of daily index total returns for the January 1, 1990 to December 31, 2021 and January 1, 1996 to December 31, 2021 periods. Mean and standard deviation are annualized, skewness and kurtosis are computed from daily returns.

Table 2: Parameter Estimates

	HN GARCH	CHJ GARCH	DT SV		CT SV
	(1)	(2)	(3)		(4)
P-parameters					
μ	2.8785	2.8785	2.8785	μ	2.8785
$\omega \times 10^4$	0	0	0		
β	0.6668	0.6751	0.0125	κ	0.0133
	(0.0005)	(0.0005)	(0.0003)		(0.0002)
$\alpha_1 \times 10^2$	0.1305	0.1204	0.0850	$\theta \times 10^2$	0.0129
	(0.0037)	(0.0049)	(0.0012)		(0.0002)
$\alpha_2 \times 10^2$	0	0	0.0667	σ	0.0019
			(0.0009)		(0.0001)
λ	0.5664	0.5594	0.9893	ρ	-0.7609
	(0.0005)	(0.0011)	(0.0002)		(0.0078)
η_1			0.7717		0.7321
			(0.0035)		(0.0035)
Q-parameters					
$\omega^* \times 10^4$	0	0	0	κ^*	0.0070
β^*	0.6668	0.6751	0.0125	$\theta^* \times 10^2$	0.0244
$\alpha_1^* \times 10^2$	0.1305	0.1305	0.0851		
$\alpha_2^* \times 10^2$	0	0	0.0667		
λ^*	0.5702	0.5629	0.9921		
P-Q wedge					
ζ	1	1.0837	1.0005	ξ	1544.21
		(0.0235)	(0.0002)		(149.38)
Model fit					
$\log L^R$	26,312.18	26,329.24	103,460.60		102,971.15
$\log L_r^R$	26,312.18	26,329.24	26,529.77		26,573.26
$\log L_v^R$			76,930.83		76,397.89
$\log L^O$	75,332.45	75,305.84	94,561.60		89,224.82
Weighted $\log L$	124,007.45	124,042.57	366,568.68		361,899.78
Vega-weighted RMSE	0.0372	0.0372	0.0231		0.0263

Notes: We report parameter estimates for the HN GARCH, CHJ GARCH, discrete-time (DT) SV, and continuous-time (CT) SV models. All models are estimated using the joint likelihood of underlying returns and option prices. Robust standard errors are in parentheses. The option sample period is from January 1996 to December 2021; the return sample starts in January 1990. $\log L^R = \log L_r^R + \log L_v^R$, where $\log L_r^R$ and $\log L_v^R$ account for the transition probabilities of daily returns and variances, respectively.

Table 3: Model Properties

	HN GARCH	CHJ GARCH	DT SV	CT SV
	(1)	(2)	(3)	(4)
Panel A: Preference Parameters and Model-Implied Moments				
Risk preference parameters				
γ	2.8785	-33.1800	2.2919	0.7043
ξ	0	26,625.97	347.95	1,544.21
Physical properties				
Annualized expected log return	0.1074	0.0972	0.1039	0.1017
Annualized long-run variance	0.0349	0.0306	0.0334	0.0325
Annualized long-run volatility of variance	0.4490	0.3890	0.7458	0.5156
Autocorrelation of daily variance	0.9877	0.9881	0.9912	0.9867
Conditional correlation $(\log R(t+1), v(t+1))$	-0.9905	-0.9906	-0.7851	-0.7609
Risk-neutral properties				
Annualized long-run variance	0.0534	0.0536	0.0882	0.0616
Annualized long-run volatility of variance	0.6890	0.6835	1.9670	0.9777
Autocorrelation of daily variance	0.9920	0.9920	0.9967	0.9930
Conditional correlation $(\log R(t+1), v^*(t+1))$	-0.9939	-0.9937	-0.7863	-0.7609
Panel B: Integrated Means, Variances, and Risk Premiums (One-Month Horizon)				
Annualized conditional physical moments				
Mean	0.1188	0.1111	0.1223	0.1176
Variance	0.0330	0.0302	0.0342	0.0325
Annualized conditional risk-neutral moments				
Mean	0.0245	0.0245	0.0245	0.0245
Variance	0.0344	0.0344	0.0362	0.0345
Annualized conditional risk premia				
Equity premium	0.0943	0.0866	0.0978	0.0931
Variance premium	-0.0015	-0.0042	-0.0020	-0.0020

Notes: Panel A reports model properties for the HN GARCH, CHJ GARCH, discrete-time (DT) SV, and continuous-time (CT) SV models, under both the physical and risk-neutral measures. Panel B reports descriptive statistics for the model-implied mean and variance of returns, as well as the equity and variance risk premiums.

Table 4: Model Pricing Errors (RMSEs) by Moneyness and Maturity

Panel A: Vega-Weighted RMSE by Moneyness							
	$F/K \leq 0.94$	$0.94 < F/K \leq 0.98$	$0.98 < F/K \leq 1.02$	$1.02 < F/K \leq 1.06$	$1.06 < F/K \leq 1.10$	$F/K > 1.10$	All
HN GARCH	0.0349	0.0384	0.0349	0.0347	0.0363	0.0424	0.0372
DT SV	0.0320	0.0209	0.0182	0.0168	0.0197	0.0278	0.0231
CT SV	0.0308	0.0240	0.0208	0.0209	0.0247	0.0338	0.0263
Panel B: IV RMSE by Moneyness							
	$F/K \leq 0.94$	$0.94 < F/K \leq 0.98$	$0.98 < F/K \leq 1.02$	$1.02 < F/K \leq 1.06$	$1.06 < F/K \leq 1.10$	$F/K > 1.10$	All
HN GARCH	0.0389	0.0353	0.0342	0.0369	0.0419	0.0567	0.0419
DT SV	0.0281	0.0188	0.0172	0.0194	0.0248	0.0371	0.0255
CT SV	0.0262	0.0213	0.0203	0.0241	0.0303	0.0446	0.0295
Panel C: Vega-Weighted RMSE by Maturity							
	DTM ≤ 30	$30 <$ DTM ≤ 60	$60 <$ DTM ≤ 90	$90 <$ DTM ≤ 120	$120 <$ DTM ≤ 180	DTM > 180	All
HN GARCH	0.0443	0.0392	0.0370	0.0374	0.0354	0.0342	0.0372
DT SV	0.0271	0.0209	0.0199	0.0196	0.0237	0.0245	0.0231
CT SV	0.0285	0.0241	0.0237	0.0240	0.0262	0.0284	0.0263
Panel D: IV RMSE by Maturity							
	DTM ≤ 30	$30 <$ DTM ≤ 60	$60 <$ DTM ≤ 90	$90 <$ DTM ≤ 120	$120 <$ DTM ≤ 180	DTM > 180	All
HN GARCH	0.0573	0.0459	0.0413	0.0412	0.0378	0.0353	0.0419
DT SV	0.0375	0.0252	0.0216	0.0206	0.0237	0.0241	0.0255
CT SV	0.0419	0.0302	0.0265	0.0257	0.0264	0.0276	0.0295

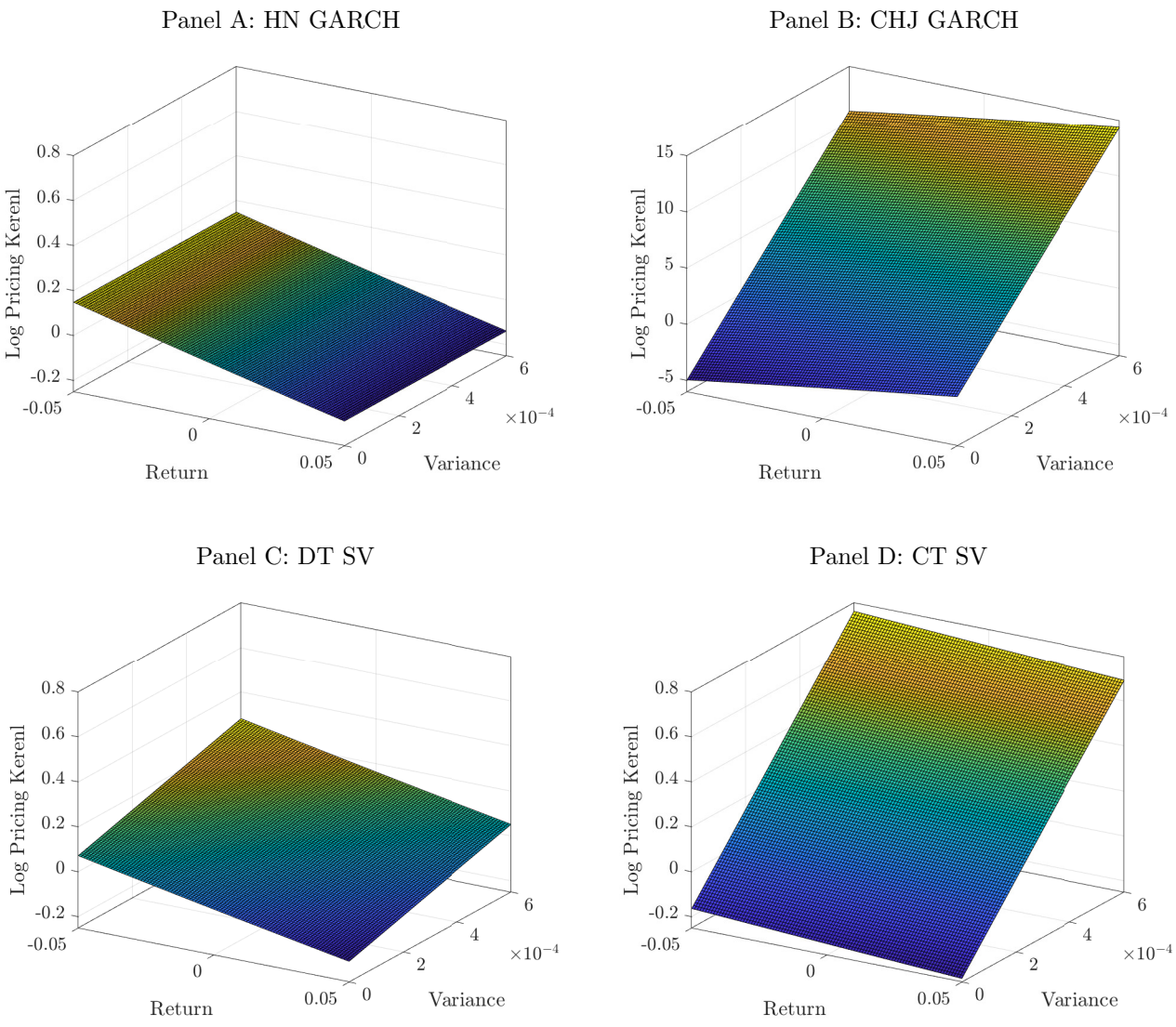
Notes: We report model pricing errors (RMSEs) by moneyness and maturity for the GARCH, discrete-time (DT) SV, and continuous-time (CT) SV models. Moneyness is defined as the implied futures price $F = Se^{(r-q)\tau}$ divided by the strike price K. The sample period is from January 1996 to December 2021.

Table 5: Model Pricing Errors (RMSEs) by Conditional Variance

Panel A: Vega-Weighted RMSE by Conditional Variance and Moneyness							
	$v(t) \leq v_{10},\ m_f \leq 1$	$v(t) \leq v_{10},\ m_f > 1$	$v_{10} < v(t) \leq v_{90},\ m_f \leq 1$	$v_{10} < v(t) \leq v_{90},\ m_f > 1$	$v(t) > v_{90},\ m_f \leq 1$	$v(t) > v_{90},\ m_f > 1$	All
DT SV	0.0155	0.0174	0.0189	0.0192	0.0500	0.0373	0.0231
CT SV	0.0331	0.0236	0.0217	0.0229	0.0408	0.0474	0.0263
Panel B: IV RMSE by Conditional Variance and Moneyness							
	$v(t) \leq v_{10},\ m_f \leq 1$	$v(t) \leq v_{10},\ m_f > 1$	$v_{10} < v(t) \leq v_{90},\ m_f \leq 1$	$v_{10} < v(t) \leq v_{90},\ m_f > 1$	$v(t) > v_{90},\ m_f \leq 1$	$v(t) > v_{90},\ m_f > 1$	All
DT SV	0.0132	0.0172	0.0184	0.0242	0.0405	0.0488	0.0255
CT SV	0.0282	0.0232	0.0193	0.0288	0.0354	0.0606	0.0295
Panel C: Vega-Weighted RMSE by Conditional Variance and Maturity							
	$v(t) \leq v_{10},\ \tau \leq 90$	$v(t) \leq v_{10},\ \tau > 90$	$v_{10} < v(t) \leq v_{90},\ \tau \leq 90$	$v_{10} < v(t) \leq v_{90},\ \tau > 90$	$v(t) > v_{90},\ \tau \leq 90$	$v(t) > v_{90},\ \tau > 90$	All
DT SV	0.0140	0.0188	0.0179	0.0200	0.0435	0.0435	0.0231
CT SV	0.0194	0.0318	0.0208	0.0237	0.0465	0.0429	0.0263
Panel D: IV RMSE by Conditional Variance and Maturity							
	$v(t) \leq v_{10},\ \tau \leq 90$	$v(t) \leq v_{10},\ \tau > 90$	$v_{10} < v(t) \leq v_{90},\ \tau \leq 90$	$v_{10} < v(t) \leq v_{90},\ \tau > 90$	$v(t) > v_{90},\ \tau \leq 90$	$v(t) > v_{90},\ \tau > 90$	All
DT SV	0.0145	0.0172	0.0237	0.0209	0.0502	0.0405	0.0255
CT SV	0.0188	0.0289	0.0277	0.0238	0.0586	0.0429	0.0295

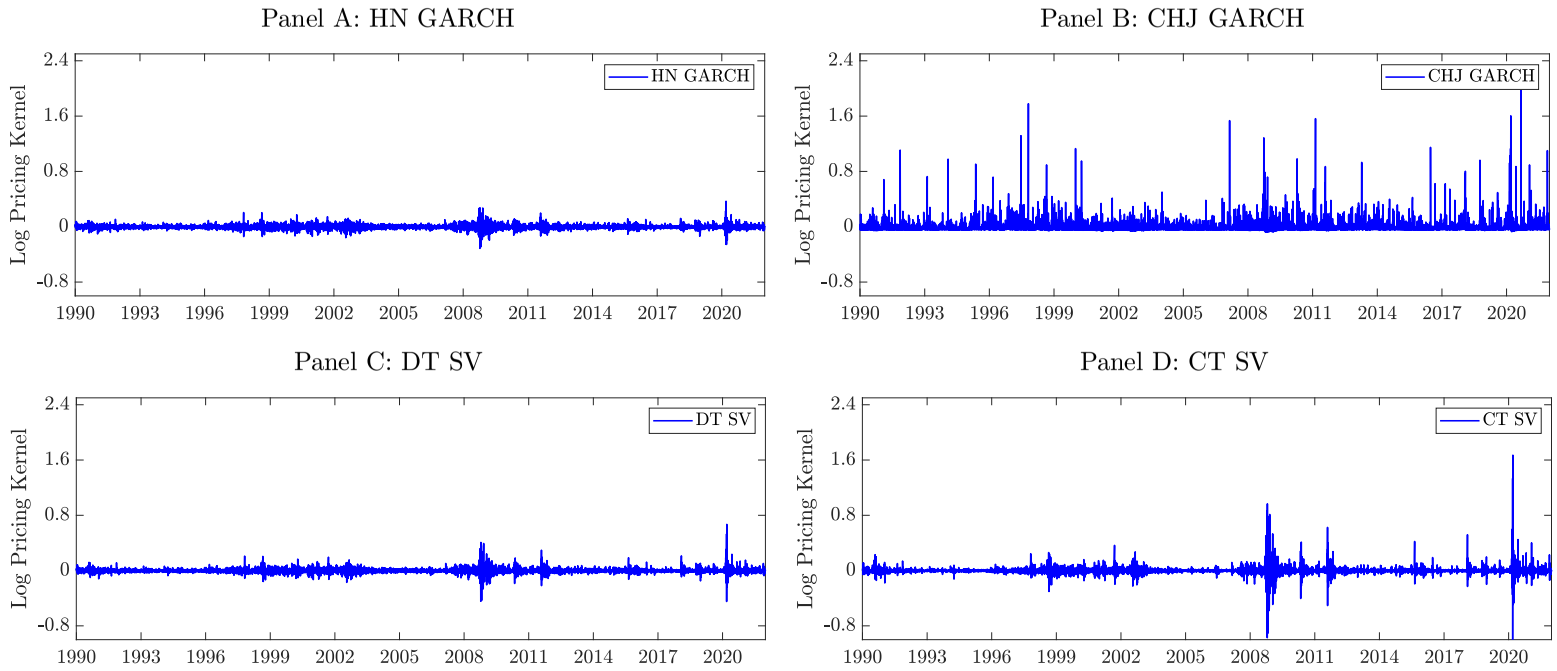
Notes: We report model pricing errors (RMSEs) by conditional variance and moneyness (m_f) or maturity (τ) for the discrete-time (DT) SV, and continuous-time (CT) SV models. Moneyness is defined as the implied futures price $F = Se^{(r-q)\tau}$ divided by the strike price K. The sample period is from January 1996 to December 2021.

Figure 1: Bivariate Daily Log Pricing Kernels



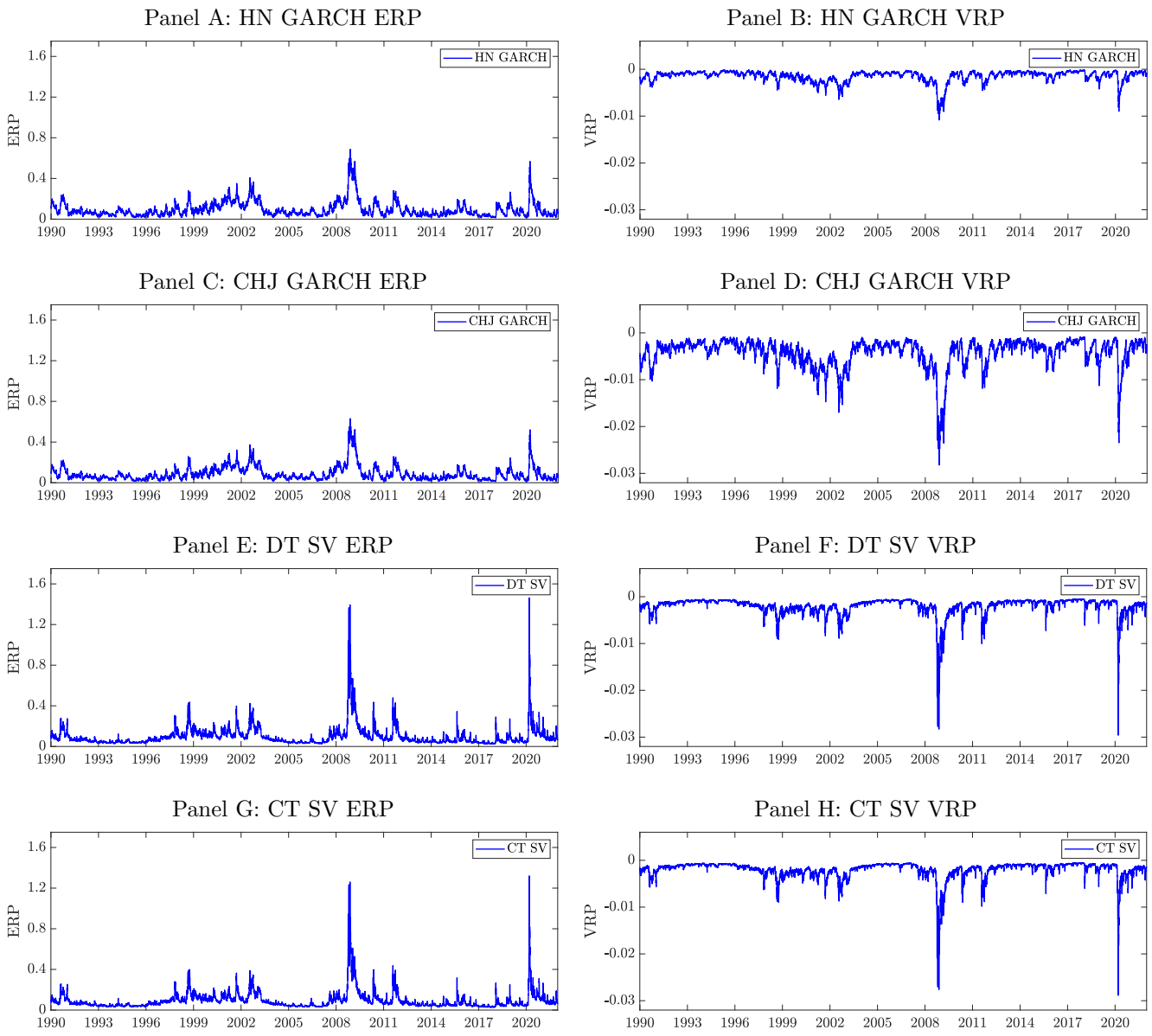
Notes: We plot the log pricing kernel as a function of the index return and variance for four models. The calculations are based on the parameters in Table 2. In the HN GARCH model, the kernel does not depend on the variance.

Figure 2: Time Series of Realized Daily Log Pricing Kernel



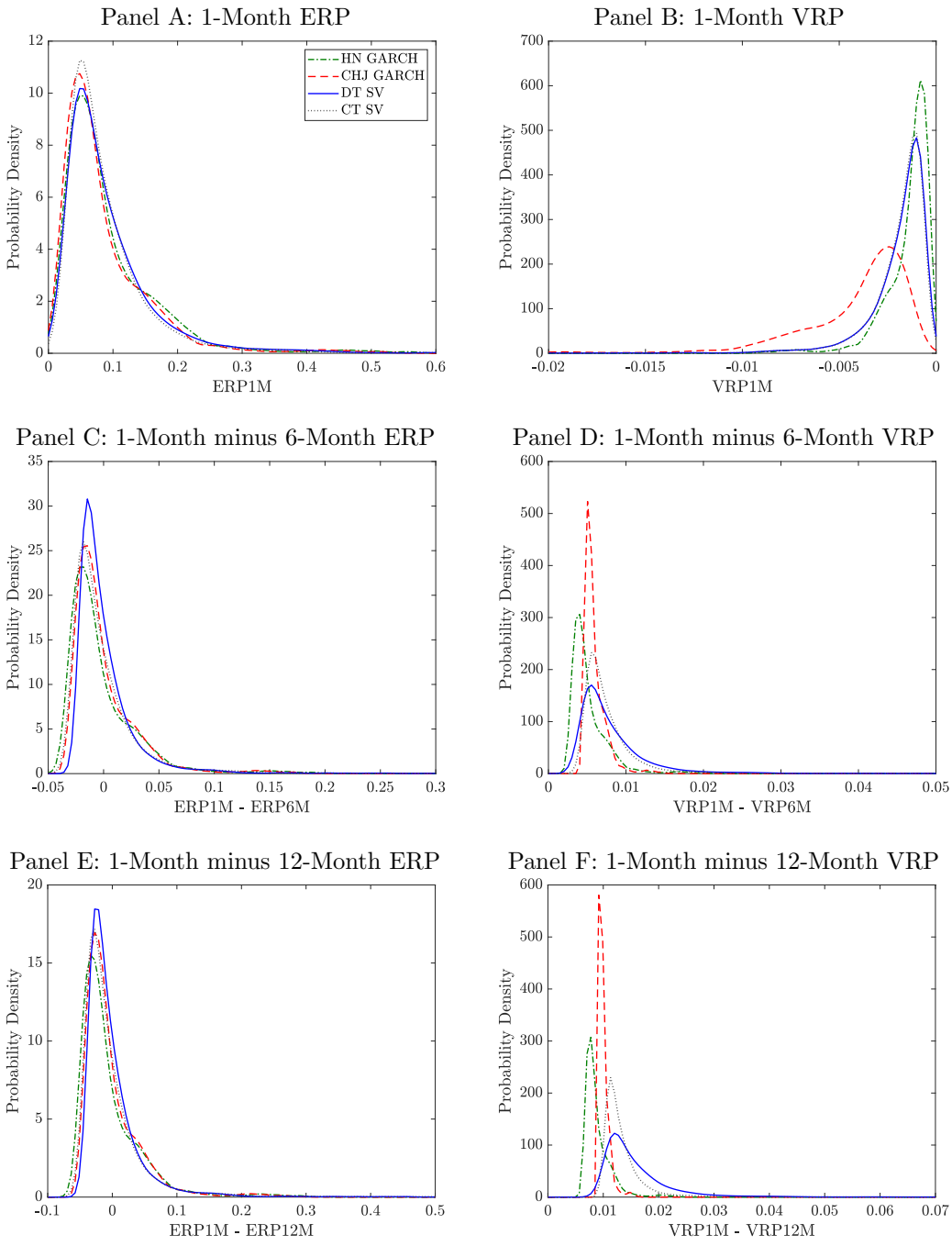
Notes: We plot the time series of the daily log pricing kernel for four models. The calculations are based on the parameters in Table 2.

Figure 3: Time Series of the One-Month Conditional Equity and Variance Risk Premiums



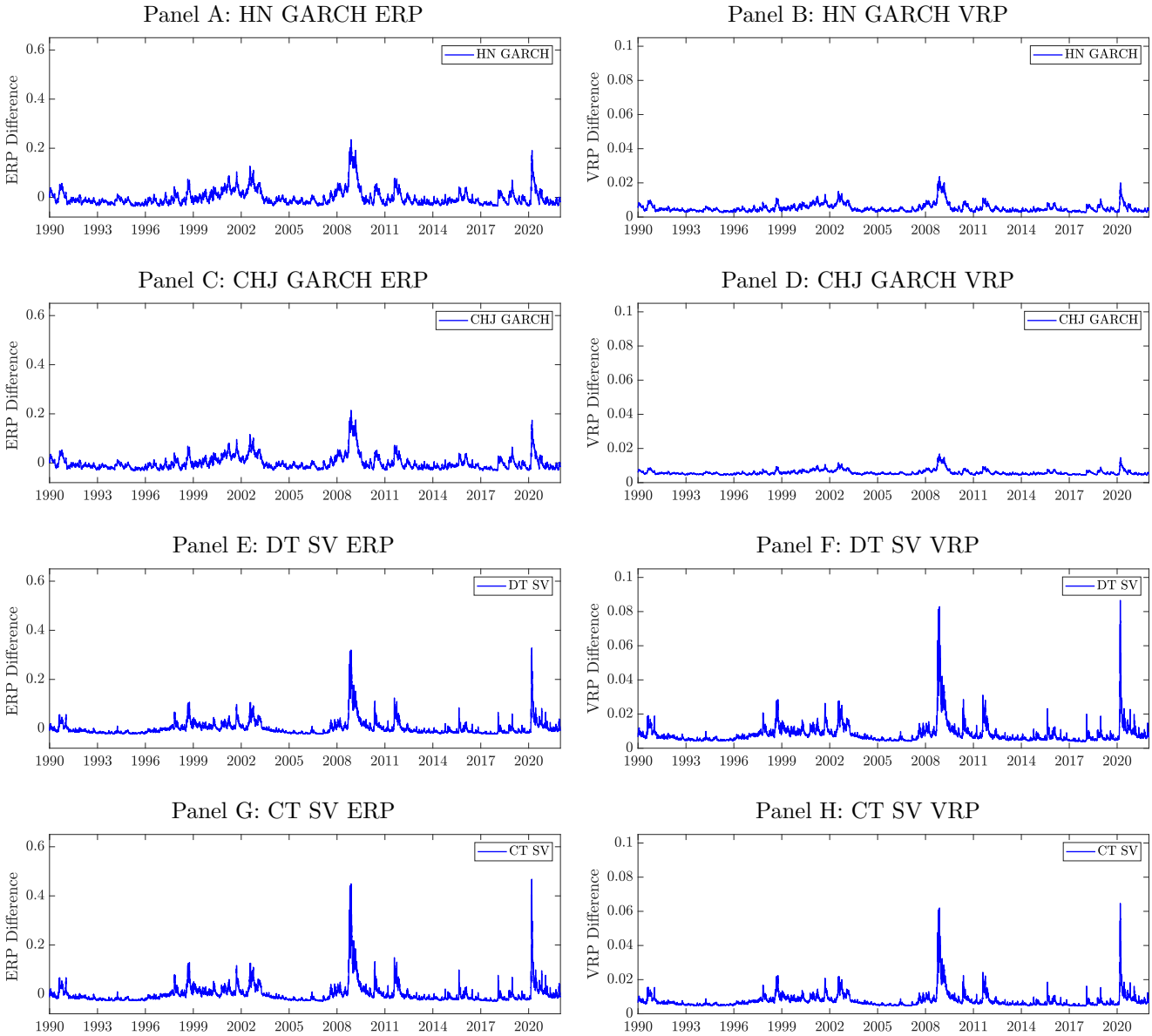
Notes: We plot the annualized time series of the one-month equity risk premium (ERP) and variance risk premium (VRP) for four models. The calculations are based on the parameters in Table 2.

Figure 4: Distribution of the Level and Term Structure of the Risk Premiums



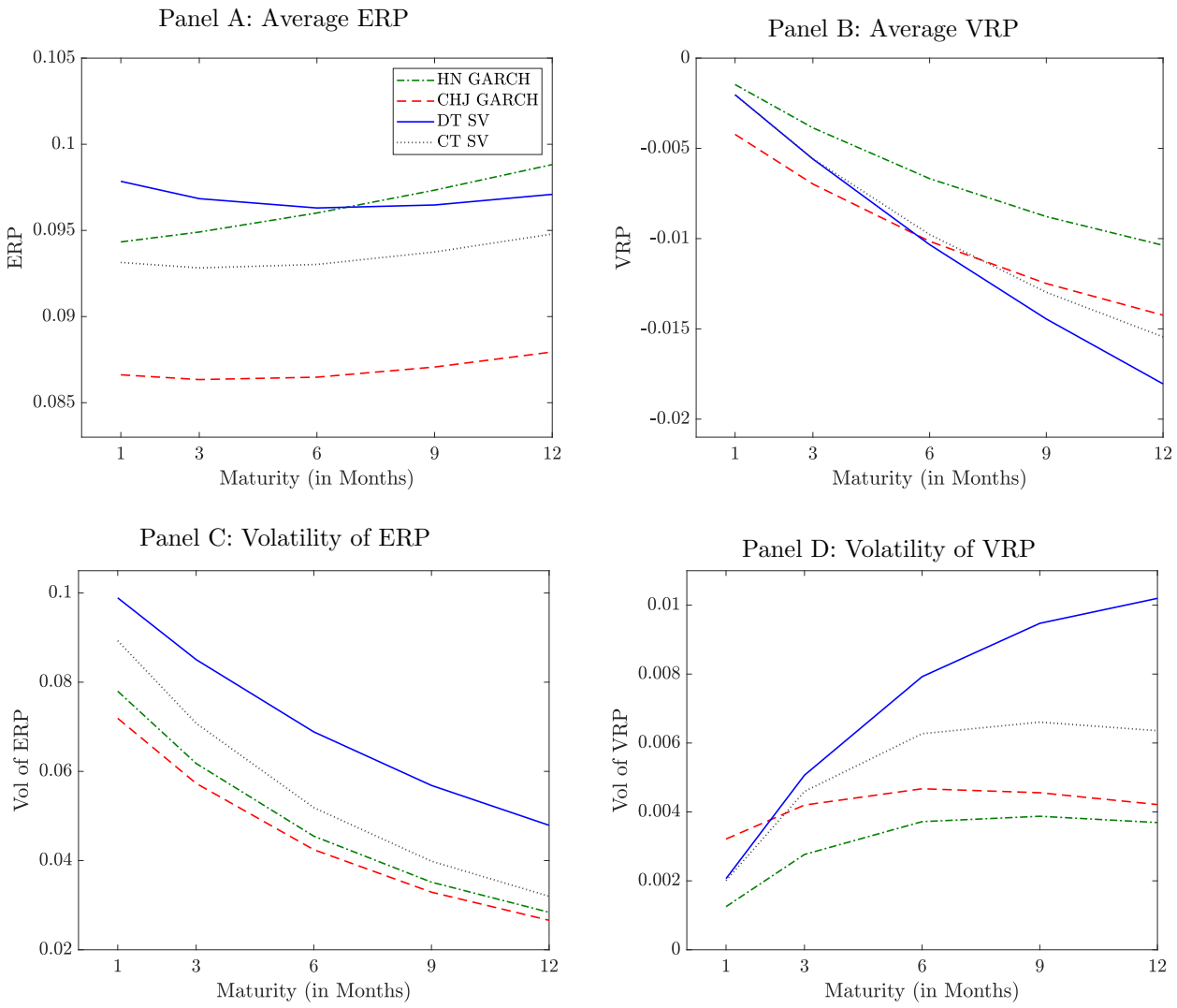
Notes: We plot the empirical distributions of the annualized one-month equity and variance risk premiums as well as the difference between the one-month and either the six-month or the twelve-month risk premiums. The empirical distributions are computed using kernel density estimation with the Gaussian kernel and Scott's (1992) rule-of-thumb bandwidth.

Figure 5: Time Series of One-Month Minus Six-Month Equity and Variance Risk Premiums



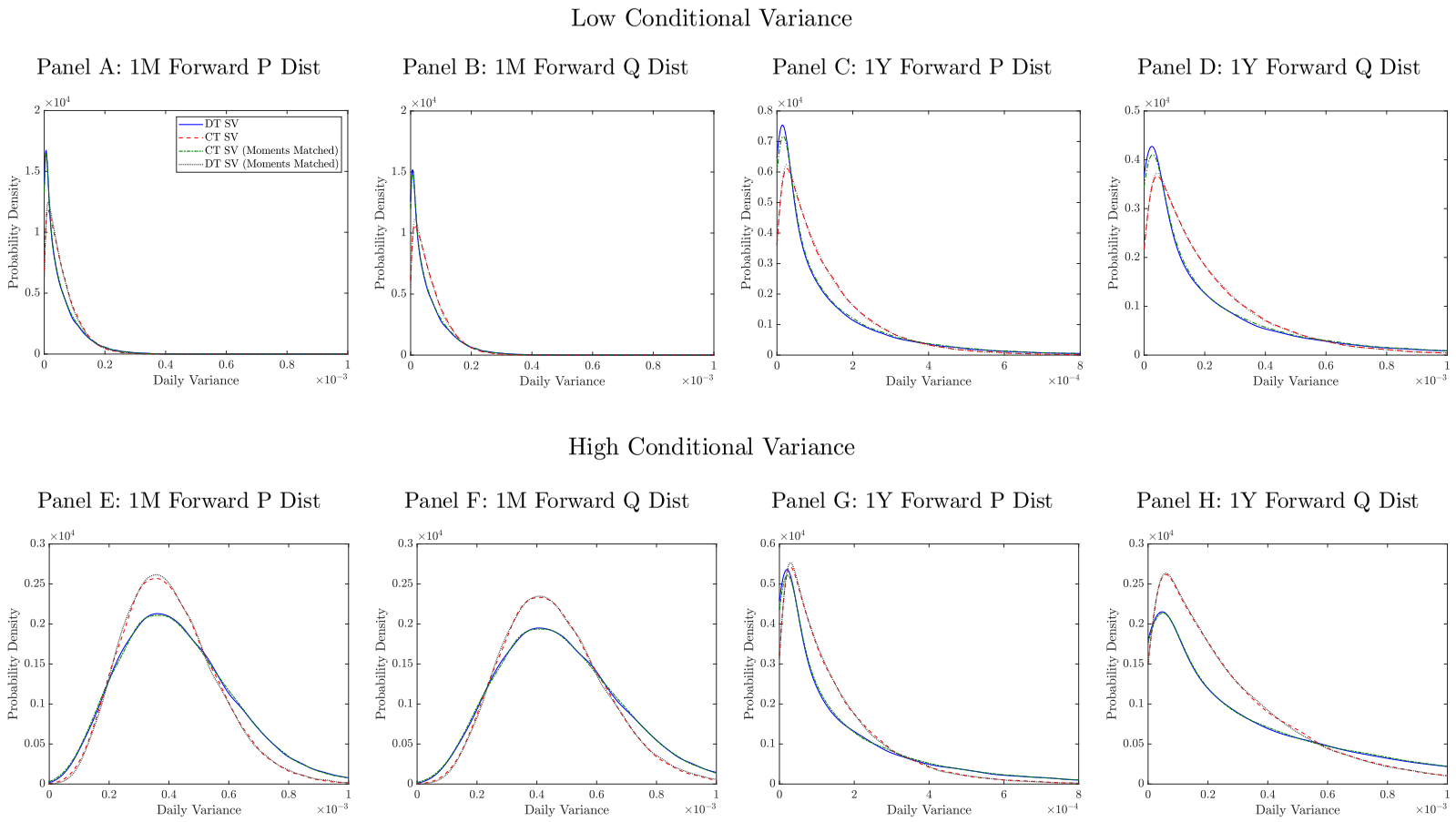
Notes: We plot the annualized time series of the (inverted) term structure of the equity risk premium (ERP) and variance risk premium (VRP) for four models. We plot the one-month ERP minus the six-month ERP, as well as the one-month VRP minus the six-month VRP. The calculations are based on the parameters in Table 2.

Figure 6: Unconditional Term Structure of Annualized Risk Premiums



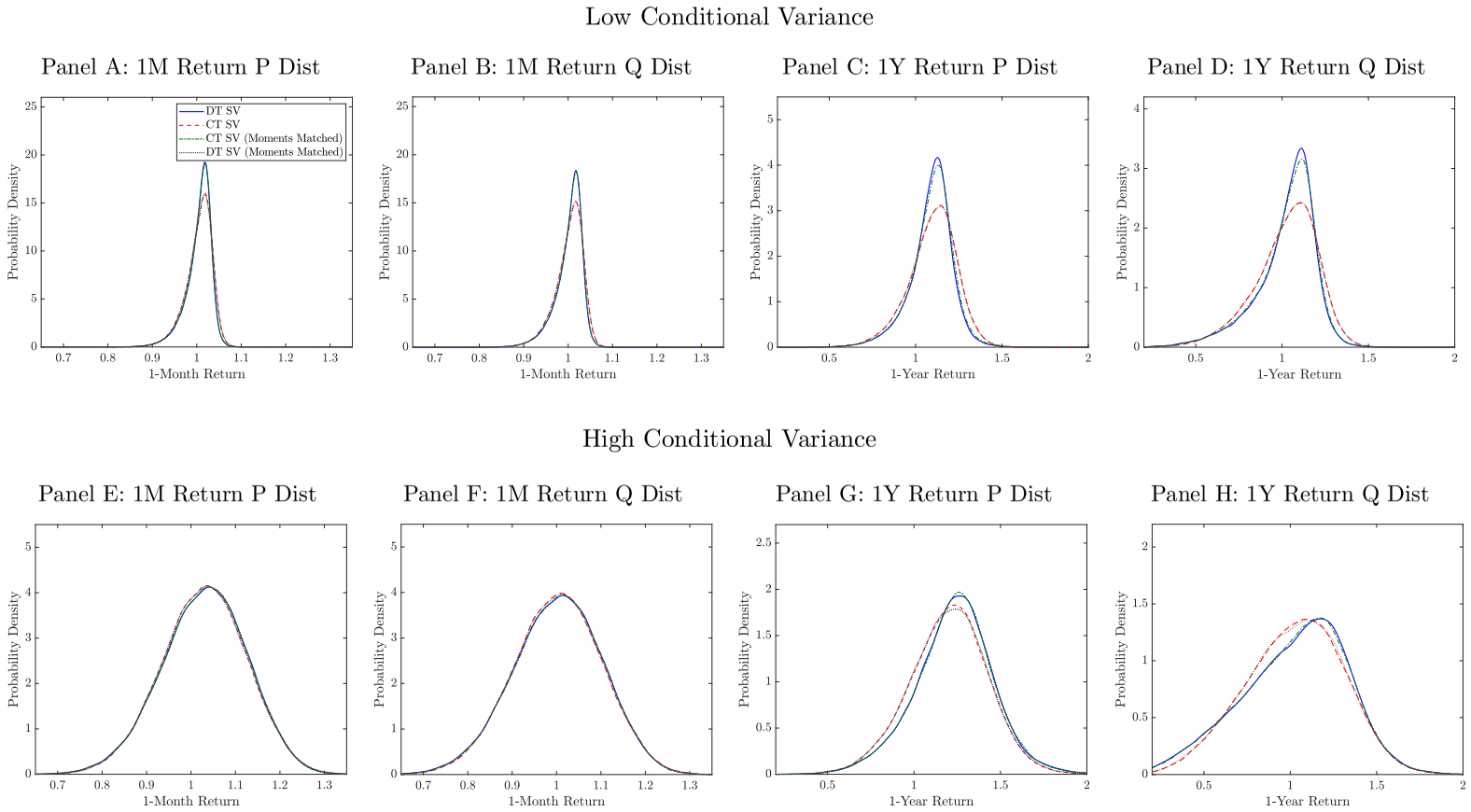
Notes: We plot the annualized average and standard deviation of the time series of the term structure of the equity and variance risk premiums for maturities between one month and one year. The calculations are based on the parameters in Table 2.

Figure 7: Daily Variance Distribution



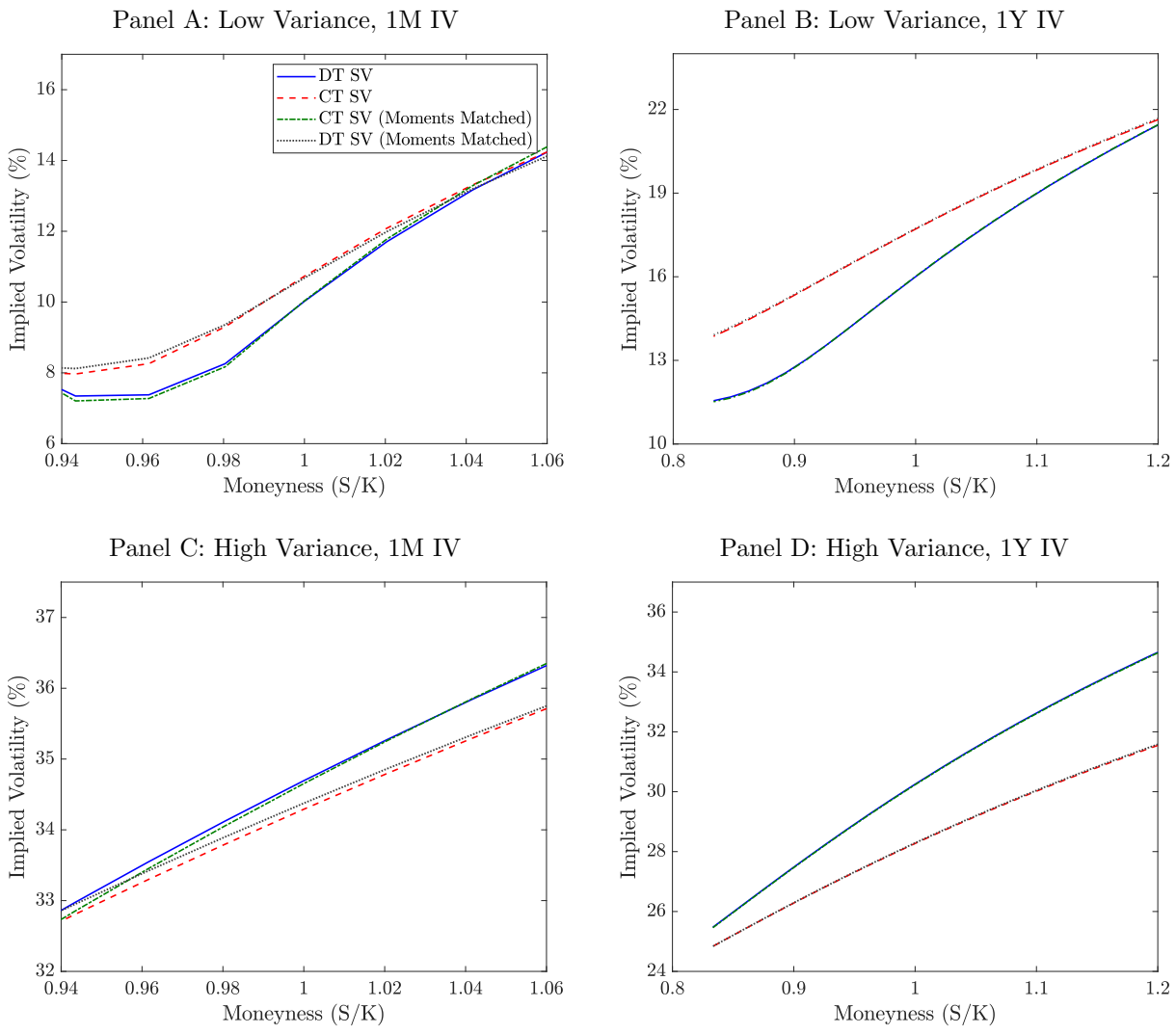
Notes: We plot the simulated physical (P) and risk-neutral (Q) distributions of the one-month and one-year forward daily variance. Each model's dynamics are simulated 50,000 times at a daily frequency over one year. The simulations for the discrete-time (DT) SV and continuous-time (CT) SV models are based on the parameters in Table 2. The parameters for CT SV (Moments Matched) and DT SV (Moments Matched) are determined by matching the mean, variance, and autocorrelation of variance, as well as the correlation between return and variance, of the optimal DT SV and CT SV models, respectively. Low (high) conditional variance corresponds to the 2.5th (97.5th) percentile of its time series distribution.

Figure 8: Forward Return Distribution



Notes: We plot the simulated physical (P) and risk-neutral (Q) distributions of the one-month and one-year forward return. Each model's dynamics are simulated 50,000 times at a daily frequency over one year. The simulations for the discrete-time (DT) SV and continuous-time (CT) SV models are based on the parameters in Table 2. The parameters for CT SV (Moments Matched) and DT SV (Moments Matched) are determined by matching the mean, variance, and autocorrelation of variance, as well as the correlation between return and variance, of the optimal DT SV and CT SV models, respectively. Low (high) conditional variance corresponds to the 2.5th (97.5th) percentile of its time series distribution.

Figure 9: Annualized Implied Volatilities



Notes: We plot the annualized one-month and one-year implied volatilities across different moneyness levels. The calculations for the discrete-time (DT) SV and continuous-time (CT) SV models are based on the parameters in Table 2. The parameters for CT SV (Moments Matched) and DT SV (Moments Matched) are determined by matching the mean, variance, and autocorrelation of variance, as well as the correlation between return and variance, of the optimal DT SV and CT SV models, respectively. Low (high) conditional variance corresponds to the 2.5th (97.5th) percentile of its time series distribution.

Appendix

A Closed-Form Expressions for the Moment Generating Function

A.1 The SV(1,1) Dynamics

Assume that the conditional moment generating function (MGF) has the following form:

$$\begin{aligned} f^{MG}(\phi) &= E_t\left[S(T)^{\phi}\right] \\ &= S(t)^{\phi}\exp\left(A(t)+B(t)v(t)\right). \end{aligned}$$

Let $x(t)=\log(S(t))$. We need to find $A(t)$ and $B(t)$ in

$$f^{MG}(\phi)=\exp\left(\phi x(t)+A(t)+B(t)v(t)\right)$$

with a terminal condition $A(T)=B(T)=0$. By the Law of Iterated Expectations,

$$\begin{aligned} f^{MG}(t;\phi) &= E_t\left[f^{MG}(t+1;\phi)\right] \\ &= E_t\left[\exp\left(\phi x(t+1)+A(t+1)+B(t+1)v(t+1)\right)\right] \end{aligned} \tag{A.1}$$

Substituting the dynamics of $x(t)$ and $v(t)$ into the above equation,

$$\begin{aligned} f^{MG}(t;\phi) &= E_t\left[\exp\left(\phi\left(x(t)+\left(r+\left(\mu-\frac{1}{2}\right)v(t)\right)+\sqrt{v(t)}z_1(t+1)\right)\right.\right. \\ &\qquad \left.\left.+A(t+1)+B(t+1)\left(\omega+\beta v(t)+\left(\alpha_1 z_1(t+1)+\alpha_2 z_2(t+1)-\lambda\sqrt{v(t)}\right)^2\right)\right)\right] \\ &= \exp\left(\phi\left(x(t)+\left(r+\left(\mu-\frac{1}{2}\right)v(t)\right)\right)+A(t+1)+B(t+1)(\omega+\beta v(t))\right) \\ &\qquad \times \underbrace{E_t\left[\exp\left(\phi\sqrt{v(t)}z_1(t+1)+B(t+1)\left(\alpha_1 z_1(t+1)+\alpha_2 z_2(t+1)-\lambda\sqrt{v(t)}\right)^2\right)\right]}_{(*)} \end{aligned} \tag{A.2}$$

where $(*)$ can be written as:

$$(*) = \frac{\exp\left[\frac{\left(-2\lambda^2 B(t+1)+4\alpha_1\lambda\phi B(t+1)+\phi^2\left(-1+2\alpha_2^2 B(t+1)\right)\right)v(t)}{-2+4\left(\alpha_1^2+\alpha_2^2\right)B(t+1)}\right]}{\sqrt{1-2\left(\alpha_1^2+\alpha_2^2\right)B(t+1)}}.$$

since

$$E\left[a\sqrt{v}z_1+\psi\left(b_1z_1+b_2z_2-c\sqrt{v}\right)^2\right] = \frac{\exp\left[\frac{(-2c^2\psi+4b_1ca\psi+a^2(-1+2b_2^2\psi))v}{-2+4(b_1^2+b_2^2)\psi}\right]}{\sqrt{1-2(b_1^2+b_2^2)\psi}}. \tag{A.3}$$

By equating equations (A.1) and (A.2), we obtain the following recursive coefficients:

$$\begin{aligned}
E_t\left[S(T)^\phi\right] &= S(t)^\phi \exp\left(A(t)+B(t)v(t)\right), \\
A(t) &= A(t+1)+\phi r+\omega B(t+1)-\frac{1}{2}\log\left(1-2\left(\alpha_1^2+\alpha_2^2\right)B(t+1)\right) \\
B(t) &= \phi\left(\mu-\frac{1}{2}\right)+\beta B(t+1)+\frac{2\lambda^2B(t+1)-4\alpha_1\lambda\phi B(t+1)+\phi^2(1-2\alpha_2^2B(t+1))}{2-4(\alpha_1^2+\alpha_2^2)B(t+1)}, \\
A(T) &= B(T)=0.
\end{aligned}$$

A.2 The SV(p,q) Dynamics

This section presents the MGF for the SV(p,q) dynamics. The discrete-time SV(p,q) dynamic is as follows:

$$\begin{aligned}
\ln(S(t+1)) &= \ln(S(t))+\left(r+\left(\mu-\frac{1}{2}\right)v(t)\right)+\sqrt{v(t)}z_1(t+1), \qquad (A.4)\\
v(t+1) &= \omega+\sum_{i=1}^{p}\beta_i v(t+1-i) \\
&\quad +\sum_{i=1}^{q}k_i\left(\alpha_1z_1(t+2-i)+\alpha_2z_2(t+2-i)-\lambda_i\sqrt{v(t+1-i)}\right)^2,
\end{aligned}$$

where $k_1 = 1$. This SV(p,q) dynamics nest the GARCH(p,q) dynamics in Heston and Nandi (2000) and Christoffersen, Jacobs, and Ornthanalai (2013) with the slight difference in notation as discussed in Section 2.4; specifically, when $\alpha_2 = 0$, the SV(p,q) model exactly corresponds to the GARCH(p,q) dynamic.

To find the conditional MGF of the stock price, we conjecture

$$f^{MG}(t;\phi) \;=\; \exp\left(\phi x(t) + A(t) + \sum_{i=1}^{p} B_i(t) v(t+1-i) + \sum_{i+j\le q} C_{i,j}(t) Z_{i,j}(t)\right), \qquad \text{(A.5)}$$

where $Z_{i,j}(t) = \left(\alpha_1 z_1(t+1-i) + \alpha_2 z_2(t+1-i) - \lambda_{i+j}\sqrt{v(t-i)}\right)^2$. Proceeding similarly to the SV(1,1) case in Appendix A.1,

$$\begin{aligned}
f^{MG}(t;\phi) &= E_t\left[f^{MG}(t+1;\phi)\right] \\
&= E_t\left[\exp\left(\phi x(t+1) + A(t+1) + \sum_{i=1}^{p} B_i(t+1) v(t+2-i) + \sum_{i+j\le q} C_{i,j}(t+1) Z_{i,j}(t+1)\right)\right] \\
&= \exp\left(\phi(x(t)+r) + A(t+1) + B_1(t+1)\omega + \left(\phi\left(\mu - \frac{1}{2}\right) + B_1(t+1)\beta_1 + B_2(t+1)\right) v(t)\right. \\
&\quad + \sum_{i=2}^{p} \left(B_1(t+1)\beta_i + B_{i+1}(t+1)\right) v(t+1-i) \\
&\quad \left. + \sum_{i=1}^{q-1} B_1(t+1) k_i Z_{i,1}(t) + \sum_{i=2}^{q-1} \sum_{j\le q-i} C_{i,j}(t+1) Z_{i-1,j+1}(t)\right) \\
&\quad \times \underbrace{E_t\left[\exp\left(\phi\sqrt{v(t)} z_1(t+1) + B_1(t+1) k_1 \left(\alpha_1 z_1(t+1) + \alpha_2 z_2(t+1) - \lambda_1\sqrt{v(t)}\right)^2\right.\right.}_{(**)} \\
&\quad \underbrace{\left.\left. + \sum_{1\le j\le q-1} C_{1,j}(t+1)\left(\alpha_1 z_1(t+1) + \alpha_2 z_2(t+1) - \lambda_{1+j}\sqrt{v(t)}\right)^2\right)\right]}_{(**)}. \qquad \text{(A.6)}
\end{aligned}$$

The exponent component inside the expectation $(**)$ can be rearranged as

$$\begin{aligned}
&\phi\sqrt{v(t)}z_1(t+1) + \left(B_1(t+1)k_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\right) \\
&\qquad \times \left(\alpha_1 z_1(t+1) + \alpha_2 z_2(t+1) - \frac{B_1(t+1)k_1\lambda_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\lambda_{1+j}}{B_1(t+1)k_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)}\sqrt{v(t)}\right)^2 \\
&- \left[\frac{\left(B_1(t+1)k_1\lambda_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\lambda_{1+j}\right)^2}{\left(B_1(t+1)k_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\right)} - \left(B_1(t+1)k_1\lambda_1^2 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\lambda_{1+j}^2\right)\right] v(t).
\end{aligned}$$

Substituting into equation (A.3) and then equating equations (A.5) and (A.6), we obtain the following recursions for the MGF coefficients:

$$\begin{aligned}
A(t) &= A(t+1) + \phi r + \omega B_1(t+1) - \frac{1}{2}\log\left(1 - 2\left(\alpha_{1,1}^2 + \alpha_{2,1}^2\right)\left(B_1(t+1)k_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\right)\right) \\
B_1(t) &= \phi\left(\mu - \frac{1}{2}\right) + \beta_1 B_1(t+1) + B_2(t+1) + \frac{\phi^2\left(\frac{1}{2} - \alpha_2^2\left(B_1(t+1)k_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\right)\right)}{1 - 2\left(\alpha_1^2 + \alpha_2^2\right)\left(B_1(t+1)k_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\right)} \\
&\quad + \frac{2\left(\alpha_1^2 + \alpha_2^2\right)\left(B_1(t+1)k_1\lambda_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\lambda_{1+j}\right)^2}{1 - 2\left(\alpha_1^2 + \alpha_2^2\right)\left(B_1(t+1)k_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\right)} \\
&\quad + \frac{-2\alpha_1\left(B_1(t+1)k_1\lambda_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\lambda_{1+j}\right)\phi}{1 - 2\left(\alpha_1^2 + \alpha_2^2\right)\left(B_1(t+1)k_1 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\right)} \\
&\quad + B_1(t+1)k_1\lambda_1^2 + \sum_{1\le j\le q-1} C_{1,j}(t+1)\lambda_{1+j}^2,
\end{aligned}$$

$$
\begin{aligned}
B_i(t) &= \beta_i B_1(t+1) + B_{i+1}(t+1) \qquad \text{for } 2 \le i \le p-1, \\
B_p(t) &= \beta_p B_1(t+1), \\
C_{i,j}(t) &= \begin{cases} k_{i+1} B_1(t+1) & \text{if } j = 1 \text{ and } i + j \le q, \\ C_{i+1,j-1}(t+1) & \text{if } j \ne 1 \text{ and } i + j \le q, \\ 0 & \text{otherwise,} \end{cases} \\
A(T) &= B_i(T) = C_{i,j}(T) = 0.
\end{aligned}
$$

As a special case, when $\alpha_2 = 0$, this solution for the MGF coefficients exactly matches the formula in Appendix C of Christoffersen, Jacobs, and Ornthanalai (2013) for the GARCH(p,q) model.

B No-Arbitrage Restrictions for the Exponential-Affine Pricing Kernel

B.1 The Pricing Kernel

We re-write the pricing kernel in equation (10) as follows:

$$
\frac{M(t+1)}{M(t)} = \left(\frac{S(t+1)}{S(t)} \right)^{-\gamma} \exp(\delta + \eta v(t) + \xi(v(t+1) - v(t))), \tag{B.1}
$$

where time is expressed in days. From equation (2), we have

$$
\begin{aligned}
\frac{S(t+1)}{S(t)} &= \exp\left(r + \left(\mu - \frac{1}{2} \right) v(t) + \sqrt{v(t)} z_1(t+1) \right), \\
v(t+1) - v(t) &= \omega + (\beta - 1) v(t) + \left(\alpha_1 z_1(t+1) + \alpha_2 z_2(t+1) - \lambda \sqrt{v(t)} \right)^2 .
\end{aligned} \tag{B.2}
$$

Substituting equation (B.2) into equation (B.1), we obtain:

$$
\begin{aligned}
\frac{M(t+1)}{M(t)} &= \exp\left[-\gamma r - \gamma \left(\mu - \frac{1}{2} \right) v - \gamma \sqrt{v} z_1 + \delta + \eta v + \xi \omega + \xi(\beta - 1) v + \xi \left(\alpha_1 z_1 + \alpha_2 z_2 - \lambda \sqrt{v} \right)^2 \right] \\
&= \exp\left[-\gamma r + \delta + \xi \omega - \left(\gamma \left(\mu - \frac{1}{2} \right) - \eta - \xi(\beta - 1) \right) v - \gamma \sqrt{v} z_1 + \xi \left(\alpha_1 z_1 + \alpha_2 z_2 - \lambda \sqrt{v} \right)^2 \right]
\end{aligned}
$$

The no-arbitrage condition implies $E_t\left[\frac{M(t+1)}{M(t)}\right] = \exp(-r)$. The conditional expectation of the pricing kernel becomes

$$\begin{aligned}
E_t\left[\frac{M(t+1)}{M(t)}\right] &= \exp\left[-\gamma r + \delta + \xi\omega + \left(-\gamma\left(\mu - \frac{1}{2}\right) + \eta + \xi(\beta-1)\right)v\right] \\
&\quad \times E_t\left[\exp\left(-\gamma\sqrt{v}z_1 + \xi\left(\alpha_1 z_1 + \alpha_2 z_2 - \lambda\sqrt{v}\right)^2\right)\right] \\
&= \exp\left[-\gamma r + \delta + \xi\omega + \left(-\gamma\left(\mu - \frac{1}{2}\right) + \eta + \xi(\beta-1)\right)v\right] \\
&\quad \times \frac{\exp\left[\frac{(-2\lambda^2\xi - 4\alpha_1\lambda\gamma\xi + \gamma^2(-1+2\alpha_2^2\xi))v}{-2+4(\alpha_1^2+\alpha_2^2)\xi}\right]}{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}} \qquad \text{by equation (A.3)} \\
&= \exp\left[-\gamma r + \delta + \xi\omega - \frac{1}{2}\log\left(1 - 2\left(\alpha_1^2+\alpha_2^2\right)\xi\right)\right. \\
&\quad \left. + \left(-\gamma\left(\mu - \frac{1}{2}\right) + \eta + \xi(\beta-1) + \frac{-2\lambda^2\xi - 4\alpha_1\lambda\gamma\xi + \gamma^2(-1+2\alpha_2^2\xi)}{-2+4(\alpha_1^2+\alpha_2^2)\xi}\right)v\right] \qquad \text{(B.3)}
\end{aligned}$$

By equating equation (B.3) to e^{-r}, we get:

$$\begin{aligned}
&(-\gamma+1)r + \delta + \xi\omega - \frac{1}{2}\log\left(1 - 2\left(\alpha_1^2+\alpha_2^2\right)\xi\right) = 0, \\
&-\gamma\left(\mu - \frac{1}{2}\right) + \eta + \xi(\beta-1) + \frac{-2\lambda^2\xi - 4\alpha_1\lambda\gamma\xi + \gamma^2(-1+2\alpha_2^2\xi)}{-2+4(\alpha_1^2+\alpha_2^2)\xi} = 0.
\end{aligned}$$

Rearranging, we obtain

$$\begin{aligned}
\delta &= (\gamma-1)r - \xi\omega + \frac{1}{2}\log\left(1 - 2\left(\alpha_1^2+\alpha_2^2\right)\xi\right), \qquad \text{(B.4)} \\
\eta &= \left(\mu - \frac{1}{2}\right)\gamma + \xi(1-\beta) - \frac{-2\lambda^2\xi - 4\alpha_1\lambda\gamma\xi + \gamma^2(-1+2\alpha_2^2\xi)}{-2+4(\alpha_1^2+\alpha_2^2)\xi}. \qquad \text{(B.5)}
\end{aligned}$$

Proceeding in a similar fashion starting with $E_t\left[\frac{S(t+1)}{(S_t)}\frac{M(t+1)}{M(t)}\right] = 1$, we get:

$$\begin{aligned}
1 = \exp\Bigg[&(1-\gamma)r + \delta + \xi\omega - \frac{1}{2}\log\left(1 - 2\left(\alpha_1^2+\alpha_2^2\right)\xi\right) \\
&+ \left((1-\gamma)\left(\mu - \frac{1}{2}\right) + \eta + \xi(\beta-1) + \frac{-2\lambda^2\xi + 4\alpha_1\lambda(1-\gamma)\xi + (1-\gamma)^2(-1+2\alpha_2^2\xi)}{-2+4(\alpha_1^2+\alpha_2^2)\xi}\right)v\Bigg]. \qquad \text{(B.6)}
\end{aligned}$$

Substituting equations (B.4) and (B.5) in equation (B.6), we get:

$$\left(\mu - \frac{1}{2}\right) + \frac{4\alpha_1\lambda\xi + (1-2\gamma)(-1+2\alpha_2^2\xi)}{-2+4(\alpha_1^2+\alpha_2^2)\xi} = 0.$$

By rearranging the above equation, γ is given by

$$\gamma = \mu - \frac{\alpha_1\xi(2\alpha_1\mu - \alpha_1 + 2\lambda)}{1-2\alpha_2^2\xi}. \tag{B.7}$$

B.2 The Risk-Neutral Probability Distribution

The physical and risk-neutral probability distributions are related as follows:

$$\begin{aligned} f_t^*(S(t+1), v(t+1)) &= \frac{f_t(S(t+1), v(t+1))M(t+1)}{E_t(M(t+1))} \\ &= f_t(S(t+1), v(t+1)) \times \frac{S^{-\gamma}(t+1)\exp\left(\delta\times(t+1) + \eta\sum_{s=0}^{t} v(s) + \xi v(t+1)\right)}{E_t\left[S^{-\gamma}(t+1)\exp\left(\delta\times(t+1) + \eta\sum_{s=0}^{t} v(s) + \xi v(t+1)\right)\right]}. \end{aligned} \tag{B.8}$$

Equation (2) implies

$$\begin{aligned} S^{-\gamma}(t+1) &= S^{-\gamma}(t)\exp\left(-\gamma r - \gamma\left(\mu - \frac{1}{2}\right)v(t) - \gamma\sqrt{v(t)}z_1(t+1)\right), \\ v(t+1) &= \omega + \beta v(t) + \left(\alpha_1 z_1(t+1) + \alpha_2 z_2(t+1) - \lambda\sqrt{v(t)}\right)^2, \end{aligned}$$

Equation (B.8) can then be expressed as

$$f_t^*(S(t+1), v(t+1)) = f_t(S(t+1), v(t+1)) \times \frac{\exp\left(-\gamma\sqrt{v}z_1 + \xi\left(\alpha_1 z_1 + \alpha_2 z_2 - \lambda\sqrt{v}\right)^2\right)}{E_t\left[\exp\left(-\gamma\sqrt{v}z_1 + \xi\left(\alpha_1 z_1 + \alpha_2 z_2 - \lambda\sqrt{v}\right)^2\right)\right]}, \tag{B.9}$$

where we have simplified the notation as follows: $z_1 = z_1(t+1)$, $z_2 = z_2(t+1)$, and $v = v(t)$, and we can calculate the expectation in the denominator by using equation (A.3):

$$E_t\left[\exp\left(-\gamma\sqrt{v}z_1 + \xi\left(\alpha_1 z_1 + \alpha_2 z_2 - \lambda\sqrt{v}\right)^2\right)\right] = \frac{1}{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}}\exp\left[\frac{(-2\lambda^2\xi - 4\alpha_1\lambda\gamma\xi + \gamma^2(-1+2\alpha_2^2\xi))v}{-2+4(\alpha_1^2+\alpha_2^2)\xi}\right].$$

To find the joint distribution $f_t(S(t+1), v(t+1))$, we use Bayes' rule:

$$f_t(S(t+1), v(t+1)) = f_t(v(t+1)|S(t+1))f_t(S(t+1)).$$

Note that $f_t(S(t+1))$ follows a normal distribution, governed by z_1, and that $f_t(v(t+1)|S(t+1))$ follows a non-central chi-squared distribution because $\left(\frac{\alpha_1}{\alpha_2}z_1 + z_2 - \frac{\lambda}{\alpha_2}\sqrt{v}\right)$ follows a normal distribution with nonzero mean, conditional on z_1 and v. Since $f_t(S(t+1), v(t+1))$ can be obtained from $f_t(z_1, y)$ with the help of a Jacobian, where $y \equiv \left(\frac{\alpha_1}{\alpha_2}z_1 + z_2 - \frac{\lambda}{\alpha_2}\sqrt{v}\right)^2$, we simply focus on $f_t(z_1, y)$.

The joint risk-neutral probability distribution of z_1 and y is given by:

$$\begin{aligned} f_t^*(z_1, y) &= f_t(y|z_1)f_t(z_1) \times \frac{\exp\left(-\gamma\sqrt{v}z_1 + \xi\alpha_2^2 y\right)}{E_t\left[\exp\left(-\gamma\sqrt{v}z_1 + \xi\alpha_2^2 y\right)\right]} \\ &= NC\chi_y^2(1,\psi)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z_1^2}{2}\right) \times \frac{\exp\left(-\gamma\sqrt{v}z_1\right)\exp\left(\xi\alpha_2^2 y\right)}{E_t\left[\exp\left(-\gamma\sqrt{v}z_1 + \xi\alpha_2^2 y\right)\right]}, \end{aligned} \tag{B.10}$$

where $NC\chi_y^2(1,\chi)$ is the non-central chi-squared distribution function with degree of freedom 1 and non-centrality parameter $\chi = \left(\frac{\alpha_1}{\alpha_2}z_1 - \frac{\lambda}{\alpha_2}\sqrt{v}\right)^2$. $f_t(y|z_1) = NC\chi_y^2(1,\chi)$ is given as

$$\begin{aligned} f_t(y|z_1) &= \frac{1}{2}e^{-\frac{y+\chi}{2}}\left(\frac{y}{\chi}\right)^{-\frac{1}{4}} I_{\left(-\frac{1}{2}\right)}\left(\sqrt{\chi y}\right) \\ &= \frac{1}{2\sqrt{2\pi y}}\left[\exp\left(-\frac{1}{2}\left(\sqrt{\chi}+\sqrt{y}\right)^2\right) + \exp\left(-\frac{1}{2}\left(\sqrt{\chi}-\sqrt{y}\right)^2\right)\right], \end{aligned} \tag{B.11}$$

where I is a modified Bessel function of first kind.

We first find $f_t^*(z_1)$ by integrating $f_t^*(z_1, y)$ with respect to y over its support. That is,

$$f_t^*(z_1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_1^2}{2}\right) \times \frac{\exp\left(-\gamma\sqrt{v}z_1\right)}{E_t\left[\exp\left(-\gamma\sqrt{v}z_1 + \xi\alpha_2^2 y\right)\right]} \times \underbrace{\int_0^\infty NC\chi_y^2(1,\chi)\exp\left(\xi\alpha_2^2 y\right) dy.}_{(*)} \quad \text{(B.12)}$$

The integral part $(*)$ is a moment generating function (MGF) of y. As MGF of a non-central chi-squared distribution with one degree of freedom is given by $E\left[e^{\phi y}\right] = \exp\left(\frac{\phi\chi}{1-2\phi}\right)/(1-2\phi)^{1/2}$, we can rewrite equation (B.12) as the following:

$$\begin{aligned}
f_t^*(z_1) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_1^2}{2}\right) \times \frac{\exp\left(-\gamma\sqrt{v}z_1\right)}{E_t\left[\exp\left(-\gamma\sqrt{v}z_1 + \xi\alpha_2^2 y\right)\right]} \times \frac{1}{\sqrt{1-2\xi\alpha_2^2}} \exp\left(\frac{\xi\alpha_2^2\chi}{1-2\xi\alpha_2^2}\right) \\
&= \frac{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}}{\sqrt{1-2\xi\alpha_2^2}} \exp\left[\frac{(-2\lambda^2\xi - 4\alpha_1\lambda\gamma\xi + \gamma^2(-1+2\alpha_2^2\xi))v}{2-4(\alpha_1^2+\alpha_2^2)\xi}\right] \\
&\quad \times \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_1^2}{2} - \gamma\sqrt{v}z_1 + \frac{\xi\alpha_2^2\left(\frac{\alpha_1}{\alpha_2}z_1 - \frac{1}{\alpha_2}\lambda\sqrt{v}\right)^2}{1-2\xi\alpha_2^2}\right) \\
&= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(z_1 - E^*(z_1))^2}{2\sigma^2}\right], \qquad \text{(B.13)}
\end{aligned}$$

where $E^*(z_1) = \frac{((-1+2\xi\alpha_2^2)\gamma - 2\xi\alpha_1\lambda)\sqrt{v}}{1-2(\alpha_1^2+\alpha_2^2)\xi}$ and $\sigma = \frac{\sqrt{1-2\xi\alpha_2^2}}{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}}$. We now show that z_1 also follows a normal distribution under the new measure. Substituting in for γ from equation (B.7):

$$\begin{aligned}
E^*(z_1) &= \frac{((-1+2\xi\alpha_2^2)\gamma - 2\xi\alpha_1\lambda)\sqrt{v}}{1-2(\alpha_1^2+\alpha_2^2)\xi} \\
&= -\left(\mu + \frac{\alpha_1^2\xi}{1-2(\alpha_1^2+\alpha_2^2)\xi}\right)\sqrt{v}.
\end{aligned}$$

We can then define z_1^*, which follows a standard normal distribution under the risk-neutral measure, *i.e.* with mean 0 and variance 1, as follows:

$$z_1^*(t+1) = \frac{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}}{\sqrt{1-2\xi\alpha_2^2}}\left(z_1(t+1) + \left(\mu + \frac{\alpha_1^2\xi}{1-2(\alpha_1^2+\alpha_2^2)\xi}\right)\sqrt{v}\right). \quad \text{(B.14)}$$

Now we focus on $f_t^*(y|z_1) = f_t^*(y, z_1)/f_t^*(z_1)$, which follows a non-central chi-squared distribution.

Using equations (B.10) and (B.13) to obtain the conditional distribution, we get:

$$\begin{aligned} f_t^*(y|z_1) &= \sqrt{1-2\xi\alpha_2^2} \times NC\chi_y^2(1,\chi) \times \exp\left(\xi\alpha_2^2 y - \frac{\left(\alpha_1 z_1 - \lambda\sqrt{v}\right)^2 \xi}{1-2\xi\alpha_2^2}\right) \\ &= \sqrt{1-2\xi\alpha_2^2} \times NC\chi_y^2(1,\chi) \times \exp\left(\xi\alpha_2^2 y - \frac{\alpha_2^2\chi\xi}{1-2\xi\alpha_2^2}\right). \end{aligned} \tag{B.15}$$

Recall that the non-central chi-squared distribution $NC\chi_y^2(1,\chi)$ is given by equation (B.11). Equation (B.15) is then given by:

$$\begin{aligned} f_t^*(y|z_1) &= \sqrt{1-2\xi\alpha_2^2} \\ &\quad \times \frac{1}{2\sqrt{2\pi y}}\left[\exp\left(-\frac{1}{2}\left(\sqrt{\chi}+\sqrt{y}\right)^2\right) + \exp\left(-\frac{1}{2}\left(\sqrt{\chi}-\sqrt{y}\right)^2\right)\right] \exp\left(\xi\alpha_2^2 y - \frac{\alpha_2^2\chi\xi}{1-2\xi\alpha_2^2}\right) \\ &= \sqrt{1-2\xi\alpha_2^2} \times \frac{1}{2\sqrt{2\pi y}}\left[\exp\left(-\frac{1}{2}\left(\sqrt{(1-2\xi\alpha_2^2)y} + \sqrt{\frac{\chi}{1-2\xi\alpha_2^2}}\right)^2\right)\right. \\ &\qquad \left. + \exp\left(-\frac{1}{2}\left(\sqrt{(1-2\xi\alpha_2^2)y} - \sqrt{\frac{\chi}{1-2\xi\alpha_2^2}}\right)^2\right)\right]. \end{aligned}$$

Let $y^* = (1-2\xi\alpha_2^2)y$ and $\tilde{\chi} = \chi/(1-2\xi\alpha_2^2)$. By a change of random variable,

$$\begin{aligned} f_t^*(y^*|z_1) &= f_t^*(y|z_1) \times J \\ &= \frac{1}{2\sqrt{2\pi y^*}}\left[\exp\left(-\frac{1}{2}\left(\sqrt{y^*}+\sqrt{\tilde{\chi}}\right)^2\right) + \exp\left(-\frac{1}{2}\left(\sqrt{y^*}-\sqrt{\tilde{\chi}}\right)^2\right)\right], \end{aligned} \tag{B.16}$$

where the Jacobian $J = \frac{1}{\left|\frac{\partial y^*}{\partial y}\right|} = \frac{1}{1-2\xi\alpha_2^2}$. Equation (B.16) defines a non-central chi-squared distribution with one degree of freedom and non-centrality parameter $\tilde{\chi}$.

While y is defined as $\left(z_2+\sqrt{\chi}\right)^2$, it does not follow a non-central chi-squared distribution under f_t^*. Therefore, we would like to find a risk-neutral measure z_2^* such that it follows a standard normal distribution and y^* has a form of $y^* = \left(z_2^*+\sqrt{\tilde{\chi}}\right)^2$. Since $y^* = (1-2\xi\alpha_2^2)y$, we can eventually find

a relationship between z_2^* and z_2 such that $\left(z_2^* + \sqrt{\tilde{\chi}}\right)^2 = (1 - 2\xi\alpha_2^2)\left(z_2 + \sqrt{\chi}\right)^2$, namely:

$$\begin{aligned} z_2^* &= -\sqrt{\tilde{\chi}} + \sqrt{(1-2\xi\alpha_2^2)(z_2+\sqrt{\chi})^2} \\ &= \sqrt{1-2\xi\alpha_2^2}\left[(z_2+\sqrt{\chi}) - \frac{\sqrt{\chi}}{1-2\xi\alpha_2^2}\right] \\ &= \sqrt{1-2\xi\alpha_2^2}\left[z_2 - \frac{2\xi\alpha_2\left(\alpha_1 z_1 - \lambda\sqrt{v}\right)}{1-2\xi\alpha_2^2}\right]. \end{aligned} \tag{B.17}$$

Using equations (1), (2), (B.14), and (B.17), we obtain the restrictions between the physical and risk-neutral parameters. First, by rearranging equation (B.14),

$$\begin{aligned} \ln S(t+1) &= \ln S(t) + r + \left(\mu - \frac{1}{2}\right)v(t) + \frac{\sqrt{1-2\xi\alpha_2^2}\sqrt{v(t)}}{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}} z_1^*(t+1) - \mu v(t) - \frac{\alpha_1^2\xi}{1-2(\alpha_1^2+\alpha_2^2)\xi} v(t) \\ &= \ln S(t) + r - \frac{1}{2}v^*(t) + \sqrt{v^*(t)}z_1^*(t+1), \end{aligned}$$

where $v^*(t) = \frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}v(t)$. As a special case, if $\xi = 0$, *i.e.* in the absence of a variance risk premium, $v^*(t) = v(t)$.

Furthermore, we can find the dynamics of $v^*(t)$ in equation (2) by multiplying by $\frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}$:

$$v^*(t+1) = \frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}\omega + \beta v^*(t) + \frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}\left(\alpha_1 z_1 + \alpha_2 z_2 - \lambda\sqrt{v(t)}\right)^2.$$

By replacing z_1 and z_2 with equations (B.14) and (B.17),

$$\begin{aligned} v^*(t+1) &= \frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}\omega + \beta v^*(t) + \frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}\left[\frac{\alpha_2}{\sqrt{1-2\xi\alpha_2^2}} z_2^*(t+1)\right. \\ &\quad \left. + \frac{\alpha_1}{1-2\xi\alpha_2^2}\left(\frac{\sqrt{1-2\xi\alpha_2^2}}{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}} z_1^*(t+1) - \left(\mu + \frac{\alpha_1^2\xi}{1-2(\alpha_1^2+\alpha_2^2)\xi}\right)\sqrt{v(t)}\right) - \frac{\lambda\sqrt{v(t)}}{1-2\xi\alpha_2^2}\right]^2 \\ &= \omega^* + \beta^* v^*(t) + \left(\alpha_1^* z_1^*(t+1) + \alpha_2^* z_2^*(t+1) - \lambda^*\sqrt{v^*(t)}\right)^2, \end{aligned}$$

where

$$\omega^* = \frac{1-2\xi\alpha_2^2}{1-2(\alpha_1^2+\alpha_2^2)\xi}\omega,\ \beta^* = \beta,\ \alpha_1^* = \frac{\alpha_1}{1-2(\alpha_1^2+\alpha_2^2)\xi},\ \alpha_2^* = \frac{\alpha_2}{\sqrt{1-2(\alpha_1^2+\alpha_2^2)\xi}},$$
$$\text{and}\ \ \lambda^* = \left(\frac{\alpha_1\mu+\lambda}{1-2\xi\alpha_2^2} + \frac{\alpha_1^3\xi}{(1-2\xi\alpha_2^2)(1-2(\alpha_1^2+\alpha_2^2)\xi)}\right). \tag{B.18}$$

C The Daily Stochastic Variance and the VIX

Recall the following risk-neutral stochastic variance dynamics:

$$v^*(t+1) = \omega^* + \beta^* v^*(t) + \left(\alpha_1^* z_1^*(t+1) + \alpha_2^* z_2^*(t+1) - \lambda^*\sqrt{v^*(t)}\right)^2.$$

By expanding the squared term, we can rewrite as follows:

$$\begin{aligned} v^*(t+1) = {} & \underbrace{\omega^* + (\alpha_1^*)^2 + (\alpha_2^*)^2}_{\theta_v} + \underbrace{\left(\beta^* + (\lambda^*)^2\right)}_{\rho_v} v^*(t) \\ & + \underbrace{\left((\alpha_1^* z_1^*)^2 + (\alpha_2^* z_2^*)^2 - (\alpha_1^*)^2 - (\alpha_2^*)^2 + 2\alpha_1^*\alpha_2^* z_1^* z_2^* - 2\alpha_1^*\lambda^*\sqrt{v(t)}z_1^* - 2\alpha_2^*\lambda^*\sqrt{v(t)}z_2^*\right)}_{\epsilon(t+1)}. \end{aligned} \tag{C.1}$$

Since z_1^* and z_2^* are independent random variables following the standard normal distribution, $E_t^*(z_1^*) = E_t^*(z_2^*) = 0$, $E_t^*\left((z_1^*)^2\right) = E_t^*\left((z_2^*)^2\right) = 1$, and $E_t^*(z_1^* z_2^*) = 0$. Then, $E_t^*(\epsilon(t+1)) = 0$.

By defining $\text{VIX}_d^2(t) = \frac{1}{252}\left(\frac{\text{VIX}(t)}{100}\right)^2$, we get:

$$\text{VIX}_d^2(t) \ = \ \frac{1}{n}\sum_{k=1}^{n} E_t^*(v^*(t+k))$$

where n is the number of trading days per month. Using the law of iterated expectation,

$$\begin{aligned}
E_t^*(v^*(t+k)) &= E_t^*\left(\theta_v + \rho_v v^*(t+(k-1)) + \epsilon(t+k)\right) \\
&= E_t^*\left(\theta_v + E_{t+(k-1)}^*\left(\rho_v v^*(t+(k-1)) + \epsilon(t+k)\right)\right) \\
&= E_t^*\left(\theta_v + \rho_v v^*(t+(k-1))\right) \\
&= \cdots \\
&= \left(\theta_v + \rho_v\theta_v + \cdots + \rho_v^{k-1}\theta_v\right) + \rho_v^k v^*(t) \\
&= \frac{\theta_v\left(1-\rho_v^k\right)}{1-\rho_v} + \rho_v^k v^*(t).
\end{aligned}$$

Thus, $\text{VIX}_d^2(t)$ can be expressed as:

$$\text{VIX}_d^2(t) = \psi_0 + \psi_1 v^*(t)$$

where $\psi_0 = \frac{\theta_v(1-\psi_1)}{1-\rho_v}$ and $\psi_1 = \frac{\rho_v(1-\rho_v^n)}{n(1-\rho_v)}$.