NBER WORKING PAPER SERIES

# EMPIRICAL OPTION PRICING MODELS

David S. Bates

Working Paper 29554
http://www.nber.org/papers/w29554

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2021

This paper received no external financing other than that provided to me by the University of Iowa as a professor. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Empirical Option Pricing Models
David S. Bates
NBER Working Paper No. 29554
December 2021
JEL No. G13

**ABSTRACT**

This paper is an overview of empirical options research, with primary emphasis on research into systematic stochastic volatility and jump risks relevant for pricing stock index options. The paper reviews evidence from time series analysis, option prices and option price evolution regarding those risks, and discusses required compensation.

David S. Bates
Henry B. Tippie College of Business
Department of Finance
University of Iowa
Iowa City, IA 52242-1000
and NBER
david-bates@uiowa.edu

Options are derivatives, with prices and payoffs sensitive to the stochastic evolution of the underlying asset price. Much of the academic interest in options originates in their potential information regarding the nature of that stochastic evolution. In particular, options research has attempted to shed light on the following questions:

1) What risks are of concern to investors?
2) What compensation do investors require for bearing those risks?

Early options research surveyed in Smith (1976) sought to determine how stock options should be priced to generate appropriate compensation for their content of underlying stock price risk. A resolution in Black & Scholes (1973) is that option returns are Gaussian in continuous time under the assumption of geometric Brownian motion for the underlying stock price, yielding a CAPM-based justification of the Black-Scholes-Merton option pricing formula. (The no-arbitrage foundations of the formula in Black & Scholes (1973) and Merton (1973) understandably received greater attention.)

Subsequent option pricing research expanded the types of risks under consideration to include stochastic volatility, stochastic interest rates, and jump risk. This research developed in parallel with – and at times interacting with – research into the stochastic properties of underlying asset returns. The continuous-time stochastic volatility option pricing models of the 1980's were developed at the same time as extensive discrete-time ARCH modeling of the conditional volatility of daily asset returns. Multifactor models of stochastic volatility are similar to the component GARCH model of Engle & Lee (1999). Jump risk models such as Merton (1976) were influenced by earlier research by Mandelbrot (1963) and Fama (1965) into non-Gaussian properties of asset returns.

This article primarily reviews empirical research into models used to price stock index options. Sections 1-4 discuss estimates of systematic stochastic volatility and jump risks underlying stock market evolution from time series analysis, from option prices, and from how option prices evolve. Section 5 discusses average option returns (including the implicit pricing kernel puzzle), while Section 6 concludes with interpretations of those returns.

## 1. Models

The papers reviewed in this article assume the underlying asset price $S_t$ follows a low-order Markov process – i.e., a relatively small number $K$ of underlying state variables $\boldsymbol{Y}_t = (Y_{1t}, \ldots, Y_{Kt})$ fully summarize conditional distributions of future asset returns at any point in time. Continuous-time models of the joint stochastic evolution of $S_t$, $\boldsymbol{Y}_t$ and the positive pricing kernel $M_t$ used when pricing options typically involve special cases of the following jump-diffusions:

$$\frac{dS_t}{S_t} = \left[b(\boldsymbol{Y}_t)dt - Cov_t\left(\frac{dS_t}{S_t}, \frac{dM_t}{M_t}\right)\right] + \sigma(\boldsymbol{Y}_t)dW_t + \left[(e^{\gamma} - 1)\, dN_t - \lambda(\boldsymbol{Y}_t)\bar{k}_t dt\right]$$
$$dY_{it} = \mu_{Yi}(\boldsymbol{Y}_t)dt + \sigma_{Yi}(\boldsymbol{Y}_t)dW_{Yit} + [\gamma_{Yi} dN_{Yit} - \lambda_{Yi}(\boldsymbol{Y}_t)\bar{\gamma}_{Yit} dt], i = 1, \ldots, K \quad 1.$$
$$\frac{dM_t}{M_t} = -r(\boldsymbol{Y}_t)dt + \sigma_m(\boldsymbol{Y}_t)dW_{mt} + \left[(e^{\gamma_m} - 1)\, dN_{mt} - \lambda_m(\boldsymbol{Y}_t)\bar{k}_{mt} dt\right].$$

$b(\boldsymbol{Y}_t)$ is the cost of carry on the underlying asset: zero for futures, $r(\boldsymbol{Y}_t) - \delta$ if the instantaneous riskless rate is $r(\boldsymbol{Y}_t)$ and the asset pays a continuous dividend yield $\delta$. $(W_t, W_{Yit}, W_{mt})$ are correlated Wiener processes, while $(N_t, N_{Yit}, N_{mt})$ are correlated (and possibly identical) Poisson counters with intensities $\{\lambda(\boldsymbol{Y}_t), \lambda_{Yi}(\boldsymbol{Y}_t), \lambda_{Yi}(\boldsymbol{Y}_t)\}$ that identify random jumps $(\gamma, \gamma_{Yi}, \gamma_m)$ in $(\ln S_t, Y_{it}, \ln M_t)$. The jump components have conditional means $(\bar{k}_t, \bar{\gamma}_{Yit}, \bar{k}_{mt}) = E[(e^{\gamma} - 1, \gamma_{Yi}, e^{\gamma_m} - 1)|\boldsymbol{Y}_t]$. Papers that explore Lévy processes use infinite- rather than finite-activity jump processes, but can still have a low-order Markov representation (Bates 2012).

Options are then priced at expected discounted terminal payoff using an equivalent "risk-neutral" process that incorporates required $M_t$-based compensation for the various risks:

$$\frac{dS_t}{S_t} = b(\boldsymbol{Y}_t)\, dt + \sigma(\boldsymbol{Y}_t)dW_t^* + \left[(e^{\gamma^*} - 1)dN_t^* - \lambda^*(\boldsymbol{Y}_t)\bar{k}_t^* dt\right] \quad 2.$$
$$dY_{it} = \mu_{Yi}^*(\boldsymbol{Y}_t)dt + \sigma_{Yi}(\boldsymbol{Y}_t)dW_{Yit}^* + [\gamma_{Yi}^* dN_{Yit}^* - \lambda_{Yi}^*(\boldsymbol{Y}_t)\bar{\gamma}_{Yit}^* dt], \qquad i = 1, \ldots, K,$$

where $\mu_{Yi}^*(\boldsymbol{Y}_t) = \mu_{Yi}(\boldsymbol{Y}_t) + Cov_t(dY_{it}, dM_t/M_t)/dt$. The variance-covariance matrix of diffusive $(S_t, \boldsymbol{Y}_t)$ shocks is identical under the objective and risk-neutral measures, but jump intensities and distributions generally change.[1] As noted in Heston (1993) and extended by various authors (including myself), European option pricing is substantially more tractable if the drifts, variances, covariances and jump intensities in Equation 2 are linear (affine) functions of the underlying state variables $\boldsymbol{Y}_t$. As summarized in Duffie et al. (2000), affine models imply the risk-neutral conditional characteristic function for future log returns $\ln(S_T/S_t)$ is an exponentially linear function of $\boldsymbol{Y}_t$, permitting European option evaluation by numerical Fourier inversion.

[1] See, e.g., Bates (1991, Appendix I) or Bates (2006, pp. 944-946).

There is a substantial parallel literature on GARCH option pricing models, using various GARCH specifications to estimate underlying volatility state variables from past daily returns. GARCH models are typically but not invariably low-order Markov; univariate GARCH models such as Bollerslev's (1986) GARCH(1,1) have one underlying state variable, while Engle & Lee's (1999) component GARCH model has two. "Long-memory" volatility processes studied inter alia by Ding et al. (1993) lack a low-order Markov representation, but LeBaron (2001) argues they can be well approximated by a three-factor stochastic volatility model.

Markov processes imply option prices are of the form $O(S_t, \boldsymbol{Y}_t; \tau, X)$ for maturity $\tau$ and strike price $X$. It is common, however, to summarize observed option prices by their implied volatilities, and by the implied volatility surface $IV(X/S_t, \tau)$ across different strike prices and maturities at any point in time. Implied volatility is the value of diffusive volatility $\sigma$ for which the constant-volatility Black-Scholes-Merton option pricing formula $O^{BSM}(S_t, \tau; X, \sigma)$ matches the observed market price of options. Implied volatility surfaces describe the pricing failures of that model, in the same way that yields inferred from a bond pricing model premised on identical discount rates for all maturities are used to describe non-flat term structures of bond yields. Implied volatilities express option prices in relatively intuitive terms that can be compared with time series estimates of annualized conditional volatility, while controlling for other factors (maturity, strike price, dividends, interest rates) that render direct inspection of option prices uninformative. A U-shaped ("smile") pattern in implied volatility patterns across different strike prices indicates a risk-neutral distribution more leptokurtic than the lognormal, while the "smirk" pattern in stock index options' implied volatilities since 1987 indicates substantial negative skewness.

Litterman & Scheinkmann (1991) use principal components analysis to describe how the term structure of bond yields evolves, and conclude that three factors (level, steepness and curvature) capture 94 – 98% of the variance of bond returns at various maturities. Similar analysis for the evolution of two-dimensional IV surfaces is complicated by various methodological issues; in particular, whether the "moneyness" of representative options at various maturities should be measured in absolute terms (e.g., $X/S_t$ or $X/F_{t,T}$ for forward price $F_{t,T}$) or in standard deviation units (e.g., $\ln(X/F_{t,T})/\sigma^{ATM}\sqrt{\tau}$, using the implied volatility from an at-the-money option of maturity $\tau = T - t$). Nevertheless, analyses of the IV surface of S&P 500 options indicate that two to three factors can capture almost all of the variance of IV surface changes.[2] The first is roughly a "level" factor: S&P IV's for all strike prices and maturities tend to rise or fall in unison (albeit not perfectly in parallel), and are negatively correlated with S&P index returns. The next two factors capture variations in the tilt of the moneyness smirk and in the slope of the term structure of IV's. Such studies suggest that an appropriately specified low-order Markov specification in Equation 2 might indeed match observed option prices.

## 2. Estimates of volatility dynamics from time series analysis

Much of what we know about the dynamics of the conditional volatility of asset returns comes from the GARCH literature. More recent numerically intensive approaches for estimating the stochastic volatility models more commonly used in pricing options are similar in also delivering point estimates of short-horizon conditional variances, or "spot variance". Univariate GARCH models recursively update spot variance estimates based on the latest daily return $y_t$; e.g. $h_{t+1} \equiv Var[y_{t+1}|\boldsymbol{\mathcal{I}}_t] = f(y_t, h_t)$ for information set $\boldsymbol{\mathcal{I}}_t = \{y_1, \dots, y_t\}$ and a parameterized news impact curve $f(y_t, h_t)$. Univariate stochastic volatility models use Bayes' rule to update what is known about spot variance $V_t$, as summarized by its conditional distribution

$$p(V_t|\boldsymbol{\mathcal{I}}_t) = \frac{p(y_t, V_t|\boldsymbol{\mathcal{I}}_{t-1})}{p(y_t|\boldsymbol{\mathcal{I}}_{t-1})} = \frac{\int p(y_t, V_t|V_{t-1}) p(V_{t-1}|\boldsymbol{\mathcal{I}}_{t-1}) dV_{t-1}}{\int p(y_t|V_{t-1}) p(V_{t-1}|\boldsymbol{\mathcal{I}}_{t-1}) dV_{t-1}}. \qquad 3.$$

The latest return $y_t$ is used to update yesterday's conditional distribution $p(V_{t-1}|\boldsymbol{\mathcal{I}}_{t-1})$ to today's conditional distribution $p(V_t|\boldsymbol{\mathcal{I}}_t)$. The point estimate of spot variance is then $\hat{V}_{t|t} = \int V_t p(V_t|\boldsymbol{\mathcal{I}}_t) dV_t$.

Both approaches are empirically driven; there is no theoretical justification for choosing one model over another. The GARCH literature explores many different specifications of the news impact curve $f(y_t, h_t)$, whereas the stochastic volatility filtration in Equation 3 relies on whatever Markov specification in Equation 1 underlies the joint transition density $p(y_t, V_t|V_{t-1})$. GARCH and stochastic volatility approaches to spot volatility estimation are similar in the absence of price jumps (Nelson 1992), but can differ substantially if there are jumps (Bates 2006, Table 5). GARCH models assume that conditional spot variance $h_t$ can be accurately estimated from past data (subject to parameter uncertainty), whereas

---

[2] See Skiadopolous et al. (1999) and Cont & da Fonseca (2002) for early studies, and Carr & Wu (2020) for a more recent study. Andersen et al. (2015) have a principal components analysis of IV surfaces, rather than of the change in IV surfaces. They find the first two principal components match IV surfaces over time with $R^2$'s of 96.4% and 98.5%, respectively. The corresponding RMSE's for IV's are 3.38% and 2.2% – respectable fits given IV's can range from 6% to more than 100%.

stochastic volatility models also have state uncertainty $Var(V_t|\mathcal{I}_t)$ from the filtration in Equation 3 – an additional source of risk that may be of concern to investors.[3] There has also been extensive exploration of other data that might be more informative for conditional spot volatility than daily returns; e.g., daily high-low ranges or intradaily realized volatility.[4]

### *2.1. GARCH-based estimates*

The starting point for the GARCH literature was the discovery of substantial autocorrelations in squared and absolute daily returns. Sometimes markets are noisy, sometimes quiet; and recent behavior tends to persist. Influenced by Box & Jenkins (1970) ARMA analysis, Engle (1982) models conditional variance as a constrained AR process, while Bollerslev (1986) and Taylor (1986) extend this to ARMA specifications. Attempts to more accurately model autocorrelations in squared returns at different horizons led to long-memory specifications and to the multifactor component GARCH model of Engle & Lee (1999).

Black (1976) and Christie (1982) observe that stock volatility shocks and equity returns tend to be negatively correlated – a phenomenon commonly labeled "leverage" because of a possible origin in operational or financial leverage by individual firms.[5] Nelson (1991) models the effect using an EGARCH model with an asymmetric news impact curve $f(y_t, h_t)$, and finds the asymmetry quite pronounced for CRSP stock market returns over 1962-1987. Various authors propose alternate specifications for the news impact curve; see Hentschel (1995) for a structure that nests many of the alternatives.

Under the standard GARCH assumption of conditionally Gaussian log-differenced prices, the objective and risk-neutral GARCH processes have identical news impact curves $f(y_t, h_t)$ but different conditional means for asset returns if daily log pricing kernel shocks are also conditionally Gaussian.[6] European options can consequently be priced via Monte Carlo simulation of risk-neutral asset return sample paths. The GARCH model of Heston & Nandi (2000) has a conditional characteristic function for future stock returns that is exponentially affine in spot variance, permitting quicker option evaluation by

---

[3] Shaliastovich (2015) explores state uncertainty as a relevant state variable for equity and option markets.
[4] See Alizadeh et al. (2002) for the former, and Corsi (2009) as an example of the latter.
[5] An alternate explanation is that higher volatility induces heavier discounting of future cash flows, and therefore lowers stock prices and the stock market (Campbell & Hentschel 1992). Poterba & Summers (1986) are skeptical given their estimates of volatility dynamics. A third explanation in Grossman & Zhou (1996) is that leverage is an equilibrium outcome for market volatility from portfolio insurers selling to limit losses as the market falls.
[6] This follows from applying the discrete-time equilibrium change of measure of Rubinstein (1976) and Brennan (1979) to a one-day horizon. Duan (1995) is an early example of this approach for GARCH option pricing.

Fourier inversion. Christoffersen et al. (2009) extend this to a version of the multifactor component GARCH model of Engle & Lee (1999). There has also been examination of non-Gaussian GARCH models, including implications for option pricing. Christoffersen et al. (2004) explore inverse Gaussian shocks, while Bates & Craine (1999) and Maheu & McCurdy (2004) estimate GARCH-jump models that mimic jump-diffusions.

Below are some stylized facts from the GARCH literature that are relevant when pricing options.

1) Spot variance mean-reverts, with a typical half-life from univariate models estimated on daily data of 0.5 to 2 months.
2) Multivariate models improve upon univariate models. The long-run component in a component GARCH model is highly persistent, with a near unit root, while the short-run component typically has a half-life of days.
3) GARCH models of stock market volatility exhibit leverage, for both short- and long-run volatility in multivariate models such as Christoffersen et al. (2009).
4) Time-varying volatility accounts for much but not all of the unconditional leptokurtosis of daily asset returns. Daily asset returns are also conditionally leptokurtic and possibly skewed.

The first two points are relevant for modeling the term structure of at-the-money implied volatilities, which are downwardly biased relative to the square root of risk-neutral expected future variance. Symmetric GARCH models generate leptokurtic distributions and implied volatility smiles at multiday horizons via volatility randomization, while leverage generates skewed distributions and implied volatility smirks.

GARCH options research by its nature tends to examine the option pricing implications of GARCH models fitted to time series data on underlying asset returns. Babaoglu et al. (2018) conclude that volatility risk aversion, multifactor volatility models and fat-tailed shocks all play important roles in matching observed stock index option prices.

### 2.2 Estimates of continuous-time stochastic volatility models

Continuous-time models of the evolution of asset prices or of underlying fundamentals are of interest in finance for issues such as the optimal lifetime consumption/saving strategy, asset market equilibrium, and the pricing of bonds and options. Affine models can generate closed-form solutions to these issues, making them a popular modeling approach. There is substantial flexibility in affine models, but also some significant constraints when pricing options. All variances and covariances of diffusive shocks in Equation 2 must be linear functions of state variables. Jump intensities must be linear in state variables, while jump distributions must be invariant. Spot variance jumps cannot be negative, to preclude negative values.

Estimating continuous-time stochastic volatility models on discrete-time asset return data and using them to pricing options typically involves two numerically intensive steps:

1) Estimating the parameters of a postulated structure of asset price evolution (Equation 1);
2) Using Bayesian filtration (Equation 3) to estimate current state variable values $\boldsymbol{Y}_t$ that are inputs to the option pricing formula $O(S_t, \boldsymbol{Y}_t; \tau, X)$.

Melino & Turnbull (1990) estimate a stochastic volatility process for the Canadian dollar/U.S. dollar exchange rate via GMM, and use an extended Kalman filter to estimate current spot volatility and price FX options. Chernov et al. (2003) use an EMM/SNP approach that builds upon GARCH estimation procedures to estimate an extensive assortment of affine and non-affine models of stock market evolution over 1953-99. Eraker et al. (2003) use Bayesian Monte Carlo Markov Chain (MCMC) for the first step (for stock indices), while Johannes et al. (2009) use a particle filter for the second step. I use robust Kalman filtration in Bates (2006, 2012, 2019) to estimate and filter affine stochastic volatility models of stock market evolution without and with jumps. Christoffersen et al. (2010) use a particle filter approach to estimate six affine and non-affine stochastic volatility models of stock market evolution, without and with jumps.

Continuous-time stochastic volatility models differ from the discrete-time GARCH approach in a focus on diffusive versus jump specifications – primarily for the asset price process, but increasingly for the evolution of underlying state variables as well. Jumps capture outliers; but what counts as an outlier relative to a Gaussian benchmark depends upon the data frequency and the level of conditional volatility. A daily absolute movement more than 5% in the S&P 500 would normally be considered an outlier, occurring only 0.3% of the time over 1962-2020; but is much more likely if daily conditional volatility is at 3% than at its unconditional level of 1%. Intradaily market movements of 0.5 – 2% are considered jumps at the five-minute frequency,[7] while Barro & Ursúa (2008, 2012) define consumption disasters as a drop of more than 10% peak-to-trough over a two- to four-year window.

[7] See Tauchen & Zhou (2011) for estimates of intradaily jump magnitudes.

Table 1 contains examples of some continuous-time stochastic volatility models without and with jumps that have been used to price options and/or estimated on stock market data. Estimates on daily data broadly agree with GARCH estimates that stock market volatility is mean-reverting and has shocks that are negatively correlated with market returns. Using a central-tendency model, Bates (2012) finds results similar to component GARCH: spot variance mean-reverts to a central tendency with a half-life of a week, while the central tendency over 1926-2006 has a half-life of a year. Furthermore, the central-tendency model fits the term structure of at-the-money implied volatilities better than univariate models, although substantial divergences remain.

**[Table 1 about here]**

Jumps and stochastic volatility interact in complicated fashions that can affect the estimation of both. Estimates may also be affected by a bias towards tractability: affine versus non-affine, as well as which affine models are considered. Univariate affine models must impose time-invariant jump distributions, although the jump intensity can vary over time.[8] Furthermore, affine stochastic volatility models have known conditional characteristic functions in only two cases:

1) constant- or stochastic-intensity price jumps without volatility jumps (e.g., Bates 1996, 2000);
2) constant-intensity price/volatility cojumps (Duffie et al. 2000).

Stochastic-intensity price/volatility cojumps – a structure implying self-exciting volatility surges – lack a known conditional characteristic function.

In Bates (2006), I find stochastic-intensity price jump models substantially outperform constant-intensity processes for daily stock market returns, but have difficulty matching the 1987 stock market crash. In Bates (2012), I explore alternate stochastic-intensity Lévy specifications, again without volatility jumps. I find little difference between the various fat-tailed specifications regarding time series fit and implications for pricing options; and matching the 1987 crash is again an issue. In Bates (2019) I estimate a stochastic-intensity price/volatility cojump model on intradaily (15-minute) and overnight returns. The most general three-factor model is able to capture daily stock market movements exceeding 10% in magnitude in 1987 and 2008 as runs of self-exciting but small intradaily price/volatility cojumps.

[8] Multivariate models allow distributional variation by allowing the frequency of big jumps versus small jumps to vary over time.

I suggested stock market volatility might jump in a working paper presented at the 1995 WFA convention, based upon implausibly large movements in $V_t$ estimates inferred from option prices.[9] Eraker et al. (2003) find strong evidence for constant-intensity price/volatility cojumps using MCMC methods on daily stock market data spanning the 1987 crash. (The strongest evidence, of course, is the large daily stock market movements immediately following the crash). Aït-Sahalia et al. (2015) and Fulop et al. (2015) find evidence of self-exciting price/volatility cojumps in daily stock market data.

Affine models are tractable but may be seriously constraining parameter estimates. Chernov et al. (2003) find that leptokurtosis in daily asset returns can be explained either by price jumps or by explosive but mean-reverting behavior in a non-affine diffusive stochastic volatility component of a multifactor model. (The estimates in Bates (2019) on intradaily data support the latter explanation.) Christoffersen et al. (2010) estimate six models of diffusive spot variance evolution on daily stock market data. They conclude the volatility of diffusive spot variance shocks is proportional to its level, as opposed to the square-root restriction of affine models. They find this conclusion is robust to adding constant-intensity price jumps, but do not explore volatility jumps or price/volatility cojumps as alternate models of the volatility of volatility.

I view jump processes as a representation of outliers at a specific data frequency, rather than literal continuous-time truth. Big stock market movements at a daily frequency are typically the aggregation of smaller intradaily movements. The stock market crash on October 19, 1987 did not occur within five minutes, for instance; it took all day for the S&P 500 index to fall 23%. Similarly, while daily stock market movements of -12.3%, -10.2% and +12.5% on October 28-30, 1929 were dramatic precursors to the Great Depression, it was over the subsequent 2.7 years (through June 8, 1932) that the S&P 90 fell 81%, accompanied by a major increase in market volatility. Lower-frequency leverage, possibly in the form of self-exciting price/volatility cojumps for a more persistent volatility component, may be a better description of disaster risk than the stochastic-intensity rare-events approach of Gabaix (2008) and Wachter (2013).

## 3. Risk-neutral processes inferred from option prices

Pricing options under jumps and/or stochastic volatility benefitted greatly from Heston's (1993)

---

[9] The paper included estimates of a constant-intensity jump-diffusion with exponentially distributed jumps for the evolution of implicit $V_t$ – the process subsequently expanded into a cojump model by Duffie et al. (2000). I omitted those estimates from the Bates (2000) published version because a stochastic-intensity model was used to infer $V_t$.

Fourier inversion approach to pricing European options; there are many more distributions with analytical characteristic functions than with analytical probability density functions. Option pricing evolved from models without price jumps (Black & Scholes 1973; Merton 1973) to constant-volatility models with constant-intensity jumps (Merton 1976; Carr et al. 2002) to stochastic-intensity jumps (Bates 2000, 2012; Carr et al. 2003). While Hull & White (1987) and Scott (1987) use a non-affine diffusive log variance process without leverage, the affine diffusive stochastic volatility model (with leverage) of Hull & White (1988) and Heston (1993) was followed by diffusive stochastic volatility plus constant-intensity price jumps (Bates 1996) or stochastic-intensity jumps (Bates 2000). Duffie et al. (2000) introduce a constant-intensity cojump model in which spot variance also jumps, synchronously and in correlated fashion with price jumps. Andersen et al. (2015), Carr & Wu (2017) and Bates (2019) have cojump models with a self-exciting jump intensity. Early models use univariate specifications of spot variance evolution; more recent models are multivariate.

Because implied volatilities (IV's) are proxies for option prices, it is useful to understand which distributional features generate which IV patterns. Backus et al. (1997) use a Gram-Charlier expansion to approximate IV patterns at maturity $\tau$ as

$$IV(d_1)\sqrt{\tau} \approx SD\left[1 - \frac{SKEW}{3!} d_1 + \frac{XKURT}{4!}(d_1^2 - 1)\right], \qquad 4.$$

where

$$d_1 = \frac{\ln(F/X) + ½SD^2}{SD} \approx -\frac{\ln X - \ln F}{SD} \qquad 5.$$

measures how far options are in- or out-of-the-money in standard deviation units. Gram-Charlier approximations are not ideal – they can generate negative probability densities in the tails – but the approximation illustrates the broad relationship of IV patterns to the finite-horizon risk-neutral moments of log-differenced asset prices. Excess kurtosis generates a U-shaped "smile" pattern whereas skewness tilts the pattern, creating a "smirk". Symmetric stochastic volatility and jumps generate excess kurtosis but not skewness (the basis for the skewness premium of Bates 1991, 1997), whereas leverage and asymmetric jumps generate skewness. Finite-variance processes with constant diffusive volatility and constant-intensity jumps have skewness and excess kurtosis proportional to $1/\sqrt{\tau}$ and $1/\tau$ respectively, counterfactually predicting flatter IV patterns at longer maturities.

An alternate approach is in Aït-Sahalia et al. (2021a, b), who relate the properties of Taylor expansion representations of the IV surface $IV(X/S_t, \tau|V_t)$ conditional on diffusive spot variance $V_t$ to the instantaneous risk-neutral joint evolution of the asset price and $V_t$. (At-the-money implied volatilities converge to $\sqrt{V_t}$ as $\tau$ approaches zero, rather than to total spot volatility $\sqrt{V_t + \lambda_t^* E_t^*(\gamma^2)}$ .) Those properties are strongly affected by price jumps; models without jumps imply conditionally Gaussian returns as $\tau$ approaches zero, and an instantaneously flat IV pattern. Models with finite-activity jumps are better represented by Taylor expansions in $\sqrt{\tau}$ rather than in $\tau$. The overall patterns are largely as discussed above (e.g., volatility smirks can arise from leverage or from asymmetric jumps), but the approach permits scrutiny of the relationship between the level and evolution of $V_t$ implicit in option prices without imposing substantial model structure.

Structural models generate option prices $O(S_t, \boldsymbol{Y}_t; \tau, X; \boldsymbol{\Theta})$ given asset price $S_t$, underlying state variables $\boldsymbol{Y}_t$, maturity $\tau$, strike price $X$, and parameters $\boldsymbol{\Theta}$. While early papers such as Bates (1991) and Bakshi et al. (1997) estimate the parameters $\boldsymbol{\Theta}_t$ each day essentially as a form of data description, it has become more common since Bates (1996) to impose constant parameters and rely on implicit state variables $\widehat{\boldsymbol{Y}}_t$ to match observed option prices. Taylor & Xu (1994) and Bates (2000) use two-factor models of volatility evolution to price currency and stock index options, respectively; more recent papers that infer state variables from option prices use three-factor models. One criterion of success is how closely the models fit option prices or implied volatilities.

Andersen et al. (2015) examine various stochastic-intensity price/volatility cojump models, using two alternative structures to price stock index options:

1) Asset price jumps $\gamma = \rho\gamma_V + \gamma_s$, where $\gamma_V$ is an exponentially distributed jump in diffusive spot variance and $\gamma_s$ is an independent Gaussian or double-exponential shock; or
2) Spot variance jumps $\gamma_V = \mu_1\gamma_-^2 + \mu_2\gamma_2^2$, where $\gamma_-$ is the negative component of a double-exponential asset price jump and $\gamma_2$ is an independent exponentially distributed shock.

The issue is constructing a negative correlation between volatility jumps and log price jumps, while ensuring the former is strictly positive to preclude negative volatility. The former structure includes a stochastic-intensity version of the Duffie et al. (2000) cojump model; the latter is GARCH-like in going from asset price jumps (squared) to spot variance jumps. The latter approach has a more aggressive volatility response to asset price jumps through the squared-exponential structure, and fits option prices better. The most general three-factor model achieves an overall RMSE for IV's of 1.71%. The authors note that their third "U" factor captures movements in the left tail of the risk-neutral distribution (especially during crises), is strongly correlated with the Hu et al.'s (2013) "noise factor" measure of financial stress, and is especially useful in predicting stock market returns.

Carr & Wu (2017) model the diffusive spot variance and jump intensities underlying total asset value as evolving independently; the former as a diffusion, the latter as a self-exciting pure-jump process. Shocks to both are negatively correlated with asset price shocks. Their model also has a third factor to capture equity leverage. Implicit jump intensities were high during various crisis periods (especially 2008-9), suggesting it is capturing the same phenomena as Andersen et al.'s U factor.

Gruber (2015) uses the Wishart process of Bru (1991) to separate but interconnect diffusive variance and jump intensity factors in a matrix affine structure. Both must of course be nonnegative, which additive models achieve by positively weighted sums of underlying nonnegative processes such as the square-root diffusion. Wishart processes generalize additive and concatenated models by working with a matrix $\boldsymbol{X}_t$ that is always positive definite because of the positive definite structure of its shocks. Gruber then specifies diffusive spot variance in terms of the strictly positive diagonal elements of $\boldsymbol{X}_t$, while allowing the off-diagonal element $X_{12t}$ (which can and does go negative) to influence the jump intensity.

The above three papers agree that specifying partially independent evolution of diffusive spot variance and jump intensities is needed to match observed prices of stock index options. Variation in jump intensities captures variation in prices of OTM put options that protect against downside risk – especially major put price increases observed during the 2008-9 financial crisis.

## 4. Option price evolution

Various authors use the evolution of state variables inferred from option prices to describe how option prices evolve, and to test for compatibility with option pricing patterns. Taylor & Xu (1994) find the term structure of at-the-money implied volatilities from four PHLX currency options is broadly compatible with a central-tendency model describing how implied volatilities evolve. (Their option pricing model postulates a unit root for the central tendency, whereas their time series estimates find persistent but stationary processes with near unit roots.) Bates (1996, 2000) and Bakshi et al. (1997) focus more on the volatility of volatility – a key parameter along with leverage for generating volatility smiles and smirks in the absence of price jumps, and one that is common to the objective and risk-neutral processes in Equations

1 and 2. The papers found that "vol-of-vol" implicit parameter estimates run too high relative to the observed volatility of implicit spot variance evolution, especially when price jumps are precluded in the option pricing models. The latter two papers also present evidence of apparent jumps in volatility state variables inferred from index option prices.

An alternate and roughly model-free implicit state variable from stock index options is the VIX. Squared VIX is proportional to the price of a portfolio of put and call options that approximately replicates the log contract of Neuberger (1994). The log contract if delta-hedged approximately pays off the realized variance of the underlying asset, thereby making the price of the replicating portfolio a risk-neutral valuation of expected future realized variance (Demerfi et al. 1999). The VIX has gained widespread acceptance as an options-based assessment of future stock market volatility.[10] The Chicago Board Options Exchange has offered VIX futures since March 2004 and VIX options since February 2006, further motivating research into how the VIX evolves.

Dotsis et al. (2007) explore various univariate models of VIX evolution, as well as German and Italian equivalents, and conclude that jump-diffusion models with both positive and negative jumps capture their evolution better than diffusive models. Amengual & Xiu (2018) find the negative jumps typically coincide with FOMC announcements or speeches by Fed chairmen.[11] The VVIX measure of VIX volatility constructed from VIX options is only weakly correlated with the VIX itself (Huang et al. 2019), indicating multivariate rather than univariate models of VIX evolution are necessary. Amengual & Xiu (2018) fit a two-factor model to the S&P 500 and the term structure of variance swaps, while Bardgett et al. (2019) fit a three-factor model to S&P 500 index options and VIX options.

## 5. Option returns

Because we have almost a century of value-weighted U.S. stock market returns (and 136 years of the Dow Jones Industrial Average), we are reasonably confident that the U.S. stock market outperforms short-term money market instruments such as Treasury bills by about 6% per annum, with a standard error of 1.6%. We lack comparable histories of stock option returns, which have traded on centralized exchanges such as the CBOE only since 1973. Stock index and currency options began trading on U.S. exchanges

---

[10] Jiang & Tian (2005) explore replication issues for the VIX, as well as its biases when forecasting realized volatility. Andersen et al. (2015) highlight the sensitivity of the VIX to the strike price range of the options used in its construction, and propose a "corridor implied volatility" alternative.

[11] Patell & Wolfson (1979) similarly find that stocks' implicit volatilities increase up until earnings announcements and drop substantially thereafter.

in 1983 and 1984, at the CBOE and PHLX respectively. Furthermore, the stock market crash of 1987 led to major new constraints in how options could be traded, making pre- and post-1987 data not comparable. The popular OptionMetrics data base begins in January 1996. Options as levered investments in the underlying assets have high volatility and severe skewness and leptokurtosis, further complicating the statistical reliability of average option returns.[12]

Figure 1a gives the cumulative weekly returns of unhedged American put and call options on S&P 500 futures over 1988-2020, based on end-of-day settlement prices. All positions are positive-beta given short-put and long-call positions. The graph illustrates two stylized facts for stock index options:

1) Selling at-the-money (ATM) or out-of-the-money (OTM) put options has greatly outperformed the S&P 500, with greater returns from selling deeper OTM puts.
2) Buying ATM or OTM call options has underperformed the market.

**[Figures 1a,b about here]**

Option risk premia can be written as

$$E_t\left(\frac{dO_t}{O_t}\right) - r_t dt = E_t\left(\frac{dO_t}{O_t}\right) - E_t^*\left(\frac{dO_t}{O_t}\right)$$

$$= \underbrace{\frac{S_t}{O_t}\frac{\partial O}{\partial S}\left[E_t\left(\frac{dS_t}{S_t}\right) - b_t\right]}_{\text{Equity risk premium}} + \underbrace{\frac{1}{O_t}\frac{\partial O}{\partial \boldsymbol{Y}'}\left[E_t(d\boldsymbol{Y}_t) - E_t^*(d\boldsymbol{Y}_t)\right]}_{\boldsymbol{Y}_t \text{ risk premia}} + \underbrace{\frac{1}{O_t}(E_t - E_t^*)HOT_t}_{\text{jump risk premia}} \qquad 6.$$

where $HOT_t$ collects higher-order terms in the Taylor expansion of $dO(S_t, \boldsymbol{Y}_t; \tau, X)$. Diffusive risk contributions to $HOT_t$ are either identical or negligible under $E_t$ and $E_t^*$, so the last term represents compensation for jump risk in $S_t$ and/or $\boldsymbol{Y}_t$. Delta-hedged option returns focus attention on the last two components of Equation 6: the jump risk premia and those parts of $\boldsymbol{Y}_t$ risk premia that are not linearly attributable to equity risk. While delta-hedged returns often use Black-Scholes-Merton (BSM) deltas evaluated at the option's implied volatility, such estimates are biased when there is a volatility smirk (Bates 2005), and do not take into account the negative correlation between market returns and

[12] Broadie et al. (2009) discuss these issues, using simulated put returns. Chambers et al. (2014) dispute their results for unhedged put returns but concur with their results for delta-hedged put returns.

implied volatility changes. Figure 1b consequently shows cumulative weekly returns from short put and call positions using delta-hedged returns

$$\frac{\Delta O}{O} - \left(\frac{F_t}{O}\frac{\partial O^{BSM}}{\partial F} + c\right)\frac{\Delta F}{F} \qquad 7.$$

where $c$ is a correction term (the regression coefficient of BSM-based delta-hedged returns on futures returns) that makes hedged options returns delta-neutral (and zero-beta) over 1988-2020.

Figure 1b reinforces the anomalies graphed in Figure 1a. Delta-hedged short put positions were reliably profitable over 1988-2017, with occasional sharp weekly losses but no sustained drawdowns until 2018-20. Selling puts and delta-hedging generated Sharpe ratios that were 2.8 to 3.7 times that of the market over 1988-2017, although poor performance thereafter (including during the pandemic) lowered relative performance to two to three times that of the market over 1988-2020. Selling OTM call options and delta-hedging was also profitable, albeit less so – an anomaly highlighted by Coval & Shumway (2001) and Bakshi & Kapadia (2003). These authors speculate that the substantial negative excess returns on hedged put and call positions represent required compensation for volatility risk.

### 5.1 Implicit pricing kernels

A conceptual building block for assessing option prices and returns is the butterfly spread, which can be created from European calls or puts such as the SPX options on the S&P 500. Butterfly spreads approximate the Arrow-Debreu claims on future $S_{t+\tau}$ outcomes discussed in Breeden & Litzenberger (1978). $f(X_i)/\Delta X$ units of a butterfly spread constructed from options with strike spacing $\Delta X$ replicate an arbitrary payoff function $f(S_{t+\tau})$ at the point $S_{t+\tau} = X_i$, while a portfolio of such butterfly spreads creates a continuous, piecewise linear replication of the payoff function over the available strike price range. Options with the two lowest and two highest strike prices can be added to linearly extrapolate the approximate payoff functions below and above the available strike price range.[13]

Correspondingly, the market for European options is a market for butterfly spreads. European put (call) spreads held to maturity are equally-weighted bundles of butterfly spreads below (above) the relevant strike price, with returns determined by the returns on those butterfly spreads. European put and

[13] This is not standard practice. Andersen et al. (2015) note that the VIX replication portfolio discards all deeper out-of-the-money calls and puts once a zero bid price is observed.

call options similarly represent portfolios with linearly increasing quantities of butterfly spreads deeper out of the money.

The implicit pricing kernel literature surveyed in Cuedeanu & Jackwerth (2018) essentially examines returns of butterfly spreads on the S&P 500 across different strike prices. Wealth-based models such as the CAPM imply the pricing kernel is inversely related to the stock market return. This monotonicity hypothesis implies conditional expected excess returns on butterfly spreads should be inversely related to strike price: negative for low $X/S_t$, positive for high $X/S_t$.[14] Unconditional expected excess returns should therefore also be inversely related to $X/S_t$ – a hypothesis testable using sample excess returns. OTM European puts or calls that load heavily on low- or high-strike butterfly spreads should exhibit negative or positive excess returns respectively if held to maturity. Furthermore, calls and call spreads load more heavily on higher-strike butterfly spreads as the strike price increases and should therefore exhibit ever-higher excess returns – a result empirically rejected in Bakshi et al. (2010).

While conceptually straightforward, directly computing average excess returns on butterfly spreads is not. Early key papers by Jackwerth (2000) and Aït-Sahalia & Lo (2000) compute average payoff and average price separately: the former from estimates of the unconditional density of stock market returns, the latter from average risk-neutral densities from option prices. Both papers find near-the-money non-monotonicities in the implicit pricing kernel. Rosenberg & Engle (2002) estimate conditional densities from a GARCH model with conditionally Gaussian shocks plus leverage, and examine what distributional transformation of that conditional density best matches observed option prices. They also find near-the-money non-monotonicities in the implicit pricing kernel.

Early studies of implicit pricing kernels had limited options data to work with. Aït-Sahalia & Lo (2000) compare risk-neutral densities over 1993 (when the VXO measure of $\sigma^{ATM}$ from 1-month American S&P 100 options averaged 12.65) with stock market returns over 1989-93 (when the VXO averaged 17.59). Jackwerth (2000) uses monthly returns over the four years preceding his 1988-91 options data. Because stock market volatility exhibits long-memory phenomena, with periods of high or low volatility that can last for years, reexamination of pricing kernel puzzles using longer and better-matched data appears warranted.

---

[14] Because an equally weighted portfolio of butterfly spreads over $[0, \infty)$ replicates a riskless payoff, its conditional expected return must be the riskless rate. Binsbergen et al. (2021) confirm this using box spreads (overlapping call and put spreads with identical strikes), with an implicit interest rate from midpoint SPX option prices that is typically slightly below LIBOR.

Linn et al. (2018) follow Bliss & Panigirtzoglou (2004) in examining what transform of the conditional risk-neutral distribution from options best matches subsequent monthly stock market returns – the inverse of the approach in Rosenberg & Engle (2002).[15] They note their cubic spline approach for the transform is similar to looking at returns on butterfly spreads (smoothed across adjacent strikes), and find no evidence of near-the-money non-monotonicity in the implicit pricing kernel for monthly returns over 1996-2014.

A further complicating issue is asymmetric depth for puts versus calls on the stock index. Put options have been more actively traded since the stock market crash of 1987 (Bates 2000, Figure 3), and are readily available three or more standard deviations out-of-the-money, as measured by $\sigma^{ATM}\sqrt{\tau}$. Calls by contrast are thinly traded beyond one standard deviation out-of-the-money (Andersen et al. 2015, Table 1). We consequently have to rely on calls and call spreads as bundles of butterfly spreads when assessing the implicit pricing kernel for upside risk.

Point estimates of the implicit pricing kernel do appear to be non-monotonic, with the major caveat that standard errors are large. In particular, average returns in Bakshi et al. (2010) decrease and go negative for stock index calls and call spreads further out-of-the-money that load more heavily on deep OTM butterfly spreads. The implication is negative rather than positive returns on high-strike butterfly spreads. This combined with the negative returns observed for OTM put options (indicating negative returns on low-strike butterfly spreads) and the fact that butterfly spreads averaged across all strikes earn the riskless rate suggest a non-monotonic and broadly U-shaped pricing kernel.

## 6. Interpreting option returns

There are two competing theoretical frameworks for the structure of supply and demand in option markets. These two frameworks offer alternative interpretations of the empirical evidence that OTM options on stock indexes appear to be overpriced – especially put options.

**Framework #1**: **Equity and options markets are integrated.** Equity investors can freely buy or sell index options to manage their equity exposures. Agents may be heterogeneous; they may have different aversions to or beliefs about downside risk. However, their ability to directly trade OTM put options leads to market-clearing prices for those options. Those with strong aversions to downside risk can buy downside protection from those less averse to or less pessimistic about downside risk. Resulting

[15] An early example of this approach is Fackler & King (1990), who examine commodity options.

option prices and risk-neutral distributions reflect average ("representative") preferences or beliefs about future stock market outcomes, plus required compensation for those risks.

Identifying equilibrium in this framework is easier if markets are dynamically complete, with options spanning systematic volatility and jump risks of concern to investors. A complete-markets framework implies the dynamic competitive equilibrium can be identified from fundamental underlying risks by solving an associated central planner's problem.[16] The "representative agent" can be identified by a social utility function that is a wealth-sensitive positive weighting of individual utility functions. The representative agent shares properties with individual utility functions: expected utility if individual utilities are von Neumann-Morgenstern, concave if individual utilities are concave (Constantinides 1982; Ziegler 2007). The representative agent's preferences change over time because of wealth redistributions across heterogeneous agents, generally in a nonstationary fashion.[17] Equilibrium prices of all assets change in response, in a fashion that magnifies underlying shocks to fundamentals. Positive news about future cash flows increases stock prices both directly and because the average investor becomes less risk-averse, because of wealth transfers to risk-tolerant investors who have invested more heavily in equity.

**Framework #2**: **Options markets are partly segmented from the underlying equity markets.** Equity investors can easily buy OTM index puts, but trading restrictions instituted after the 1987 stock market crash prevent most investors from selling puts.[18] Financial intermediaries (especially option market makers) play a central role; they willingly sell overpriced OTM puts to equity investors especially averse to downside or volatility risks. Intermediaries can delta-hedge against small market movements (with spillover consequences for equity markets) but are perforce exposed in aggregate to jump and volatility risks. Equity investors less averse to downside risk who might also be willing to sell OTM puts are shut out of the options market by the trading restrictions. Options prices therefore do not represent the average preferences and beliefs of most equity investors, and are only partly relevant for the resulting equilibrium in equity markets.

---

[16]See Grossman and Zhou (1996) and Weinbaum (2009) for heterogeneous-agent models with one diffusive source of risk, and Bates (2008) for a heterogeneous-agent model with an additional jump risk that is spanned by options.

[17] See Bates (2008). Chan & Kogan (2002) show that models with external habit formation can have a stationary wealth distribution across heterogeneous agents, generating a stationary asset market equilibrium.

[18] Individual investors could still buy calls or puts relatively easily after 1987, but "sophisticated" or "accredited" investor status (including adequate income or net wealth) was required to sell uncovered puts or calls.

The fundamental theorem of asset pricing states that we have $E_t\left[M_{t,\tau}(R_{t+\tau} - R_{ft})\right] = 0$ under both frameworks, but the interpretation of the pricing kernel $M_{t,\tau}$ differs. In particular, an equilibrium relationship between individuals' pricing kernels and excess returns applies only to unconstrained agents. The first framework assumes all agents are unconstrained, who trade until the product of state-contingent marginal utilities and subjective probabilities is proportional across all agents. Risk dispersal permits calibration of such models based upon aggregate wealth or aggregate consumption of all agents, given wealth-weighted aggregate demand for options in zero net supply must equal zero.

The second framework by contrast mostly focusses upon financial intermediaries as key unconstrained agents. Their willingness to sell OTM puts is affected by how many puts they have already sold, and the aggregate wealth of the financial sector. This framework has been explored in demand-based option pricing papers such as Bollen & Whaley (2004) and Garleanu et al. (2009). The latter paper shows that option market makers have indeed been net sellers of puts and calls on stock indexes – especially puts.

In framework #1, overpriced OTM puts and calls on stock indexes are interpreted either as direct evidence of average investor preferences, or evidence of additional risks of concern to investors. Rosenberg & Engle (2002) point out the implicit pricing kernel is the expected pricing kernel conditional upon observing specific stock market returns. High values of the pricing kernel (negative returns on butterfly spreads) can be either because of direct wealth effects or because the returns are proxies for other risks affecting the pricing kernel. In this framework, OTM puts are overpriced both because of strong aversion to downside market risk and because negative market returns are correlated with crash risk or volatility risk. OTM calls are overpriced either for behavioral reasons such as a long-shot bias by investors (Hodges et al. 2004), or because large positive returns are correlated with high-volatility periods to which investors are averse. Christoffersen et al. (2012) show that aversion to volatility shocks can generate a U-shaped pricing kernel in a GARCH option pricing model, while Babaoglu et al. (2018) find volatility aversion substantially improves the fit of such models to observed option prices.

In framework #2, overpriced puts are attributed to selection bias among equity investors, with only those especially averse to downside risk or volatility risk participating (on the buy side) in the index options markets.[19] A major motivation for this framework is the apparent profit opportunities from selling OTM puts since the 1987 stock market crash. It seems easier to attribute the substantial profit

---

[19] Scheinkman and Xiong (2013) similarly argue that short-sales constraints can lead to overpriced stocks because pessimists are shut out of the market, leaving only optimists.

opportunities graphed in Figure 1 to difficulties in exploiting those opportunities rather than to rational concerns by investors about downside risk. In particular: why would anyone invest in the stock market when they can get two to three times the Sharpe ratio by selling delta-hedged puts?

Explaining overpriced OTM calls is trickier; selling covered calls is a substantially unrestricted strategy available to equity investors. However, Garleanu et al. (2009) find option market makers face net demand for calls. Given their substantial volatility and jump risk exposures from delta-hedged short put positions, market makers require volatility or jump risk premia to be willing to also sell calls.

Compensation for crash or volatility risk are two competing explanations for option return anomalies. Coval & Shumway (2001) attempt to disentangle the two by using OTM puts to partially hedge the crash risk from writing delta-neutral at-the-money straddles. They label such positions “crash-neutral”, and conclude observed excess returns are compensation for volatility risk rather than crash risk.[20] Bakshi & Kapadia (2003) also conclude volatility risk is the better explanation.

Crash and volatility risks should also affect investors’ assessments of underlying investments, and show up in the equity risk premium. Welch (2016) notes that a “crash-insured” equity investment that uses puts to limit monthly losses to at most 15% still earned 5.1% per year over 1983-2012, as opposed to the 7.2% return from unhedged equity. He concludes that disaster risk consequently cannot explain the equity premium. Jurek (2014) similarly notes that option-hedged strategies to exploit the carry trade remain profitable, ruling out currency crash risk as an explanation. Ang (2014, pp. 218-222) considers volatility risk (from both diffusive and jump sources) to be an “extremely important risk factor” that affects many investments.

It is difficult at this point to distinguish between volatility risk and jump risk, let alone to identify what precisely volatility risk represents. The empirical evidence from recent cojump models is that everything is correlated. Stock market jumps are typically synchronous and negatively correlated with jumps in the VIX and in time series estimates of diffusive volatility and jump intensities. Objective and risk-neutral measures of conditional volatility covary with major market drops that are clearly of concern to investors: the 2008-9 financial crisis, for instance, or the onset of the pandemic in February to March of 2020. Stock market volatility reached high and sustained levels during the Great Depression as the stock market fell

[20] A position closer to “crash-neutral” is the delta-gamma-neutral strategy, which involves considerably more OTM puts.

(and firm leverage rose). Volatility risk premia could consequently be compensation for related risks of concern to investors, such as major recessions or financial crises.

Estimates of volatility risk premia are based on divergences between objective and risk-neutral measures of stochastic processes, with the latter inferred from option prices. The alternate possibility from framework #2 that it reflects major overpricing of options in a market somewhat segmented from the underlying equity market should also be kept in mind. Such overpricing should be (and does appear to be) especially pronounced during financial crises, perhaps because reduced capital of financial intermediaries reduces the supply of options available to equity investors concerned about downside risk.

I noted in Bates (1991) that stochastic volatility with leverage and fears of asymmetric jumps are both possible explanations for periods of risk-neutral skewness observed prior to the stock market crash of 1987. While I focused more on the latter, I now consider the former more important. In particular, time aggregation of stock market returns appears to play a key role in downside risk. Major daily stock market movements are typically the accumulation of smaller intradaily price/volatility comovements, whether diffusive (Chernov et al. 2003) or via self-exciting cojumps (Bates 2019). Component GARCH models such as Christoffersen et al. (2009) suggest similar mechanisms at longer horizons. These forms of longer-term leverage appear implicit in option prices as well: risk-neutral skewness from stock index options increases rather than decreases with option maturity at short horizons. Runs of bad shocks that lead to major cumulative losses – against which put-based portfolio insurance offers only a partial defense – are plausibly of greater concern to investors than occasional major outliers in daily returns. The degree to which required compensation for this type of risk is responsible for observed returns on stock index options and for the equity premium is an open question.

## References

Aït-Sahalia Y, Cacho-Diaz J , Laeven RJA. 2015. Modeling financial contagion using mutually exciting jump processes. *J. Financ. Econ.* 117:585-606

Aït-Sahalia Y, Li C, Li CX. 2021a. Closed-form implied volatility surfaces for stochastic volatility models with jumps. *J. Econ.* 222:364-392

Aït-Sahalia Y, Li C, Li CX. 2021b. Implied stochastic volatility models. *Rev. Financ. Stud.* 34:394-450

Aït-Sahalia Y, Lo AW. 2000. Nonparametric risk management and implied risk aversion. *J. Econ.* 94:9-51

Alizadeh S, Brandt MW, Diebold FX. 2002. Range-based estimation of stochastic volatility models. *J. Finance* 57:1047-1091

Amengual D, Xiu D. 2018. Resolution of policy uncertainty and sudden declines in volatility. *J. Econ.* 203:297-315

Andersen TG, Bondarenko O, Gonzalez-Perez MT. 2015. Exploring return dynamics via corridor implied volatility. *Rev. Financ. Stud.* 28:2902-2945

Andersen TG, Fusari N, Todorov V. 2015. The risk premia embedded in index options. *J. Financ. Econ.* 117:558-584

Ang A. 2014. *Asset Management: a systematic approach to factor investing.* Oxford: Oxford University Press

Babaoglu K, Christoffersen P, Heston SL, Jacobs K. 2018. Option valuation with volatility components, fat tails, and nonmonotonic pricing kernels. *Rev. Asset Pricing Stud.* 8:183-231

Backus DK, Foresi S, Li K, Wu L. 1997. Accounting for biases in Black-Scholes. New York University working paper

Bakshi G, Cao C, Chen Z. 1997. Empirical performance of alternative option pricing models. *J. Finance* 52:2003-2049

Bakshi G, Kapadia N. 2003. Delta-hedged gains and the negative market volatility risk premium. *Rev. Financ. Stud.* 16:527-566

Bakshi, Gurdip, Madan, Dilip B., and Panayotov, George. Returns of claims on the upside and the viability of U-shaped pricing kernels. *J. Financ. Econ.* 97, 130-154. 2010

Bardgett C, Gourier E, Leippold M. 2015. Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *J. Financ. Econ.* 131:593-618

Barro RJ, Ursúa JF. 2008. Consumption disasters in the twentieth century. *Am. Econ. Rev.* 98:48-63

Barro RJ, Ursúa JF. 2012. Rare macroeconomic disasters. *Annu. Rev. Econ.* 4:83-109

Bates DS. 1991. The Crash of '87: was it expected? The evidence from options markets. *J. Finance* 46:1009-1044

Bates DS. 1996. Jumps and stochastic volatility: exchange rate processes implicit in PHLX deutsche mark options. *Rev. Financ. Stud.* 9:69-107

Bates DS. 1997. The skewness premium: option pricing under asymmetric processes. Advances in Futures and Options Research 9:51-82

Bates DS. 2000. Post-'87 crash fears in the S&P 500 futures option market. *J. Econ.* 94:181-238

Bates DS. 2005. Hedging the smirk. *Financ. Res. Lett.* 2:195-200

Bates DS. 2006. Maximum likelihood estimation of latent affine processes. *Rev. Financ. Stud.* 19:909-965

Bates DS. 2008. The market for crash risk. *J Econ Dyn Control* 32:2291-2321.

Bates DS. 2012. U.S. stock market crash risk, 1926-2010. *J. Financ. Econ.* 105:229-259

Bates DS. 2019. How crashes develop: intradaily volatility and crash evolution. *J. Finance* 74:193-237

Bates DS, Craine R. 1999. Valuing the futures market clearinghouse's default exposure during the 1987 crash. *J. Money Credit Bank.* 31:248-272

von Binsbergen JH, Diamond WF, Grotteria M. 2021. Risk-free interest rates. *J. Financ. Econ.* forthcoming

Black F. 1976. Studies of stock price volatility changes. Proceedings of the 1976 Meetings of the American Statistical Association 177-181

Black F, Scholes M. 1973. The pricing of options and corporate liabilities. *J. Political Econ.* 81:637-659

Bliss RR, Panigirtzoglou N. 2004. Option-implied risk aversion estimates. *J. Finance* 59:407-446

Bollen NPB, Whaley RE. 2004. Does net buying pressure affect the shape of implied volatility functions? *J. Finance* 59:711-753

Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. *J. Econ.* 31:307-327

Box GEP, Jenkins GM. 1970. *Time series analysis, forecasting and control*. San Francisco: Holden-Day

Breeden DT, Litzenberger RH. 1978. Prices of state-contingent claims implicit in option prices. *J. Bus.* 51:621-651

Brennan MJ. 1979. The pricing of contingent claims in discrete time models. *J. Finance* 34:53-68

Broadie MN, Chernov M, Johannes M. 2009. Understanding index option returns. *Rev. Financ. Stud.* 22:4492-4529

Bru M-F. 1991. Wishart processes. *J. Theor. Prob.* 4:725-751

Campbell JY, Hentschel L. 1992. No news is good news: an asymmetric model of changing volatility in stock returns. *J. Financ. Econ.* 31:281-318

Carr P, Geman H, Madan DB, Yor M. 2002. The fine structure of asset returns: An Empirical Investigation. *J. Bus.* 75:305-332

Carr P, Geman H, Madan DB, Yor M. 2003. Stochastic volatility for Lévy processes. *Math. Financ.* 13:345-382

Carr P, Wu L. 2017. Leverage effect, volatility feedback, and self-exciting market disruptions: disentangling the multi-dimensional variations in S&P 500 index options. *J. Financ. Quant. Anal.* 52:2119-2156

Carr P, Wu L. 2020. Option profit and loss attribution and pricing: a new framework. *J. Finance* 75:2271-2316

Chambers DR, Foy M, Liebner J, Lu Q. 2014. Index option returns: still puzzling. *Rev. Financ. Stud.* 27:1915-1928

Chan YL, Kogan L. 2002. Catching up with the Joneses: heterogeneous preferences and the dynamics of asset prices. Journal of Political Economy 110:1255-1285

Chernov M, Gallant AR, Ghysels E, Tauchen G. 2003. Alternative models for stock price dynamics. *J. Econ.* 116:225-257

Christie AA. 1982. The stochastic behavior of common stock variances: value, leverage and interest rate effects. *J. Financ. Econ.* 10:407-432

Christoffersen P, Heston SL, Jacobs K. 2004. Option valuation with conditional skewness. *J. Econ.* 131:253-284

Christoffersen P, Heston SL, Jacobs K. 2012. Capturing option anomalies with a variance-dependent pricing kernel. *Rev. Financ. Stud.* 26:1963-2006

Christoffersen P, Jacobs K, Ornthanalai C, Wang Y. 2009. Option valuation with long-run and short-run volatility components. *J. Financ. Econ.* 90:272-297

Christoffersen P, Jacobs K, Mimouni K. 2010. Models for S&P500 dynamics: evidence from realized volatility, daily returns, and option prices. *Rev. Financ. Stud.* 23:3141-3189

Constantinides GM. 1982. Intertemporal asset pricing with heterogeneous consumers and without demand aggregation. *J. Bus.* 55:253-267

Cont R, da Fonseca J. 2002. Dynamics of implied volatility surfaces. *Quant. Financ.* 2:45-60

Corsi F. 2009. A simple approximate long-memory model of realized volatility. *J. Financ. Economet.* 7:174-196

Coval JD, Shumway T. 2001. Expected option returns. *J. Finance* 56:983-1009

Cuesdeanu H, Jackwerth JC. 2018. The pricing kernel puzzle: survey and outlook. *Ann. Finance* 4:289-329

Demeterfi K, Derman E, Kamal M, Zou J. 1999. A guide to volatility and variance swaps. *J. Deriv* 6:9-32

Ding Z, Granger CWJ, Engle RF. 1993. A long memory property of stock market returns and a new model. *J. Empir. Financ.* 1:83-106

Dotsis G, Psychoyios D, Skiadopoulos G. 2007. An empirical comparison of continuous-time models of implied volatility indices. *J. Bank. Financ.* 31:3584-3603

Duan J-C. 1995. The GARCH option pricing model. *Math. Financ.* 5

Duffie D, Pan J, Singleton KJ. 2000. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68:1343-1376

Engle RF. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50:987-1007

Engle RF, Lee GGJ. 1999. A permanent and transitory component model of stock return volatility. In *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*, ed. RF Engle, H. White, pp. 475-497. Oxford, UK: Oxford University Press

Eraker B, Johannes M, Polson NG. 2003. The impact of jumps in volatility and returns. *J. Finance* 58:1269-1300

Fackler PL, King RP. 1990. Calibration of option-based probability assessments in agricultural commodity markets. *Am. J. Ag. Econ.* 72:73-83

Fama EF. 1965. The behavior of stock market prices. *J. Bus.* 38:34-105

Fulop A, Li J, Yu J. 2015. Self-exciting jumps, learning, and asset pricing implications. *Rev. Financ. Stud.* 28:876-912

Gabaix X. 2008. Variable rare disasters: a tractable theory of ten puzzles in macro-finance. *Am. Econ. Rev.* 98:64-67

Garleanu N, Pedersen LH, Poteshman AM. 2009. Demand-based option pricing. *Rev. Financ. Stud.* 22:4259-4299

Grossman SJ, Zhou Z. 1996. Equilibrium analysis of portfolio insurance. *J. Finance* 51:1379-1403

Gruber PH. 2015. *Essays on Variance Risk.* Università della Svizzera Italiana thesis

Hentschel L. 1995. All in the family: nesting symmetric and asymmetric GARCH models. *J. Financ. Econ.* 39:71-104

Heston SL. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.* 6:327-344

Heston SL, Nandi S. 2000. A closed-form GARCH option pricing model. *Rev. Financ. Stud.* 13:585-625

Hodges SD, Tompkins RG, Ziemba WT. 2004. The favorite/long-shot bias in S&P 500 and FTSE 100 index futures options: the return to bets and the cost of insurance. In *Handbook of Sports and Lottery Markets*, ed. DB Hausch, WT Ziemba, pp. 161-180. Amsterdam: Elsevier

Huang D, Schlag C, Shaliastovich I, Thimme J. 2019. Volatility-of-volatility risk. *J. Financ. Quant. Anal.* 54:2423-2452

Hull J, White A. 1987. The pricing of options on assets with stochastic volatility. *J. Finance* 42:281-300

Hull J, White A. 1988. An analysis of the bias in option pricing caused by stochastic volatility. Advances in Futures and Options Research 3:29-61

Jackwerth JC. 2000. Recovering risk aversion from option prices and realized returns. *Rev. Financ. Stud.* 13:433-451

Jiang GJ, Tian YS. 2005. Model-free implied volatility and Its information content . *Rev. Financ. Stud.* 18:1305-1342

Johannes M, Polson NG, Stroud J. 2009. Optimal filtering of jump-diffusions: extracting latent states from asset prices. *Rev. Financ. Stud.* 22:2759-2799

Jurek JW. 2014. Crash-neutral currency carry trades. *J. Financ. Econ.* 113:325-347

LeBaron BD. 2001. Stochastic volatility as a simple generator of apparent financial power laws and long memory. *Quant. Financ.* 1:621-631

Linn M, Shive S, Shumway T. 2018. Pricing kernel monotonicity and conditional information. *Rev. Financ. Stud.* 31:493-531

Litterman R, Scheinkman JA. 1991. Common factors affecting bond returns. *J. Fixed Income* 1:54-61

Maheu JM, McCurdy TH. 2004. News arrival, jump dynamics and volatility components for individual stock returns. *J. Finance* 59:755-793

Mandelbrot BB. 1963. The variation of certain speculative prices. *J. Bus.* 36:394-419

Melino A, Turnbull SM. 1990. Pricing foreign currency options with stochastic volatility. *J. Econ.* 45:239-265

Merton RC. 1973. Theory of rational option pricing. *Bell J Econ* 4:141-183

Merton RC. 1976. Option pricing when underlying stock returns are discontinuous. *J. Financ. Econ.* 3:125-144

Nelson DB. 1991. Conditional heteroskedasticity in asset returns:  a new approach. *Econometrica* 59:347-370

Nelson DB. 1992. Filtering and forecasting with misspecified ARCH models I: getting the right variance with the wrong model. *J. Econ.* 52:61-90

Neuberger A. 1994. The log contract. *J. Portfolio Manage.* 20:74-80

Patell JM, Wolfson MA. 1979. Anticipated information releases reflected in call option prices. *J. Account. Econ.* 1:117-140

Poterba J, Summers L. 1986. The persistence of volatility and stock market fluctuations. *Am. Econ Rev.* 76:1142-1151

Rosenberg J, Engle RF. 2002. Empirical pricing kernels. *J. Financ. Econ.* 64:341-372

Rubinstein M. 1976. The valuation of uncertain income streams and the pricing of options. *Bell J Econ* 7:407-425

Scheinkman, JA, Xiong W. 2013. Overconfidence and speculative bubbles.  *J. Polit. Econ.* 111:1183-1219

Scott LO. 1987. Option pricing when the variance changes randomly:  theory, estimation, and an application. *J. Financ. Quant. Anal.* 22:419-438

Shaliastovich I. 2015. Learning, confidence, and option prices. *J. Econ.* 187:18-42

Skiadopoulos G, Hodges S, Clewlow L. 1999. The dynamics of the S&P 500 implied volatility surface. *Rev. Deriv. Res.* 3:263-282

Smith CW. 1976. Option pricing: A review. *J. Financ. Econ.* 3:3-51

Taylor SJ. 1986. *Modelling Financial Time Series*. Chichester: John Wiley

Taylor SJ, Xu X. 1994. The term structure of volatility implied by foreign exchange options. *J. Financ. Quant. Anal.*  29:57-74

Wachter J. 2013. Can time-varying risk of rare disasters explain aggregate stock market volatility? *J. Finance* 68:987-1035

Welch I. 2016. The (time-varying) importance of disaster risk. *Financ. Anal. J.* 72:14-30

Ziegler A. 2007. Why does implied risk aversion smile? *Rev. Financ. Stud.* 20:859-904

**Table 1. Some models of variance dynamics**

| **Models** | **Jump Intensities** | **Examples** |
|---|---|---|
| **Univariate models of diffusive variance $V_t$** | | |
| $d \ln V_t = (\alpha - \beta \ln V_t)dt + \sigma_V dW_{Vt}$ | No jumps | Hull & White (1987), Scott (1987) |
| $dV_t = (\alpha - \beta V_t)dt + \sigma_V \sqrt{V_t} dW_{Vt}$ | No jumps | Hull & White (1988), Heston (1993) |
| | $\lambda_0$ | Bates (1996) |
| | $\lambda_t = \lambda_0 + \lambda_1 V_t$ | Bates (2000, 2006) |
| $dV_t = (\alpha - \beta V_t)dt + \sigma_V \sqrt{V_t} dW_{Vt} + \gamma_V dN_{Vt}$ | $\lambda_0$ | Duffie et al. (2000), Eraker et al. (2003) |
| | $\lambda_0 + \lambda_1 V_t$ | Andersen et al. (2015), Bates (2019) |
| $dV_t = \kappa V^a (\theta - V_t)dt + \sigma V_t^b dW_{Vt}$ | $\lambda_t = 0$ or $\lambda_0$ | Christoffersen et al. (2010) |
| **Multivariate models** (with diffusive variance $V_t = \sum_i V_{it}$) | | |
| additive<br>$dV_{it} = (\alpha_i - \beta_i V_{it})dt + \sigma_{Vi} \sqrt{V_{it}} dW_{Vt}$ | $\lambda_t = \lambda_0 + \boldsymbol{\lambda}' \boldsymbol{V}_t$ | Bates (2000) |
| self-exciting<br>$dV_{it} = (\alpha_i - \beta_i V_{it})dt + \sigma_{Vi} \sqrt{V_{it}} dW_{Vt} + \gamma_V dN_{Vt}$ | $\lambda_t = \lambda_0 + \boldsymbol{\lambda}' \boldsymbol{V}_t$ | Andersen et al. (2015), Bates (2019) |
| concatenated<br>$dV_t = \kappa(\theta_t - V_t) + \sigma_v \sqrt{V_t}\, dW_{Vt}$<br>$d\theta_t = \kappa(\bar{\theta} - \theta_t)dt + \sigma_\theta \sqrt{\theta_t}\, dW_{\theta t}$ | No jumps<br>$\lambda_t = \lambda_0 + \boldsymbol{\lambda}' \boldsymbol{V}_t$ | Taylor (1994)<br>Bates (2012) |
| Wishart<br>$\boldsymbol{X}_t = \begin{pmatrix} V_{1t} & X_{12t} \\ X_{12t} & V_{2t} \end{pmatrix}$<br>$d\boldsymbol{X}_t = [\beta \boldsymbol{Q}' \boldsymbol{Q} + \boldsymbol{M} \boldsymbol{X}_t + \boldsymbol{X}_t \boldsymbol{M}']dt + \sqrt{\boldsymbol{X}_t} d\boldsymbol{B}_t \boldsymbol{Q} + \boldsymbol{Q}' d\boldsymbol{B}_t' \sqrt{\boldsymbol{X}_t}$ | $\lambda_t = \lambda_0 + \boldsymbol{\Lambda} \cdot \boldsymbol{X}_t$ | Gruber (2015) |
| Non-affine<br>$d \ln V_{it} = (\alpha_{i0} + \alpha_{ii} \ln V_{it})\, dt + (\beta_{i0} + \beta_{ii} \ln V_{it})^{\gamma_i} dW_{Vit}$ | | Chernov et al. (2003) |

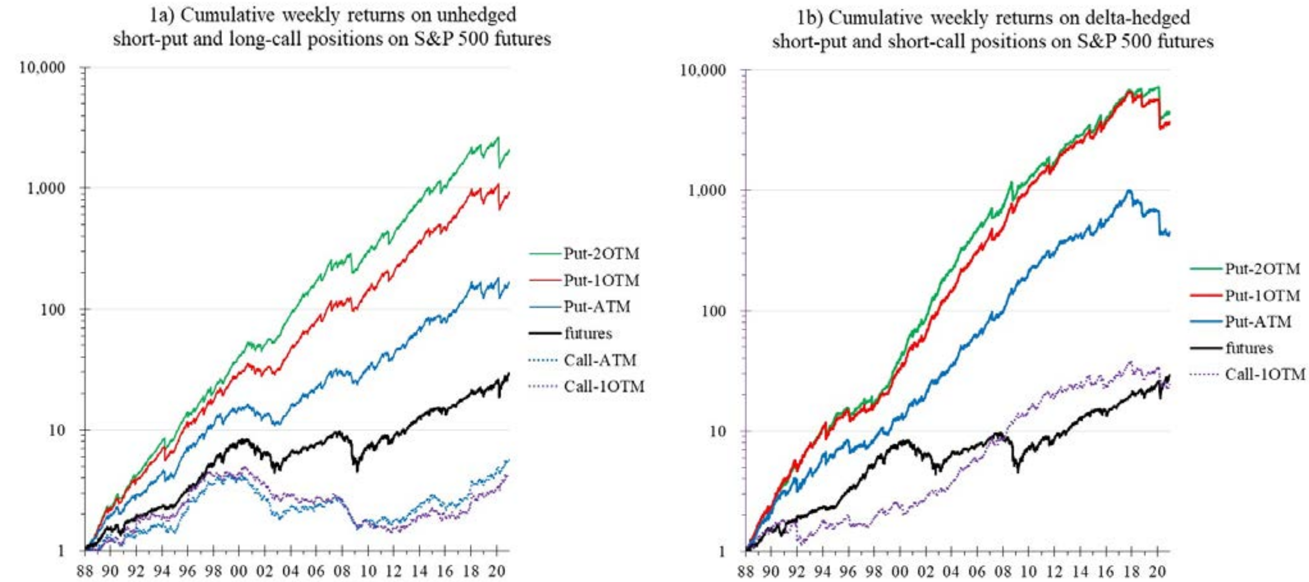


**Figure 1**. Cumulative weekly returns (Wednesday to Wednesday) of option strategies for American SP (1988-2008) and ES (2009-20) options on S&P 500 futures. All option positions are scaled to have the same 16.2% annualized volatility over 1988-2020 as S&P 500 futures (black line).

Maturities: monthly, plus or minus two weeks. At least two-week maturity on terminal Wednesday.

Moneyness: Options initially at-the-money or 1-2 standard deviations ($\sigma^{ATM}\sqrt{\tau}$) out-of-the-money.