

Understanding Index Option Returns

Mark Broadie, Mikhail Chernov, and Michael Johannes*

First Draft: September 2006

This Revision: May 3, 2007

Abstract

This paper studies the returns from investing in index options. Previous research documents significant average option returns, large CAPM alphas, and high Sharpe ratios, and concludes that put options are mispriced. We propose an alternative approach to evaluate the significance of option returns and obtain different conclusions. Instead of using these statistical metrics, we compare historical option returns to those generated by commonly used option pricing models. We find that the most puzzling finding in the existing literature, the large returns to writing out-of-the-money puts, is not even inconsistent with the Black-Scholes model. Moreover, simple stochastic volatility models with no risk premia generate put returns across all strikes that are not inconsistent with the observed data. At-the-money straddle returns are more challenging to understand, and we find that these returns are not inconsistent with explanations such as jump risk premia, Peso problems, and estimation risk.

*Broadie and Johannes are affiliated with the Graduate School of Business, Columbia University. Chernov is affiliated with London Business School and CEPR. We thank David Bates, Oleg Bondarenko, Peter Bossaerts, Pierre Collin-Dufresne, Kent Daniel, Bernard Dumas, Silverio Foresi, Vito Gala, Toby Moskowitz, Lasse Pedersen, Mirela Predescu, Todd Pulvino, Alex Reyfman, and participants of workshops at AQR Capital Management, Columbia, Goldman Sachs Asset Management, HEC-Lausanne, Lugano, Minnesota, Universidade Nova de Lisboa, Yale School of Management and the Adam Smith Asset Pricing (ASAP) conference for comments. Sam Cheung, Sudarshan Gururaj, and Pranay Jain have provided excellent research assistance. Broadie acknowledges partial support by NSF grant DMS-0410234. Chernov acknowledges support of the BNP Paribas Hedge Funds Centre.

1 Introduction

It appears to be a common perception that index options are mispriced, in the sense that certain option returns are excessive relative to their risks. In fact, some researchers go as far to refer to these returns as puzzling or anomalous.¹ In this paper, we provide a new perspective on the evidence and methods used to support these claims, and come to largely different conclusions.

The primary evidence supporting mispricing is the large magnitude of historical returns to writing S&P 500 put options. For example, Bondarenko (2003) reports that average at-the-money (ATM) put returns are -40% , not per annum, but per month, and deep out-of-the-money (OTM) put returns are -95% per month. Average option returns and CAPM alphas are statistically significant with p -values close to zero, and Sharpe ratios are larger than those of the underlying index. The returns are economically significant, as investors endowed with a wide array of utility functions find large certainty equivalent gains from selling put options (e.g., Driessen and Maenhout, 2004; Santa-Clara and Saretto, 2005).

Still unaddressed is the question of whether or not option returns remain puzzling in the context of commonly used option pricing models. In this paper, we evaluate significance by comparing observed option returns with those generated by affine jump-diffusion models that are widely accepted as plausible descriptions of S&P 500 returns.² These

¹A few quotations highlight the general sentiment of the literature: “The most likely explanation is mispricing of options... A simulated trading strategy exploiting such mispricing yields risk-adjusted expected excess returns during the post-crash period. These excess returns persist even when we account for transaction costs and hedge the downside risk” (Jackwerth (2000), p. 450); “No equilibrium model from a class of models can possibly explain the put anomaly, even when allowing for the possibility of incorrect beliefs and a biased sample. The class of rejected models is fairly broad.” (Bondarenko (2003), p. 3); “For index options, we find significantly positive abnormal returns when selling options across the range of exercise prices, with the lowest exercise prices (e.g., out-of-the-money puts) being most profitable” (Bollen and Whaley (2004), p. 714); “The analysis further shows that volatility risk and possibly jump risk are priced in the cross-section of index options, but that these systematic risks are insufficient for explaining average option returns. ...deep OTM money put options appear overpriced relative to longer-term OTM puts and calls, often generating negative abnormal returns in excess of half a percent per day” (Jones (2006), pp. 3-4), and the “empirical evidence on option returns suggest that stock index options markets are operating inefficiently” (Bates (2006), p. 2).

²In this regard, our approach follows the standard practice in asset pricing by evaluating various returns statistics using benchmark models. For example, it is common to simulate various consumption based models to assess the significance of the equity premium or the volatility of stock returns. As in this

models incorporate diffusive price shocks, price jumps and square-root diffusive stochastic volatility, which are the main drivers of index return volatility.³ Option pricing models formally account for the non-linear nature of option payoffs. Therefore, they provide a more appropriate benchmark than the standard empirical asset pricing tests that rely on average returns, CAPM alphas, or Sharpe ratios.

Methodologically, we rely on two basic tools: analytical expected option returns (EORs) formulae and simulations to assess statistical significance. Our first contribution is to compute analytical EORs. EORs are the ratio of \mathbb{P} -measure expected payoffs to \mathbb{Q} -measure discounted expected payoffs, both easily computed for affine models. EORs are useful for understanding the quantitative implications of different factors and parameterizations on option returns. They also provide a natural benchmark for anchoring null values in hypothesis tests. Although simple to derive and compute, analytical EORs have not been used in the extant option pricing literature, to our knowledge.

Statistical significance is assessed using the parametric bootstrap. Central limit approximations are problematic for option return statistics because of small samples sizes (on the order of 200 months) and the irregular nature of option return distributions. Option returns are extremely skewed, as out-of-the-money expirations generate returns of -100% . Specifically, we use historical index data to estimate parameters. Next, we simulate index sample paths and compute statistics associated with option returns along each path, thereby constructing finite sample distributions of these statistics. In addition to average returns, we also analyze standard risk-adjustments such as CAPM alphas and Sharpe ratios and also straddles. It is important to note at this stage that we do not use option prices to calibrate our models, as this would imply we are explaining option returns with option prices, making the exercise circular.

Our approach provides a number of advantages relative to existing approaches: (1) It evaluates option returns relative to reasonable option pricing benchmarks, appropriately

literature, we also calibrate our models to the underlying. In the case of asset pricing models, the underlying fundamentals are quantities like consumption and dividend growth. In the case of S&P 500 options, the underlying is the S&P 500 index. Dai and Singleton (2004) perform a similar analysis, analyzing the implications of dynamic term structure models for expectation hypothesis regressions.

³Andersen, Benzoni, and Lund (2002), Bates (2000), Broadie, Chernov, and Johannes (2007), Chernov, Gallant, Ghysels, and Tauchen (2003), Eraker (2004), Eraker, Johannes, and Polson (2003), and Pan (2002), among many others, find that these models provide an accurate fit to both index returns and options prices.

anchoring hypothesis tests; (2) It automatically accounts for the peculiar statistical features of option returns, as model based option returns embed leverage and have kinked payoffs; (3) It allows researchers to easily compute finite sample distributions; (4) It provides a formal framework for evaluating various explanations for the observed option returns that include risk premia, Peso problems, and estimation risk. Our goal is not to test the affine class of option models as this has already been done, but rather to use these accepted models as data generating processes to analyze the impact of various factors (e.g., stochastic volatility, jumps) and parametric assumptions on option returns.

We do not explicitly consider equilibrium models as our goal is not to provide an equilibrium explanation of both option price and equity market puzzles in terms of underlying dividend processes, as in Bates (1988), Naik and Lee (1990), and Liu, Pan, and Wang (2005). These models capture many aspects of option prices and equity returns, but these models have difficulties explaining important option-relevant features such as the high frequency stochastic volatility behavior of equity returns, price jumps, and the leverage effect, in addition to the usual problems that standard equilibrium models encounter (e.g., equity premium, excess volatility, etc.). Our goal is different and more modest. We seek to understand the links between index option returns and commonly assumed properties of underlying index returns such as jumps in prices and stochastic volatility. Simultaneously explaining the properties of underlying economic fundamentals, equity returns, and option prices is beyond the scope of this paper.⁴

We construct monthly S&P 500 futures option returns using a long sample from 1987 to 2005. The data and our new methodology generate a number of interesting new findings. First, we find that put returns, and especially OTM put returns, are not puzzling, at least in the context of the standard the Black and Scholes (1973) and the Heston (1993) stochastic volatility (SV) models. Monthly Black-Scholes EORs are large, on the order of -10% to -20% for ATM options and -20% to -40% for OTM options, for reasonable equity premia and volatility levels. Expected put returns are concave functions of volatility, indicating that fluctuating volatility generally makes EORs more negative. The Black-Scholes model

⁴One promising approach is the recent model of Benzoni, Collin-Dufresne, and Goldstein (2006) who introduce a continuous-time extension of Bansal and Yaron (2003). They generate realistic volatility smiles under the assumption that the highly persistent process driving aggregate consumption growth has large rare jumps in combination with Epstein and Zin (1991) recursive utility. They do not consider stochastic volatility, leverage effects, or the implications for pricing ATM money option in terms of realized versus implied volatility.

generates a p -value for 6% OTM put returns of 8%, indicating marginal significance at best, and is much larger than those previously reported.

The SV model without factor risk premia (the drift of the SV process under both measures is the same) generates even more striking findings. EORs are more negative in the SV model than in the Black-Scholes model due to the concavity mentioned above. Moreover, the impact of fluctuating volatility is quantitatively important as the p -value for 6% OTM average put returns is now 24%. This indicates roughly one in four sample paths from the SV model generates average returns that are more negative than those observed historically. Across all strikes and put return statistics, the lowest p -value for the SV model is just above 3%, certainly not overwhelming evidence of option mispricing, especially under the assumption that there are no priced risk factors.

Second, CAPM alphas are strongly biased, both in population and in finite samples. The Black-Scholes CAPM alpha for 6% OTM puts is -18% with a p -value of 13%. Although Black-Scholes is a “single-factor” model, linear risk-corrections have little impact as CAPM alphas are quite close to raw average put returns, both in population and simulations. In Heston’s SV model (again, without priced diffusive volatility risk) alphas range from -16% for ATM puts to -24% for OTM puts. This bias, along with the sampling uncertainty, generates the p -value for the alpha on 6% OTM put returns of 40%. While we are not the first to point out that alphas are biased for non-normal returns, we are the first to quantify the biases in the context of standard option pricing models. This is important because CAPM alphas are still widely used in both practice and the academic literature to risk-correct option returns.

The dramatic increase in p -values relative to the existing literature occurs because EORs should be negative (i.e., the appropriate null hypothesis is not zero) and there is substantial sampling variation due to the small samples. These results are particularly striking, as OTM put returns are most often used as evidence that options are mispriced. These results certainly do not imply that the Black-Scholes or Heston models are accurate or good option pricing models. Rather, the results indicate that put returns are too noisy to assert options are mispriced or anomalous even relative to simplest models.

Third, we find that Merton’s jump-diffusion model, somewhat surprisingly, generates less negative EORs than the Black-Scholes model if jump risk is not priced. This occurs because the presence of unpriced jump risk increases the left tail mass for both the objective and risk-neutral measures in a similar manner, increasing expected put returns toward zero.

This has a key implication: because Merton’s model without jump risk premia can generate very steep implied volatility smiles, this result dispels the common perception that steep implied volatility smiles, per se, are associated with option mispricing and large option returns.

Based on the evidence from these simple models without priced jump or stochastic volatility risk, we conclude that standard factors go a long way in explaining the magnitude and statistical significance of put returns. Put returns, especially for deep OTM strikes, are not particularly puzzling, or at least are much less puzzling than indicated by the previous literature.

The only statistic that remains challenging to understand after the introduction of un-priced stochastic volatility and jumps in prices is ATM straddle returns. ATM straddle returns are generated by the well-known wedge between ATM implied volatility and subsequently realized volatility. Over our sample, ATM implied volatility averaged 17% and realized volatility was 15%. This wedge between \mathbb{Q} (implied volatility) and \mathbb{P} (realized volatility) is not likely to be explained solely by a diffusive stochastic volatility risk premium, but that a wedge between \mathbb{Q} and \mathbb{P} jump parameters is a more plausible explanation. We analyze three commonly cited mechanisms that generate this wedge: jump risk premia, estimation risk, and Peso problems. Again, in analyzing these explanations, it is important to note we do not use option prices to estimate \mathbb{Q} -parameters, but rather calibrate the parameter values using plausible assumptions.

Each of these explanations generates significantly more negative put and straddle returns. For example, the realized historical average straddle returns observed are -15.7% per month, and these explanations generate expected straddle returns just slightly less negative, about -10% to -14% per month and p -values indicate they are not statistically significant. The same conclusion holds for CAPM alphas and Sharpe ratios for straddle returns. Thus, we conclude that option returns do not appear to be particularly puzzling, at least relative to standard models and plausible parametric assumptions.

The rest of the paper is outlined as follows. Section 2 discusses our data set and summarizes the extant evidence for option mispricing. Section 3 outlines our methodological approach, and Sections 4 and 5 report results for benchmark models (without factor risk premia) and for the three explanations of the wedge between \mathbb{P} and \mathbb{Q} measures, respectively. Section 6 concludes.

2 The evidence for mispricing

In this section, we compute index option returns for a long historical sample, review the evidence for mispricing of put options, and provide a review of the existing literature. Since we use a different methodology than existing papers, we provide a detailed description of existing approaches prior to introducing our new approach.

2.1 Data

We consider one month returns for options held to expiration for various strikes. Put returns are defined as

$$r_{t,T}^p = \frac{(K - S_{t+T})^+}{P_{t,T}(K, S_t)} - 1, \quad (1)$$

where $x^+ \equiv \max(x, 0)$, $P_{t,T}(K, S_t)$ is the observed price of a put option written on an asset S_t , at time t , struck at K , and expiring at time $t + T$.

Hold-to-expiration returns are typically analyzed in both academic studies (with a few exceptions) and in practice. Option trading involves significant costs and strategies that hold until expiration incur these costs only once. For example, ATM index option bid-ask spreads are currently on the order to 3% to 5% of the option price, and the bid-ask spreads are larger, often more than 10%, for deep OTM strikes. Following the literature and for other reasons discussed in more detail below, we also consider returns generated by model-independent trading strategies such as covered returns and straddles.

Our data consists of S&P 500 futures options from August 1987 to June 2005, a total of 215 months. This sample is considerably longer than those previously analyzed and starts in August of 1987 when one-month “serial” options were introduced. Options mature on the third Friday of each month, which implies there are 28 or 35 calendar days to maturity depending on whether it was a four- or five-week month. We are careful to account for holidays. We construct representative daily option prices using the approach in Broadie, Chernov, and Johannes (2007); details of this procedure are in Appendix A.

Using these prices, we compute option returns for fixed moneyness, measured by strike divided by the underlying, ranging from 0.94 to 1.02 (in 2% increments). This range represents the most actively traded options: 85% of one-month option transactions occur in this range. We did not include deeper OTM or ITM strikes because of missing values. We computed payoffs using the settlement values for the S&P 500 futures contract.

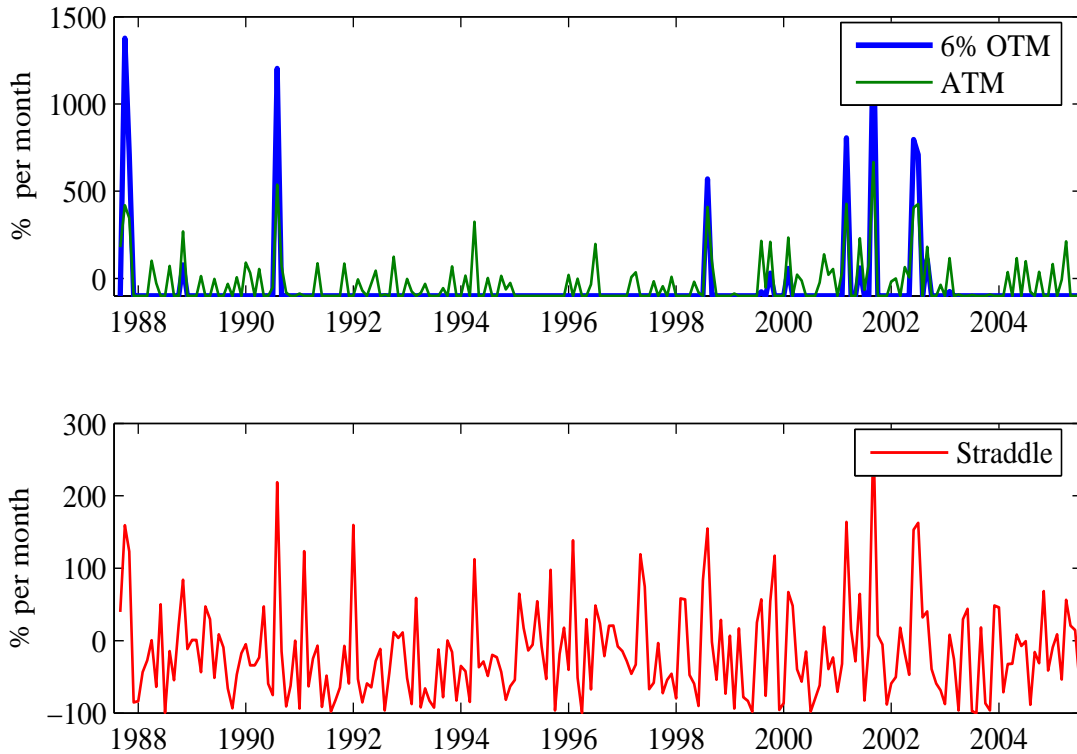


Figure 1: Time series of options returns. The top panel shows the time series of put returns with moneyness of 0.94 and 1. The bottom panel shows the time series of at-the-money straddle returns.

Figure 1 shows the time series for 6% OTM and ATM puts and ATM straddles. This figure highlights some of the potential issues that are present when evaluating the statistical significance of statistics generated by option returns. The put return time series have very large outliers and many repeated values, since OTM expirations generate returns of -100% .

2.2 Option returns summary statistics

Table 1 summarizes the distributional features of put returns. We report average returns, standard errors, t -statistics, p -values, and measures of non-normality (skewness and kurtosis).

Moneyness	0.94	0.96	0.98	1.00	1.02	Strdl
8/1987 to 6/2005	-56.8	-52.3	-44.7	-29.9	-19.0	-15.7
Standard error	14.2	12.3	10.6	8.8	7.1	4.5
<i>t</i> -stat	-3.9	-4.2	-4.2	-3.3	-2.6	-3.4
<i>p</i> -value, %	0.0	0.0	0.0	0.0	0.4	0.0
Skew	5.5	4.5	3.6	2.5	1.8	1.2
Kurt	34.2	25.1	16.7	10.5	7.1	5.2
Subsamples						
01/1988 to 06/2005	-65.2	-60.6	-51.5	-34.1	-21.6	-16.8
01/1995 to 09/2000	-85.5	-71.6	-63.5	-50.5	-37.5	-10.8
10/2000 to 02/2003	+67.2	+54.3	+44.5	+48.2	+40.4	+4.0
08/1987 to 01/2000	-83.9	-63.2	-55.7	-39.5	-25.5	-19.1

Table 1: Average put option returns. The first panel contains the full sample, with standard errors, *t*-statistics, and skewness and kurtosis statistics. The second panel analyzes subsamples. All relevant statistics are in percentages per month. The column “Strdl” refers to the statistics associated with at-the-money straddles.

sis). We also report average returns over various subsamples.

The first evidence commonly cited supporting mispricing is the large magnitude of the returns. Average monthly returns are about -60% for 6% OTM strikes and -30% for ATM strikes. These returns are highly statistically different from zero based on standard *t*-tests, as *p*-values are close to zero. The bottom panel reports average returns over subsamples. In particular, to check that our results are consistent with previous findings, we compare our statistics to the ones in the Bondarenko (2003) sample from 1987 to 2000. The returns are very close, but ours are slightly more negative for every moneyness category except the deepest OTM category. Bondarenko (2003) uses closing prices and has some missing values, which generate much of the differences. Our returns are more negative than those reported for similar time periods by Santa-Clara and Saretto (2005) using different option contracts.

Average put returns are unstable over time. For example, put returns were extremely negative during the late 1990s during the dot-com “bubble,” but were positive and large from late 2000 to early 2003. The subsample starting in January 1988 provides the same

insight: if the extremely large positive returns realized around the crash of 1987 are excluded, returns are much lower. Doing so, however, generates a case of sample selection bias, and clearly demonstrates a problem with tests using short sample periods.⁵ Note that our full sample results are significantly less negative than those in Bondarenko (2003).

It should not be surprising that average put returns are quite negative, since puts are levered short positions in the underlying. Thus it is crucial to de-lever or risk-correct option returns to account for the underlying exposure. The most common approaches for doing this are to (1) compute Sharpe ratios, (2) compute factor model alphas, (3) compute covered option positions (buying an option and the underlying index), and (4) compute straddle returns.⁶ Appendix B discusses delta-hedged returns and issues surrounding them.

The final column of Table 1 summarizes straddle returns, and Table 2 summarizes the Sharpe ratios, CAPM alphas, and covered positions.⁷ CAPM alphas and ATM straddle returns are highly statistically significant, with p -values near zero. Interestingly, the covered put positions are insignificantly different from zero for all strikes and economically small. From now on, we do not consider covered positions. The Sharpe ratios of put positions are larger than those on the underlying market. For example, the monthly Sharpe ratio for the market over our time period was about 0.1, and the put return Sharpe ratios are two to three times larger. Straddles deliver Sharpe ratios of this general magnitude also.

Based largely on this evidence and additional robustness checks (which we discuss in the following subsection), the literature concludes that put returns are puzzling and options are likely mispriced.

⁵In simulations of the Black-Scholes model, excluding the largest positive return reduces *average* option returns by about 15% for the 6% OTM strike. This outcome illustrates the potential sample selection issues and how sensitive option returns are to the rare but extremely large positive returns generated by events such as the crash of 1987.

⁶We have also computed returns to crash-neutral put positions, such as buying an ATM put option and selling an OTM put option. These portfolios do not provide any additional insights beyond standard put returns.

⁷We compute Sharpe ratios and CAPM alphas using monthly options returns.

Moneyness	0.94	0.96	0.98	1.00	1.02
CAPM α , %	-48.3	-44.1	-36.8	-22.5	-12.5
Std.err., %	11.6	9.3	7.1	4.8	2.9
t -stat	-4.1	-4.7	-5.1	-4.6	-4.2
p -value, %	0.0	0.0	0.0	0.0	0.00
Sharpe ratio	-27.3	-29.0	-29.0	-23.4	-18.5
Covered puts, %	0.20	0.11	-0.01	-0.06	-0.08
Std.err., %	0.25	0.23	0.20	0.16	0.12
Skew	0.06	0.35	0.69	1.21	2.02
Kurt	2.88	2.76	3.05	4.38	8.47

Table 2: Risk-corrected measures of average put option returns. The first panel provides put option Sharpe ratios, the second panel provides CAPM α 's with standard errors, and the third panel contains covered put returns. All relevant statistics are in percentages per month. The p -values are computed under assumption that t -statistics are t -distributed.

2.3 Previous research on option returns

Before discussing our approach and results, we provide a brief review of the existing literature analyzing index option returns.⁸ The market for index options developed in the mid to late 1980s. The Black-Scholes implied volatility smile indicates that OTM put options are expensive relative to the ATM puts, and the issue is to then determine if these put options are in fact mispriced.

Jackwerth (2000) documents that the risk-neutral distribution computed from S&P 500 index puts exhibit a pronounced negative skew after the crash of 1987. Based on a single factor model, he shows that utility over wealth has convex portions, interpreted as evidence of option mispricing. Investigating this further, Jackwerth (2000) analyzes monthly put trading strategies from 1988 to 1995 and finds that put writing strategies deliver high

⁸Prior to the development of markets on index options, a number of articles analyzed option returns on individual securities. These articles, including Merton, Scholes, and Gladstein (1978) and (1982), Gastineau and Madansky (1979), and Bookstaber and Clarke (1985). The focus is largely on returns to various historical trading strategies assuming the Black-Scholes model is correct. Sheikh and Ronn (1994) document market microstructure patterns of option returns on individual securities.

returns, both in absolute and risk-adjusted levels, with the most likely explanation being option mispricing.

In a related study, Aït-Sahalia, Wang, and Yared (2001) report a discrepancy between a risk-neutral density of S&P 500 index returns implied by the cross-section of options versus the time series of the underlying returns. The authors exploit the discrepancy to set up “skewness” and “kurtosis” trade portfolios. Depending on the current relative values of the two implicit densities, the portfolios were long or short a mix of ATM and OTM options. The portfolios were rebalanced every three months. Aït-Sahalia, Wang, and Yared (2001) find that during the 1986 to 1996 period such strategies would have yielded Sharpe ratios that are two to three times larger than those of the market.

Coval and Shumway (2001) analyze weekly option and straddle returns from 1986 to 1995. They find that put returns are too negative to be consistent with a single-factor model, and that beta-neutral straddles still have significantly negative returns. Importantly, they do not conclude that options are mispriced, but rather that the evidence points toward additional priced risk factors.

Bondarenko (2003) computes monthly returns for S&P 500 index futures options from August 1987 to December 2000. Using a novel test based on equilibrium models, Bondarenko finds significantly negative put returns that are inconsistent with single-factor equilibrium models. His test results are robust to risk adjustments, Peso problems, and the underlying equity premium. He concludes that puts are mispriced and that there is a “put pricing anomaly.” Bollen and Whaley (2003) analyze monthly S&P 500 option returns from June 1988 to December 2000 and reach a similar conclusion. Using a unique dataset, they find that OTM put returns were abnormally large over this period, even if delta-hedged. Moreover, the pricing of index options is different than individual stock options, which were not overpriced. The results are robust to transaction costs.

Santa-Clara and Saretto (2005) analyze returns on a wide variety of S&P 500 index option portfolios, including covered positions and straddles, in addition to naked option positions. They argue that the returns are implausibly large and statistically significant by any metric. Further, these returns may be difficult for small investors to achieve due to margin requirements and potential margin calls.

Most recently, Jones (2006) analyzes put returns, departing from the literature by considering daily option (as opposed to monthly) returns and a nonlinear multi-factor model. Using data from 1987 to September 2000, Jones finds that deep OTM put options

have statistically significant alphas, relative to his factor model. Both in and out-of-sample, simple put-selling strategies deliver attractive Sharpe ratios. He finds that the linear models perform as well or better than nonlinear models. Bates (2006) reviews the evidence on stock index option pricing, and concludes that options do not price risks in a manner consistent with current option-pricing models.

Given the large returns to writing put options, Driessen and Maenhout (2004a) assess the economic implications for optimal portfolio allocation. Using closing prices on the S&P 500 futures index from 1987 to 2001, they estimate expected utility using realized returns. For a wide range of expected and non-expected utility functions, investors optimally short put options, in conjunction with long equity positions. Since this result holds for various utility functions and risk aversion parameters, their finding introduces a serious challenge to explanations of the put-pricing puzzle based on heterogeneous expectations, as a wide range of investors find it optimal to sell puts.

Driessen and Maenhout (2004b) analyze the pricing of jump and volatility risk across multiple countries. Using a linear factor model, they regress ATM straddle and OTM put returns on a number of index and index option based factors. They find that individual national markets have priced jump and volatility risk, but find little evidence of an international jump or volatility factor that is priced across countries.

3 Our methodology

Existing approaches for evaluating the significance of option returns rely on utility-based tests or purely statistical methods, as reviewed in the previous section. Our approach provides an alternative testing approach. We compare market observed returns (and associated statistics) with those generated by standard option pricing models such as Black-Scholes and extensions incorporating jumps or stochastic volatility. This section describes our method in detail.

3.1 Models

We consider models that are nested versions of a general model with square-root stochastic volatility and log-normally distributed Poisson driven jumps in prices. This model, proposed by Bates (1996) and Scott (1997), which we refer to as the SVJ model, is a common

benchmark model for index option prices (see, e.g., Andersen, Benzoni, and Lund (2002), Bates (1996), Broadie, Chernov, and Johannes (2007), Chernov, Gallant, Ghysels, and Tauchen (2003), Eraker (2004), Eraker, Johannes, and Polson (2003), and Pan (2002)). As special cases of the model, we consider the Black and Scholes (1973) model, Merton's (1976) jump-diffusion model with constant volatility, and Heston's (1993) square-root stochastic volatility model.

The model assumes that the ex-dividend index level, S_t , and its spot variance, V_t , evolve according to

$$dS_t = (r + \mu - \delta) S_t dt + S_t \sqrt{V_t} dW_t^s(\mathbb{P}) + d \left(\sum_{j=1}^{N_t(\mathbb{P})} S_{\tau_{j-}} \left[e^{Z_j^s(\mathbb{P})} - 1 \right] \right) - \lambda^{\mathbb{P}} \bar{\mu}^{\mathbb{P}} S_t dt \quad (2)$$

$$dV_t = \kappa_v^{\mathbb{P}} (\theta_v^{\mathbb{P}} - V_t) dt + \sigma_v \sqrt{V_t} dW_t^v(\mathbb{P}), \quad (3)$$

where r is the risk-free rate, μ is the cum-dividend equity premium, δ is the dividend yield, W_t^s and W_t^v are two correlated Brownian motions ($E[W_t^s W_t^v] = \rho t$), $N_t(\mathbb{P}) \sim \text{Poisson}(\lambda^{\mathbb{P}} t)$, $Z_j^s(\mathbb{P}) \sim \mathcal{N}(\mu_z^{\mathbb{P}}, (\sigma_z^{\mathbb{P}})^2)$, and $\bar{\mu}^{\mathbb{P}} = \exp(\mu_z^{\mathbb{P}} + (\sigma_z^{\mathbb{P}})^2 / 2) - 1$. Black-Scholes is a special case with no jumps ($\lambda^{\mathbb{P}} = 0$) and constant volatility ($V_0 = \theta_v^{\mathbb{P}}$, $\sigma_v = 0$), Heston's model is the special case with no jumps, and Merton's model is the special case with constant volatility. When volatility is constant, we use the notation $\sqrt{V_t} = \sigma$.

Under the risk-neutral measure \mathbb{Q} , the dynamics are given by

$$dS_t = (r - \delta) S_t dt + S_t \sqrt{V_t} dW_t^s(\mathbb{Q}) + d \left(\sum_{j=1}^{N_t(\mathbb{Q})} S_{\tau_{j-}} \left[e^{Z_j^s(\mathbb{Q})} - 1 \right] \right) - \lambda^{\mathbb{Q}} \bar{\mu}^{\mathbb{Q}} S_t dt \quad (4)$$

$$dV_t = \kappa_v^{\mathbb{Q}} (\theta_v^{\mathbb{Q}} - V_t) dt + \sigma_v \sqrt{V_t} dW_t^v(\mathbb{Q}), \quad (5)$$

where $N_t(\mathbb{Q}) \sim \text{Poisson}(\lambda^{\mathbb{Q}} t)$, $Z_j(\mathbb{Q}) \sim \mathcal{N}(\mu_z^{\mathbb{Q}}, (\sigma_z^{\mathbb{Q}})^2)$, and $W_t(\mathbb{Q})$ are Brownian motions, and $\bar{\mu}^{\mathbb{Q}}$ is defined analogously to $\bar{\mu}^{\mathbb{P}}$. The diffusive equity premium is μ^c , and the total equity premium is $\mu = \mu^c + \lambda^{\mathbb{P}} \bar{\mu}^{\mathbb{P}} - \lambda^{\mathbb{Q}} \bar{\mu}^{\mathbb{Q}}$.

The parameters $\theta_v^{\mathbb{P}}$ and $\kappa_v^{\mathbb{P}}$ can both potentially change under the risk-neutral measure (Cheredito, Filipovic, and Kimmel (2003)). We explore changes in $\theta_v^{\mathbb{P}}$ and constrain $\kappa_v^{\mathbb{Q}} = \kappa_v^{\mathbb{P}}$, because, as discussed below, average returns are not sensitive to empirically plausible changes in $\kappa_v^{\mathbb{P}}$. Changes of measure for jump processes are more flexible than those for diffusion processes. We take the simplifying assumptions that the jump size distribution is log-normal with potentially different means and variances. Below we explore three explanations for differences between \mathbb{Q} and \mathbb{P} which all involve various parameterizations of the \mathbb{Q} -measure.

3.2 Expected instantaneous option returns

Before analyzing EORs using analytical and simulation methods, we first develop some intuition about signs, magnitudes, and determinants of instantaneous EORs. Appendix C applies arguments similar to those used by Black and Scholes to derive their option pricing model for the more general SVJ model. We discuss the single-factor Black-Scholes model first, and then extensions incorporating stochastic volatility and jumps.

3.2.1 The Black-Scholes model

In Black-Scholes, the link between instantaneous returns on a derivative, $f(S_t)$, and excess index returns is

$$\frac{df(S_t)}{f(S_t)} = rdt + \frac{S_t}{f(S_t)} \frac{\partial f(S_t)}{\partial S_t} \left[\frac{dS_t}{S_t} - (r - \delta) dt \right].$$

This expression displays two crucial features of the Black-Scholes model. First, instantaneous changes in the derivative's price are *linear* in the index returns, dS_t/S_t . Second, *instantaneous* option returns are conditionally normally distributed. This linearity and normality motivated Black and Scholes to assert a “local” CAPM-style model:

$$\frac{1}{dt} E_t^{\mathbb{P}} \left[\frac{df(S_t)}{f(S_t)} - rdt \right] = \frac{\partial \log[f(S_t)]}{\partial \log(S_t)} \mu.$$

In the Black-Scholes model, this expression shows that EORs are determined by the equity premium and the option's elasticity, which, in turn, are functions primarily of moneyness and volatility.

This instantaneous CAPM is often used to motivate an approximate CAPM model for finite holding period returns,

$$E_t^{\mathbb{P}} \left[\frac{f(S_{t+T}) - f(S_t)}{f(S_t)} - rT \right] \approx \beta_t \mu T,$$

and the model is often tested via an *approximate* linear factor model for option returns

$$\frac{f(S_{t+T}) - f(S_t)}{f(S_t)} = \alpha_T + \beta_t \left(\frac{S_{t+T} - S_t}{S_t} - rT \right) + \varepsilon_{t,T}.$$

As reviewed above, a number of authors use this as a statistical model of returns, and point to findings that $\alpha_T \neq 0$ as evidence of either mispricing or risk premia.

This argument, however, has a serious potential problem as the CAPM does not hold over finite time horizons. Option prices are convex functions of the underlying price,

and therefore linear regressions of option returns and underlying returns are generically misspecified. This implies, for example, that α could depend on $(S_t, K, t, T, \sigma, \mu)$ and is not zero in population. Since the results hold in continuous-time, the degree of bias depends on the length of the holding period. Since option returns are highly skewed, the errors $\varepsilon_{t,T}$ are also highly skewed, bringing into question the applicability of ordinary least squares. We show below that even the simple Black-Scholes model generates economically large alphas for put options. These results also bring into question the practice of computing alphas for multi-factor specifications such as the Fama-French model.

3.2.2 Stochastic volatility and jumps

Consider next the case of Heston's square-root stochastic volatility model. As derived in Appendix C, instantaneous realized option returns are driven by both factors,

$$\frac{df(S_t, V_t)}{f(S_t, V_t)} = rdt + \beta_t^s \left[\frac{dS_t}{S_t} - (r - \delta) dt \right] + \beta_t^v [dV_t - \kappa_v^{\mathbb{P}}(\theta_v^{\mathbb{Q}} - V_t)], \quad (6)$$

and expected excess returns are given by

$$\frac{1}{dt} E_t^{\mathbb{P}} \left[\frac{df(S_t, V_t)}{f(S_t, V_t)} - rdt \right] = \beta_t^s \mu + \beta_t^v \kappa_v^{\mathbb{P}} (\theta_v^{\mathbb{P}} - \theta_v^{\mathbb{Q}}), \quad (7)$$

where

$$\beta_t^s = \frac{\partial \log [f(S_t, V_t)]}{\partial \log S_t} \text{ and } \beta_t^v = \frac{\partial \log [f(S_t, V_t)]}{\partial V_t}.$$

Since β_t^v is positive for all options and priced volatility risk implies that $\theta_v^{\mathbb{P}} < \theta_v^{\mathbb{Q}}$, expected put returns are more negative with priced volatility risk.

Equations (6) and (7) highlight the shortcomings of standard CAPM regressions, even in continuous-time. Regressions of excess option returns on excess index returns will potentially generate negative alphas for two reasons. First, if the volatility innovations are omitted then α will be negative to capture the effect of the volatility risk premium. Second, because dS_t/S_t is highly correlated with dV_t , CAPM regressions generate biased estimates of β and α due to omitted variable bias. As in the Black-Scholes case, discretizations will generate biased coefficient estimates.

Next, consider the impact of jumps in prices via Merton's model. Here, the link between

option and index returns is far more complicated:

$$\begin{aligned} \frac{df(S_t)}{f(S_t)} &= rdt + \frac{\partial \log(f(S_t))}{\partial \log(S_t)} \left[\frac{dS_t^c}{S_t} - (r - \delta - \lambda^{\mathbb{Q}} \bar{\mu}^{\mathbb{Q}}) dt \right] \\ &+ \left[\frac{f(S_{t-e^Z}) - f(S_{t-})}{f(S_{t-})} \right] - \lambda^{\mathbb{Q}} \frac{E_t^{\mathbb{Q}} [f(S_t e^Z) - f(S_t)]}{f(S_t)} dt, \end{aligned}$$

where dS_t^c denote the continuous portion of the sample path increment and $S_t = S_{t-} e^Z$. The first line is similar to the expressions given earlier, with the caveat that excess index returns contain only the continuous portion of the increment. The second line captures the effect of discrete jumps. Expected returns are given by

$$\frac{1}{dt} E_t^{\mathbb{P}} \left[\frac{df(S_t)}{f(S_t)} - rdt \right] = \beta_t \mu^c + \frac{\lambda^{\mathbb{P}} E_t^{\mathbb{P}} [f(S_t e^Z) - f(S_t)] - \lambda^{\mathbb{Q}} E_t^{\mathbb{Q}} [f(S_t e^Z) - f(S_t)]}{f(S_t)}.$$

Because option prices are convex functions of the underlying, $f(S_{t-e^Z}) - f(S_{t-})$ cannot be linear in the jump size, e^Z , and thus even instantaneous option returns are not linear in index returns. This shows why linear factor models are fundamentally not applicable in models with jumps in prices. For contracts such as put options and standard forms of premia (e.g., $\mu_z^{\mathbb{Q}} < \mu_z^{\mathbb{P}}$), $E_t^{\mathbb{P}} [f(S_t e^Z)] < E_t^{\mathbb{Q}} [f(S_t e^Z)]$, which implies that expected put option returns are negatively impacted by any jump size risk premia. As in the case of stochastic volatility, a single-factor CAPM regression, even in continuous-time, is inappropriate. Moreover, negative alphas are fully consistent with jump risk premia and are not indicative of mispricing.

3.3 Characterizing option returns

In this section, we show how to compute exact EORs, and how we use simulations to compute the finite sample distribution of option return statistics.

3.3.1 Analytical expected option returns

In contrast to the instantaneous expected returns in the previous section, we compute exact expected option returns. Expected put option returns are given by

$$\begin{aligned} E_t^{\mathbb{P}}(r_{t,T}^p) &= E_t^{\mathbb{P}} \left[\frac{(K - S_{t+T})^+}{P_{t,T}(S_t, K)} \right] - 1 = \frac{E_t^{\mathbb{P}} [(K - S_{t+T})^+]}{P_{t,T}(S_t, K)} - 1 \\ &= \frac{E_t^{\mathbb{P}} [(K - S_{t+T})^+]}{E_t^{\mathbb{Q}} [e^{-rT} (K - S_{t+T})^+]} - 1, \end{aligned} \quad (8)$$

where in the second equality $P_{t,T}$ is known at time t . Now, the put prices, $P_{t,T}(S_t, K)$, will depend on the specific model under consideration. Our key insight is that for any model that admits “analytical” option prices, such as affine models, EORs can be explicitly computed since both the numerator and denominator are known analytically.⁹ Surprisingly, despite a large literature analyzing option returns, the fact that EORs are known has neither been noted nor applied.¹⁰ EORs do not depend on S_t . To see this, define the initial moneyness of the option as $\kappa = K/S_t$. Option homogeneity implies that

$$E_t^{\mathbb{P}}(r_{t,T}^p) = \frac{E_t^{\mathbb{P}} [(\kappa - R_{t,T})^+]}{E_t^{\mathbb{Q}} [e^{-rT} (\kappa - R_{t,T})^+]} - 1, \quad (9)$$

where $R_{t,T} = S_{t+T}/S_t$ is the gross return on the index. It is now clear that expected option’s return depends only on the moneyness, maturity, interest rate, and the distribution of index returns.¹¹

This formula provides exact EORs for finite holding periods and regardless of the risk factors of the underlying index dynamics, without using CAPM-style approximations such as those discussed in the previous section. These analytical results are primarily useful as

⁹Similarly, we can compute $E_t^{\mathbb{P}} \left[\left(r_{t,T}^p \right)^k \right]$ for $k = 2, 3, 4, \dots$, that is, we can compute other moments analytically or semi-analytically.

¹⁰This result is closely related to Rubinstein (1984), who derived it specifically for the Black-Scholes case and analyzed the relationship between hold-to-expiration and shorter holding period expected returns.

¹¹When stochastic volatility is present in a model, the expected option returns are analytical conditional on the current variance value: $E^{\mathbb{P}}(r_{t,T}^p | V_t)$. The unconditional expected returns can be computed using iterated expectations and the fact that

$$E^{\mathbb{P}}(r_{t,T}^p) = \int E^{\mathbb{P}}(r_{t,T}^p | V_t) p(V_t) dV_t.$$

The integral can be estimated via Monte Carlo simulation or by standard deterministic integration routines.

they allow us to assess the exact *quantitative* impact of risk premia or parameter configurations. Equation (8) implies that the gap between \mathbb{P} and \mathbb{Q} probability measures determines expected option returns, and the magnitude of the returns is determined by the relative shape and location of the two probability measures.¹² In models without jump or stochastic volatility risk premia, the gap is determined by the fact that the \mathbb{P} and \mathbb{Q} drifts are different by the factor μ . In models with priced stochastic volatility or jump risk, both the shape and location of the distribution can change, leading to more interesting patterns of expected returns across different moneyness categories.

3.3.2 Finite sample distribution via simulation

To assess statistical significance, we use Monte Carlo simulation to compute the distribution of various returns statistics, including average returns, CAPM alphas, and Sharpe ratios. We are motivated by concerns that the use of limiting distributions to approximate the finite sample distribution is inaccurate in this setting. The accuracy of central limit theorem approximations depends on the nature of the underlying random variables. In this setting, our concerns arise due to the relatively short sample (215 months), and due to the extreme non-normality of option returns.

To compute finite sample distribution of various option return statistics, we simulate $N = 215$ months (the sample length in the data) of index levels $G = 25,000$ times using standard simulation techniques. For each month and path pair, we compute returns for put options with a fixed moneyness via

$$r_{t,T}^{p,(g)} = \frac{\left(\kappa - R_{t,T}^{(g)}\right)^+}{P_T(\kappa)} - 1, \quad (10)$$

where

$$P_T(\kappa) \triangleq \frac{P_{t,T}(S_t, K)}{S_t} = e^{-rT} E_t^{\mathbb{Q}} \left[(\kappa - R_{t,T})^+ \right],$$

$t = 1, \dots, N$ and $g = 1, \dots, G$. Average option returns are given by

$$\bar{r}_T^{p,(g)} = \frac{1}{N} \sum_{t=1}^N r_{t,T}^{p,(g)}.$$

¹²For monthly holding periods, $1 \leq \exp(rT) \leq 1.008$ for $0\% \leq r \leq 10\%$ and $T = 1/12$ years, so this term has a negligible impact on EORs.

A set of G average returns forms the finite sample distribution. Similarly, we can construct finite sample distributions for the Sharpe ratios, CAPM alphas, straddles, and other statistics of interest.

This approach, commonly called the parametric bootstrap, provides exact finite sample inference under the null hypothesis that a given model holds. It can be contrasted with the nonparametric bootstrap, which creates artificial datasets by sampling with replacement from the observed data. The nonparametric bootstrap, which just reshuffles existing observations, has difficulties dealing with rare events. In fact, if an event has not occurred in the observed sample, it will never appear in the simulated finite sample distribution. This is an important concern when dealing with put returns which are very sensitive to rare events.

3.4 Parameter estimation

Objective, or \mathbb{P} -measure, parameter estimates are required to simulate option returns. We calibrate our models to fit the realized historical behavior of the underlying index returns over our observed sample. For parameters in the Black-Scholes model, this calibration is straightforward, but in models with unobserved volatility or jumps, the estimation is more complicated as it is not possible to estimate all of the parameter via simple sample statistics.

We calibrate the interest rate and equity premium to match those observed over our sample, $r = 4.5\%$ and $\mu = 5.4\%$. We simulate futures returns and futures options, thus $\delta = r$. We also constrain total volatility in each model to match the observed volatility of 15%. In the most general model we consider, we do this by imposing that

$$\sqrt{\theta_v^{\mathbb{P}} + \lambda^{\mathbb{P}}((\mu_z^{\mathbb{P}})^2 + (\sigma_z^{\mathbb{P}})^2)} = 15\%$$

by modifying $\theta_v^{\mathbb{P}}$. In the Black-Scholes model, we set the constant volatility to be 15%.

To obtain the values of the remaining parameters, we estimate the SVJ model using daily S&P 500 index returns spanning the same time period as our options data, August 1987 to June 2005. We use MCMC methods to simulate the posterior distribution of the parameters and state variables following Eraker, Johannes, and Polson (2003) and others. The parameter estimates (posterior means) and posterior standard deviations are reported in Table 3. The parameter estimates are in line with the values reported in previous studies (see Broadie, Chernov, and Johannes, 2007 for a review).

r	μ	$\lambda^{\mathbb{P}}$	$\mu_z^{\mathbb{P}}$	$\sigma_z^{\mathbb{P}}$	$\sqrt{\theta_v^{\mathbb{P}}}$ (SV)	$\sqrt{\theta_v^{\mathbb{P}}}$ (SVJ)	κ_v	σ_v	ρ
4.50%	5.41%	0.91	-3.25%	6.00%	15.00%	13.51%	5.33	0.14	-0.52
		(0.34)	(1.71)	(0.99)		(1.28)	(0.84)	(0.01)	(0.04)

Table 3: \mathbb{P} -parameters. We report parameter values that we use in our computational examples. Standard errors from the SVJ estimation are reported in parentheses. Parameters are given in annual terms.

Our \mathbb{P} -measure parameter estimates provide a model-based summary of what actually occurred, and this is potentially different from risk-neutral investor's expectations (the \mathbb{Q} -measure). These \mathbb{P} -measure parameters provide a summary of the historical behavior of stock returns in terms of the estimated jump intensities, jump distribution parameters, and volatility parameters. It is important that we estimate these parameters over the same sample period over which we have option returns. This allows us to generate samples for constructing finite sample distributions that mimic the properties of the observed sample.

Of particular importance are the jump parameter estimates. The estimates imply that jumps are relatively infrequent, arriving at a rate of about 0.91 per year. The jumps are modestly sized with the mean of -3.25% and a standard deviation of 6% . These parameters values imply a jump the size of the crash of 1987 event would occur every 1650 years. This assumes, counterfactually, that the entire move is attributed to the jump component with diffusive shocks not contributing. If we assume that a jump occurs simultaneously with a three-standard deviation diffusive move, $3\sqrt{\theta_v^{\mathbb{P}}}$, a crash of 1987 event occurs every 407 years.¹³

As we discuss in greater detail below, estimating jump intensities and jump size distributions is extremely difficult. The estimates are highly dependent on the observed data and on the specific model. For example, different estimates would likely be obtained if we assumed that the jump intensity was dependent on volatility (as in Bates (2000) or Pan (2002)) or if there were jumps in volatility. Again, our goal is not to exhaustively analyze every potential specification, but rather to understand option returns in common specifications and for plausible parameter values.

¹³According to our estimates, the volatility on the day of the crash of 1987 was 25% . If we assume that volatility will always be so high during the crashes, then we would expect them to occur every 141 years.

We discuss the calibration of \mathbb{Q} -measure parameters later. At this stage, we only emphasize that we do not use options data to estimate any of the parameters. Estimating \mathbb{Q} -parameters from option prices for use in understanding observed option returns would introduce a circularity, as we would be explaining option returns with option prices.

4 Option returns in the absence of risk premia

We first consider each of the models in the presence of the equity premium only. Thus, we rule out risk premia for volatility and jump shocks, and also the explanations based on estimation risk or Peso problems. We consider the simplified setting to understand the role of the underlying index dynamics in generating index returns.

4.1 Black-Scholes

4.1.1 Analytical expected returns

In the Black-Scholes model, the equity premium, volatility and moneyness levels determine EORs. Table 4 computes analytical EORs for various initial moneyness levels using equation (9). The cum-dividend equity premium ranges from 4% to 8% and volatility ranges from 10% to 20%. Black-Scholes EORs are large in magnitude, negative, and quite sensitive to the equity premium and volatility, especially for OTM strikes. For example, expected put returns are on the order of -10% to -25% per month for ATM strikes, and -10% to -50% per month for OTM strikes.

Put EORs are negatively related to the equity premium. As expected returns increase, the underlying index drifts upward more strongly resulting in fewer in-the-money (ITM) put expirations, and, conditional on an ITM expiration, lower payoffs. The impact is quantitatively large, as the expected put option return differences between high and low equity premiums is around 10% for ATM strikes and even more for deep OTM strikes. This sensitivity points to a number of important issues in interpreting historical option returns.

First, any period of time that is “puzzling” in terms of large realized equity returns, will generate option returns that are even more striking. For example, the behavior of aggregate equity index returns in the 1990s were particularly puzzling for both academics and practitioners. The realized equity premium from 1990 to 1999 was 9.4%. Assuming this realized premium was expected and combining it with the below average volatility of 13%

		Moneyness						
σ	μ	0.94	0.96	0.98	1.00	1.02	1.04	1.06
10%	4%	-27.6	-22.5	-17.6	-13.3	-9.7	-6.9	-5.0
	6%	-38.7	-32.2	-25.7	-19.7	-14.5	-10.5	-7.7
	8%	-48.3	-40.8	-33.1	-25.7	-19.2	-14.1	-10.4
15%	4%	-15.4	-13.0	-10.8	-8.8	-7.1	-5.6	-4.5
	6%	-22.5	-19.2	-16.1	-13.2	-10.7	-8.6	-6.9
	8%	-29.1	-25.0	-21.1	-17.5	-14.3	-11.5	-9.3
20%	4%	-10.3	-8.9	-7.7	-6.5	-5.5	-4.6	-3.9
	6%	-15.2	-13.3	-11.5	-9.9	-8.4	-7.1	-6.0
	8%	-20.0	-17.6	-15.3	-13.2	-11.2	-9.5	-8.1

Table 4: Population expected returns in the Black-Scholes model. The parameter μ is the cum-dividend equity premium, σ is the volatility. These parameters are reported on an annual basis, and expected options returns are monthly percentages.

over the same period, the 6% OTM and ATM EORs would, according to the Black-Scholes model, be about -40% and -23% , respectively. This shows how potentially sensitive EORs are to the underlying index returns.

Second, the impact of the equity premium on EORs is approximately linear. To see this, holding $\sigma = 15\%$ constant, the expected ATM put returns are -9% for $\mu = 4\%$, and are -18% for $\mu = 8\%$. If each outcome is equally likely, the average is -13.20% and is quite close to the expected put return of -13.27% when $\mu = 6\%$. This means that if the equity premium is time-varying and the likelihoods of high and low outcomes are roughly equal, uncertainty over the equity premium, per se, has little impact on average option returns. As the example in the previous section indicates, however, there are long time periods with relatively high or low premia, so it is important to account for the equity premium when analyzing option returns.

Third, abstracting from time-variation, estimates of a presumed constant equity premium are notoriously unreliable, heavily dependent on the data period used, and may be subject to structural breaks (see, e.g., Pastor and Stambaugh (2001)). It is not obvious how an unobserved equity premium impacts option prices. In continuous-time, the uncertainty

only affects the mean estimates, as the total volatility is unchanged.¹⁴ In discrete-time, uncertainty over the equity premium increases total volatility of returns. The key issue is the degree to which \mathbb{P} and \mathbb{Q} distributions are affected by the equity premium uncertainty.

EORs are highly sensitive to volatility. As volatility increases, expected put option returns become less negative, and the effect is substantial. For both ATM and OTM puts, increasing volatility from 10% to 20% approximately halves EORs. For 6% OTM puts, EORs change from -15% when $\sigma = 20\%$ to -40% when $\sigma = 10\%$. Thus volatility has a quantitatively large impact and its impact varies across strikes.

Further, unlike the approximately linear relationship between EORs and the equity premium, the relationship between put EORs and volatility is concave. Based on the example in the previous paragraph, if we assume that high ($\sigma = 20\%$) and low ($\sigma = 10\%$) volatility levels are equally likely, the average 6% OTM put EOR is -27% compared to -23% when $\sigma = 15\%$. This concavity implies that fully anticipated time-variation in volatility results in lower average option returns than that if volatility were constant at the average value.

In practice, this concavity is exacerbated by the fact the volatility levels are highly skewed to the right. This implies that large values of volatility are more likely than small ones and there are more volatility observations to the left of the mean than the right. These properties have important implications for put option returns because of Jensen's inequality as the following example illustrates. Suppose that volatility can take one of three values $\sigma = (10\%, 15\%, 40\%)$ with probability $(0.5, 0.4, 0.1)$, which averages to 15%. Then the average expected put return is -30% for 6% OTM puts and -16% for ATM puts, compared to -23% and -13% , respectively, when $\sigma = 15\%$. Thus, the concavity has two important implications: it decreases EORs, and its effect is stronger for deeper OTM strikes. This will be quantitatively important for interpreting observed option returns below.

4.1.2 The distribution of average option returns

Next, we evaluate the significance of the observed returns statistics using the finite sample distribution constructed from the Black-Scholes model. As an illustration, the first panel in

¹⁴If the equity premium is time-varying and unobserved, then an equilibrium model is needed to derive option prices. For work along this dimension, see David and Veronesi (2006). Buraschi and Jiltsov (2006) price options in a model with heterogeneous beliefs.

Figure 2 shows the finite sample distribution for 6% OTM average put returns. The solid vertical line is the observed sample value. We compute p -values of the observed return with respect to this finite sample distribution by recording the percentage of simulated statistics that are below the observed ones. The upper panel displays the dramatic skewness of the finite sample distribution, which is expected given the strong positive skewness of purchased put options, and points to the inaccuracies of normal approximations in samples of our size. Of note is large variability in average put return estimates: the (5%, 95%) band is -65% to $+28\%$.

The first line of the first panel of Table 5 summarizes EORs and p -values corresponding to observed average returns for various moneyness categories. The first thing to notice is that the p -values for average returns have increased dramatically relative to Table 1. For example, the p -values using standard t -statistics for the ATM options increases by roughly a factor of 10 and by a factor of more than 10000 for deep OTM put options. This dramatic increase occurs because our bootstrapping procedure accounts for the fact that expected Black-Scholes returns are quite negative for this strike, providing a proper anchor for hypothesis tests, and the distribution of average OTM returns is extremely dispersed (large sampling uncertainty).

Next, note that average 6% OTM option returns are not strongly statistically different from those generated by the Black-Scholes model. In sample, the average 6% OTM put return is about -60% , which corresponds to a p -value of just over 8%, indicating borderline insignificance or at least a lack of strong significance. Turning to the other moneyness categories, the Black-Scholes model has more difficulty generating option returns for the ATM strikes (0.98 and 1.00), although the p -values still increase dramatically.

Based only on the Black-Scholes model, we have our first striking conclusion: of all the statistics we analyze, the deep OTM put returns are always the least significant in a statistical sense. This is particularly interesting since the results in the previous literature typically conclude that the deep OTM put options are the most anomalous or mispriced. We find the exact opposite conclusion: OTM puts are not strongly inconsistent with the Black-Scholes model. This result shows the importance of properly anchoring hypothesis tests and performing finite sample inference.

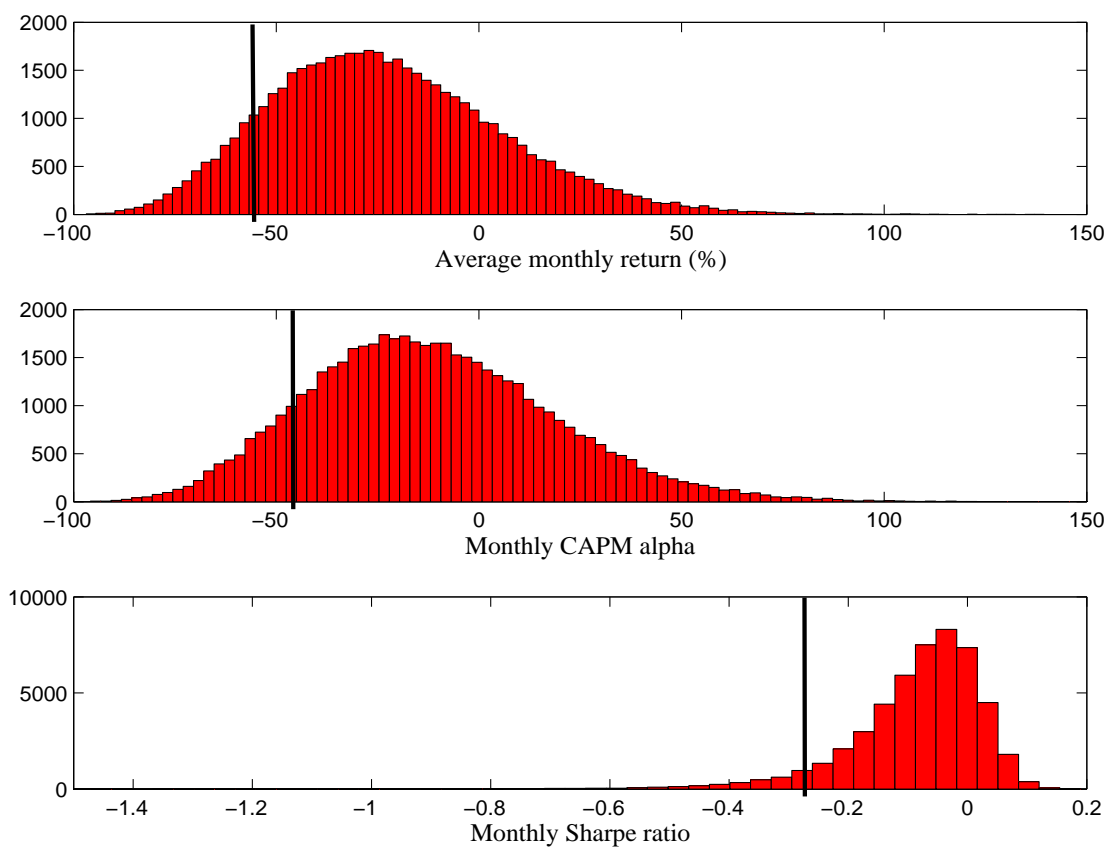


Figure 2: This figure provides the finite sample distribution of various statistics. The top panel provides the distribution of average 6% OTM put returns, the middle panel 6% OTM put CAPM alphas, and the bottom panel 6% OTM put Sharpe ratios. The solid vertical line is the observed value from the data.

		Moneyness	0.94	0.96	0.98	1.00	Strdl
		Data, %	-56.8	-52.3	-44.7	-29.9	-15.7
Average returns	BS	$E^{\mathbb{P}},\%$	-20.6	-17.6	-14.6	-12.0	1.1
		p -value,%	8.1	1.7	0.4	2.2	0.0
	Merton	$E^{\mathbb{P}},\%$	-15.0	-15.4	-14.2	-12.2	1.4
		p -value,%	9.2	2.9	0.9	3.2	0.0
	SV	$E^{\mathbb{P}},\%$	-25.8	-21.5	-17.5	-13.7	1.4
		p -value,%	24.1	9.3	3.0	7.3	0.0
	SVJ	$E^{\mathbb{P}},\%$	-10.4	-11.6	-14.2	-13.9	2.2
		p -value,%	14.1	6.4	3.3	8.3	0.1
		Data, %	-48.3	-44.1	-36.8	-22.5	-15.4
CAPM α s	BS	$E^{\mathbb{P}},\%$	-17.9	-15.3	-12.7	-10.4	0.1
		p -value,%	12.6	2.7	0.3	1.2	0.0
	Merton	$E^{\mathbb{P}},\%$	-11.1	-12.2	-11.7	-10.3	0.9
		p -value,%	12.6	3.8	0.6	1.7	0.2
	SV	$E^{\mathbb{P}},\%$	-23.6	-19.5	-15.8	-12.4	0.6
		p -value,%	39.1	14.1	3.4	8.7	0.0
	SVJ	$E^{\mathbb{P}},\%$	-7.3	-9.0	-12.3	-12.5	-0.4
		p -value,%	19.7	9.9	3.4	9.3	0.2
		Data, %	-27.3	-29.0	-29.0	-23.4	-24.3
Sharpe ratios	BS	$E^{\mathbb{P}},\%$	-5.2	-6.8	-8.1	-9.1	2.0
		p -value,%	4.9	1.9	1.2	4.0	0.0
	Merton	$E^{\mathbb{P}},\%$	-3.1	-4.9	-6.9	-8.5	1.3
		p -value,%	5.4	2.4	1.4	4.1	0.1
	SV	$E^{\mathbb{P}},\%$	-4.1	-6.9	-9.2	-10.5	0.9
		p -value,%	21.5	12.0	7.7	14.3	0.0
	SVJ	$E^{\mathbb{P}},\%$	-1.7	-2.6	-5.2	-8.8	-0.3
		p -value,%	11.1	6.2	4.5	10.9	0.1

Table 5: This table reports population expected options returns, CAPM α 's, and Sharpe ratios and finite sample distribution p -values for four models: Black-Scholes (BS), Merton, stochastic volatility (SV) and stochastic volatility with jumps (SVJ). We assume that all risk premia (except for the equity premium) are equal to zero. The column "Strdl" refers to the statistics associated with at-the-money straddles.

4.1.3 Risk adjustment/delevering

In this section we evaluate three common approaches for risk-correcting or de-levering option returns: computing CAPM alphas, computing Sharpe ratios, and analyzing straddle returns. As mentioned earlier, covered put positions, which take leverage into account by adding a long position in the underlying index to a put position, are economically and statistically insignificant, so we do not consider them.

The first risk correction we consider is the linear factor model alphas. This is one of the most common methods of risk correcting, as it has been used in the option pricing setting by Jackwerth (2000), Coval and Shumway (2001), Bondarenko (2003), Santa-Clara and Saretto (2005), and Driessen and Maenhout (2006). As mentioned earlier, these regressions only hold for the Black-Scholes model and for instantaneous returns.

We compute population CAPM alphas, which are reported in the second panel in Table 5. For the Black-Scholes model for every strike, the alphas are quite negative and their magnitudes are economically large, ranging from -18% for 6% OTM puts to -10% for ATM puts. Although Black-Scholes is a single-factor model, the alphas are strongly negatively biased in population. This outcome is due to the fact that the regression tries to fit a straight line to kinked payoff, and shows the fundamental problem that arises when applying linear factor models to option returns.

To see the issue more clearly, Figure 3 displays the results of two simulated time series. Both cases correspond to 215 monthly index returns and 6% OTM put option returns simulated from the Black-Scholes model with the OLS fitted regression line. The regression estimates in the top (bottom) panel correspond to $\alpha = 64\%$ ($\alpha = -51\%$) per month and $\beta = -58$ ($\beta = -19$). The main difference between the two simulations is a *single* large observation in the upper panel, which generates drastically different results. A single large outlier can substantially shift the constant and intercept as the estimates are obtained by minimizing squared errors. The idea that factor model alphas are inappropriate when analyzing option returns is, of course, not new. It is, however, surprising how often researchers use these regressions to risk-correct option returns.

To analyze the issues more formally, the middle panel of Figure 2 illustrates the finite sample distribution of CAPM alphas for the case of 6% OTM puts, and the middle panel of Table 5 provides finite sample p -values for the observed alphas. For the deepest OTM puts, observed CAPM alphas are insignificantly different from those generated by the Black-Scholes model. For the other strikes, the observed alphas are generally too low to be

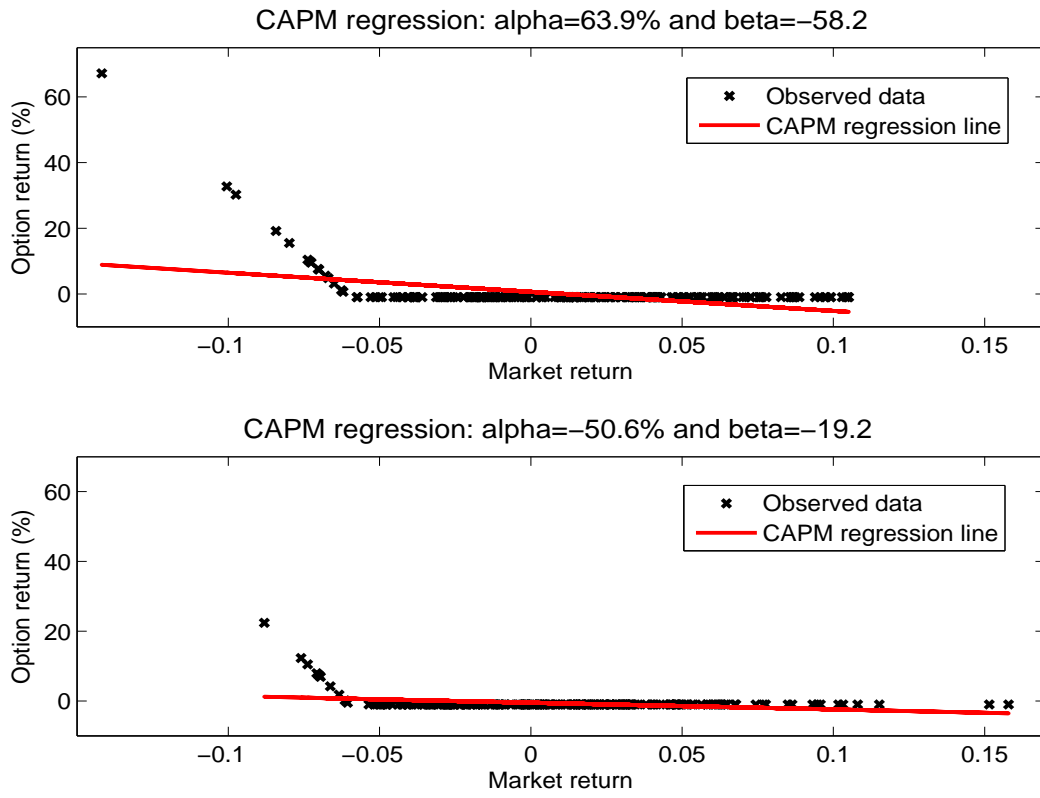


Figure 3: CAPM regressions for 6% OTM put option returns.

consistent with the Black-Scholes model, although again the p -values are much larger than those based on asymptotic theory.

We next consider Sharpe ratios. Sharpe ratios primarily account for leverage by scaling average excess returns by volatility. Formally, Sharpe ratios are only useful when returns are normally distributed, or if investors have mean-variance preferences. Despite this shortcoming, Sharpe ratios are commonly used both in academic studies and in practice and are a useful, if imperfect, metric. The bottom panel of Figure 2 illustrates the finite sample distribution of Sharpe ratios for 6% OTM puts. Since some simulated paths have very few OTM expirations, the distribution of Sharpe ratios is extremely skewed to the left. The third panel of Table 5 reports population Sharpe ratios for put options of various strikes and finite sample p -values. As a comparison, the monthly Sharpe ratio of the underlying index over our sample period is 10.5%. The Sharpe ratios are modestly statistically significant for every strike, with p -values between 1% and 5%.

Finally, we report average ATM straddle returns in the last column of the first line of the first panel. These options positions are approximately delta-neutral, as the call and put exposures roughly offset exposure to the underlying. The results indicate that straddle returns are highly statistically significant for the Black-Scholes model. The p -values imply that the straddles are in fact the most significant of the statistics that we consider.

We conclude that although the Black-Scholes model does appear to be inconsistent with option return data, we see that deep OTM average put returns are not strongly significant, either in levels or risk-corrected. The more difficult statistics to explain are ATM put and straddle returns.

4.2 Stochastic Volatility and Jumps

4.2.1 Merton's model

We next consider option returns generated by Merton's jump-diffusion model. This model accounts for rare crashes, which generate occasionally large positive put option returns, and in doing so, generates implied volatility smiles. This will allow us to assess the impact of jumps on option returns. We first consider the case without priced jump risk, allowing us to focus on the direct impact of jumps in the data-generating process.

First, consider expected put returns generated by Merton's model, computed analytically in Table 5 in the first panel in the row labeled ' E^P .' The results indicate, surprisingly,

that expected returns in Merton’s model are less negative than those in the Black-Scholes model. Moreover, the expected returns are slightly non-monotonic, as they increase for the 6% OTM strike relative to the 4% OTM strike. For ATM (OTM) options, Merton’s model generates EORs of -12% (-15%) compared to -12% (-21%) for the Black-Scholes model.

Merton EORs are less negative than Black-Scholes EORs because jump risk is not priced. The presence of price jumps increases the left tail mass in the distribution of returns in a similar manner under both \mathbb{P} and \mathbb{Q} measure. To see this, note that equation (9) implies that a factor increasing tail mass similarly under both measures will actually make EORs less negative. For example, for Merton’s model the numerator and denominator are 0.182 and 0.155 (per hundred dollars in ATM strike) compared to 0.144 and 0.111 for Black-Scholes, which generates less negative returns for Merton’s model.

This result provides a useful insight into the relationship between the implied volatility smile and the determinants of option returns. Despite the fact that Merton’s model can generate steep implied volatility curves, the model generates less negative EORs. Moreover, the more negative the jumps, the less negative the EORs. For example, if we decrease the jump mean parameter to -10% under both probability measures, the expected returns become even less negative, only -7% for 6% OTM options. Unpriced jump risks increase the prices of OTM put options relative to Black-Scholes, but also increase the \mathbb{P} -measure expected payoffs, thus both the numerator and denominator of the expected returns in (8) increase. As in the Black-Scholes model, the only difference between \mathbb{P} and \mathbb{Q} measures is the difference in drifts generated by the equity premium, which, due to the higher \mathbb{P} and \mathbb{Q} measure expectations, generates a smaller impact for OTM options. Thus, ATM returns are similar in the two models, but deep OTM option returns are less negative.

This provides another clue to understanding the sources of put and straddle returns: an extremely steeply sloped implied volatility smile will not help in generating realistic put or straddle returns, unless the steepness is generated by a gap between the \mathbb{P} and \mathbb{Q} measures. Thus, a steep implied volatility curve, in and of itself, provides no information about whether or not options are overpriced or misspriced.

Turning to the finite sample results, we see that jumps in prices also have an important impact. In every case, Merton’s model generates higher p -values, despite the fact that the EORs are less negative. Jumps significantly fatten the tails of the finite sample distribution as simulated samples with slightly fewer jumps than expected have more negative average returns, and those with slightly more jumps than expected have less negative returns,

generating more finite sample uncertainty. In terms of risk-corrections, population alphas are less negative than in Black-Scholes, straddle returns are larger, and Sharpe ratios are smaller than in the Black-Scholes model. In conclusion, the addition of jumps in prices does not generate more realistic EORs, and in fact, generates option returns that are smaller in absolute value than those in the Black-Scholes model.

4.2.2 Stochastic volatility and jumps

Next, we consider how the addition of stochastic volatility affects our previous conclusions, as we consider the SV and SVJ models, which extend Black-Scholes and Merton by incorporating fluctuating volatility. Table 5 provides population average returns, CAPM alphas, Sharpe ratios, and straddle returns for the SV model, as well as p -values. We do not assume that this volatility risk is priced, that is, we set $\theta_v^Q = \theta_v^P$.

As argued in Section 3.2.1, EORs are a concave function of volatility, which implies that fluctuations in volatility, even if fully anticipated, will increase absolute EORs. The results indicate that fluctuating volatility has an important quantitative impact on put returns. Average returns decrease about 2% for ATM puts and more than 5% for 6% OTM strikes. More importantly, the p -values increase dramatically. For example, for 6% OTM strikes, the p -values for average returns, CAPM alphas, and Sharpe ratios are 24%, 39%, and 22%, respectively. This indicates that roughly one in four simulated sample paths generate average 6% OTM put returns that are *more negative* than those observed in the data. Moreover, this is in a model in which volatility risk is not priced in options. This indicates that there is absolutely nothing puzzling about deep OTM put returns, at least relative to standard models. This conclusion is in strong contrast to the existing literature, and is one of our primary results.

Regarding the risk-adjusted put return statistics, we see that most moneyness/statistic combinations are insignificant, with the exception of the 0.98 moneyness average returns and CAPM alphas, which have p -values around 3%. Of note, CAPM alphas are even more biased in population as the α for ATM (6% OTM) strikes is now -12% (-24%). In every case, the SV model generates much higher p -values than the Black-Scholes model, in some cases more than five times higher. For completeness, we also consider the case of the SVJ model. The SVJ model, consistent with the results from Merton's model, generates less negative population values than the SV model does, but p -values are of similar magnitudes. Of particular note is the fact that the p -values for the straddles barely change as we move

through the models.

These results show the importance of our methodology, and in particular, of properly anchoring hypothesis tests and basing tests on finite sample distributions. After computing exact expected returns and properly accounting for finite sample variation of average option returns, the statistics associated with put returns do not seem particularly surprising. Moreover, this conclusion is based on standard models and factors without relying on alternative explanations such as factor risk premia, estimation risk, or Peso problems. Thus, we conclude that there is nothing puzzling about put returns, especially OTM put returns. This finding is in strong contrast to the papers cited earlier.

Our methodology also uncovers a statistic that is potentially puzzling: ATM straddle returns. These returns are generated by the well known gap between realized and implied volatility. In our sample, realized volatility is approximately 15% while ATM implied volatility averages 17%. This gap, which is largely robust over subsamples, generates the large negative straddle returns.¹⁵ For example, during the last two one-year periods from July 2003 to July 2004 and July 2004 to July 2005 the gap was 5.3% and 1.9%, respectively. Options are consistently priced with higher volatility than is subsequently realized. In the next section, we explore potential explanations for this gap and the steepness of the observed implied volatility smile that include jump risk premia, estimation risk, and Peso problems.

5 Risk premia, estimation risk, and Peso problems

5.1 Differences between \mathbb{P} and \mathbb{Q}

In this section we evaluate how gaps between \mathbb{P} and \mathbb{Q} measures generated by factor risk premia, estimation risk, and Peso problems impact put option returns, and more importantly, straddle returns.

We first note that one potential explanation for the negative straddle and put returns is a diffusive volatility risk premium, generated by a gap between $\theta_v^{\mathbb{P}}$ and $\theta_v^{\mathbb{Q}}$ (see, e.g., Coval and Shumway, 2001). This, however, is unlikely to be a main or even a significant driver of short-dated straddle returns. The argument is relatively simple. A volatility risk premium affects expected returns through the term $\beta_t^v \kappa_v^{\mathbb{P}} (\theta_v^{\mathbb{P}} - \theta_v^{\mathbb{Q}})$ in equation (7).

¹⁵Bakshi and Madan (2006) link this gap to the skewness and kurtosis of the underlying returns via the representative investor's preferences. Chernov (2007) relates this gap to volatility and jump risk premia.

Because volatility is highly persistent, $\kappa_v^{\mathbb{P}}$ is small. Combined with the fact that we are analyzing short-dated, monthly option returns, $\theta_v^{\mathbb{Q}}$ would need to be much larger than $\theta_v^{\mathbb{P}}$ to generate negative enough monthly straddle returns.¹⁶ For example, our computations show that $\sqrt{\theta_v^{\mathbb{Q}}} = 22\%$ would generate straddle returns that are statistically insignificant from those observed. However, such a high risk-neutral average volatility implies that the term structure of implied volatilities would be steep and upward sloping, on average, which can be rejected based on observed implied volatility term structures (see also Broadie, Chernov, and Johannes, 2007). Therefore, a very high volatility risk premium can be rejected as the sole explanation for the gap generated short-dated straddle returns.

A more promising explanation is that risk-neutral and objective measure perceptions of the jump parameters are different. Differences in jump parameters between \mathbb{P} and \mathbb{Q} have the advantage that they have a first-order impact on short-dated options. From a mechanical standpoint, it is always possible to increase the jump risk under \mathbb{Q} (via $\mu_z^{\mathbb{Q}}$, $\sigma_z^{\mathbb{Q}}$ or $\lambda^{\mathbb{Q}}$) to generate the straddle returns. One way to do this is to estimate these parameters solely from option data, as in Broadie, Chernov, and Johannes (2007). They find that these estimates are consistent with observed option returns. This, of course, is circular as they used the option prices to estimate the parameter in the first place.

We take a different approach. As reviewed above, the literature has introduced (at least) three potential explanations for the observed option returns: factor risk premia, estimation risk, and Peso problems. All three of these explanation generate differences between \mathbb{P} and \mathbb{Q} . In analyzing these explanations, our goal is to re-evaluate them using our methodology and common option pricing models. At some level, we are going to quantify, in a parametric sense, how far these explanations need to be pushed to generate option returns consistent with the data.

In the following three subsections, we discuss how we calibrate these explanations, with specific results.

5.1.1 Factor risk premia

General equilibrium models provide a natural starting point for generating factor risk premia. For our purposes, we need (a) models that incorporate important option-relevant features such as time-varying volatility and price jumps and (b) estimates of the \mathbb{P} -measure

¹⁶We constrain $\kappa_v^{\mathbb{P}} = \kappa_v^{\mathbb{Q}}$. Some authors have found that $\kappa_v^{\mathbb{Q}} < \kappa_v^{\mathbb{P}}$, which implies that $\theta_v^{\mathbb{Q}}$ would need to be even larger to generate a noticeable impact on expected option returns.

parameters that closely match the historical experience of observed stock index returns. In particular, we need to choose \mathbb{P} -measure parameters to exactly match the observed equity premium and volatility over our sample. Otherwise, if we assumed that options were priced with a lower equity premium (consistent with some simple equilibrium models), we would substantially understate put option returns.

The main problem with directly applying standard general equilibrium models such as Bates (1988) or Naik and Lee (1990) is one of calibration. These authors introduce extensions of the standard log-normal diffusion model incorporating jumps in dividends. The main problem with these models is that, when calibrated to dividends, they lead to well-known asset pricing puzzles such as equity premium and excess volatility puzzles.¹⁷ When these models are applied to option pricing applications the problems are even more severe, as we know very little about the equilibrium sources of jumps in prices and stochastic volatility or their connections to dividend and consumption growth. For example, the connections between fluctuating stock index volatility, jumps in prices, or the leverage effect and underlying economic uncertainty is not at all clear. At some level, the equilibrium models and option-relevant features of stock index returns operate on a different time scale.

Given this caveat, we would still like to explore how jump risk premia affect option returns. To do this, we use the functional forms of the risk correction for the jump parameters, but we fix the overall equity premium and the level of volatility to be consistent with our observed historical data on index returns, 5.4% and 15%, respectively. The risk corrections are given by (see Bates, 1988, or Naik and Lee, 1990)

$$\lambda^{\mathbb{Q}} = \lambda^{\mathbb{P}} \exp \left(\mu_z^{\mathbb{P}} \gamma + \frac{1}{2} \gamma^2 (\sigma_z^{\mathbb{P}})^2 \right) \quad (11)$$

$$\mu_z^{\mathbb{Q}} = \mu_z^{\mathbb{P}} - \gamma (\sigma_z^{\mathbb{P}})^2, \quad (12)$$

where γ is the risk aversion parameter, and the \mathbb{P} -measure parameters are those estimated from stock index returns (and not dividend or consumption data). Notice that the volatility of jump sizes, σ_z , is the same across both probability measures. We re-iterate that our goal is not to impose a particular equilibrium model in order to understand the connections between dividends, consumption, and stock index returns, but rather to understand the

¹⁷Benzoni, Collin-Dufresne, and Goldstein (2006) extend the Bansal and Yaron (2004) model to incorporate rare jumps in the latent dividend growth rates. They show that this model can generate a reasonable volatility smile, but they do not analyze the issues of straddle returns, or equivalently, the difference between implied and realized volatility. Their model does not incorporate stochastic volatility.

	$\lambda^{\mathbb{Q}}$	$\mu_z^{\mathbb{Q}}$	$\sigma_z^{\mathbb{Q}}$	$\sqrt{\theta_v^{\mathbb{Q}}}$
Risk Premia	1.51	-6.85%	$\sigma_z^{\mathbb{P}}$	$\sqrt{\theta_v^{\mathbb{P}}}$
Estimation Risk	1.25	-4.96%	6.99%	14.79%
Peso problem	2.73	$\mu_z^{\mathbb{P}}$	$\sigma_z^{\mathbb{P}}$	$\sqrt{\theta_v^{\mathbb{P}}}$

Table 6: \mathbb{Q} -parameters. We report the parameter values for the various \mathbb{P} -measure adjustment scenarios that we explore. In addition, in the estimation risk scenario, we value options with the spot volatility $\sqrt{V_t}$ incremented by 0.5%.

connections between stock index returns and option returns using standard risk adjustments.

We consider the benchmark case of $\gamma = 10$. This is certainly in the range of values considered to be reasonable in applications. From (11) and (12), this value generates $\lambda^{\mathbb{Q}} = 1.51$ and $\mu_z^{\mathbb{Q}} = -6.85\%$, i.e., investors price options as if there will be 1.51 jumps per year even though on average only 0.91 will be realized ($\lambda^{\mathbb{Q}}/\lambda^{\mathbb{P}} = 1.65$) and that $\mu_z^{\mathbb{P}} - \mu_z^{\mathbb{Q}} = 3.6\%$, i.e., investors price options as if mean jump sizes are 3.6% less than those realized. The \mathbb{Q} -parameter values are given in Table 6. We do not consider a stochastic volatility risk premium, $\theta_v^{\mathbb{P}} < \theta_v^{\mathbb{Q}}$, as standard equilibrium models do not incorporate time-varying volatility.

There are other theories that generate similar gaps between \mathbb{P} and \mathbb{Q} jump parameters. Given the difficulties in estimating the jump parameters, Liu, Pan, and Wang (2004) consider a representative agent who is averse to the uncertainty over jump parameters. Although their base parameters differ, the \mathbb{P} and \mathbb{Q} measure gaps they generate for their base parametrization and the “high-uncertainty aversion” case are $\mu_z^{\mathbb{P}} - \mu_z^{\mathbb{Q}} = 3.9\%$ and $\lambda^{\mathbb{Q}}/\lambda^{\mathbb{P}} = 1.96$, which are similar in magnitude to those that we consider.¹⁸ We do not have a particular vested interest in the standard risk-aversion explanation vis-a-vis an uncertainty aversion explanation, our only goal is to use a reasonable characterization for the difference between \mathbb{P} and \mathbb{Q} jump parameters.¹⁹

¹⁸Specifically, Liu, Pan, and Wang (2003) assume that $\gamma = 3$, the coefficient of uncertainty aversion $\phi = 20$, and the penalty coefficient $\beta = 0.01$. The \mathbb{P} -measure parameters they use are $\lambda^{\mathbb{P}} = 1/3$, $\mu_z^{\mathbb{P}} = -1\%$ and $\sigma_z^{\mathbb{P}} = 4\%$. We thank Jun Pan for helpful discussions regarding the details of their calibrations.

¹⁹An additional explanation for gaps between \mathbb{P} and \mathbb{Q} jump parameters is the argument in Garleanu, Pedersen, and Poteshman (2006). Although they do not provide a formal parametric model, they argue

5.1.2 Estimation risk

Another explanation for observed option returns is estimation risk. Estimation risk captures the idea that parameters are unobserved and inaccurately estimated. Alternatively, the market makers cannot perfectly hedge in Garleanu, Pedersen, and Poteshman (2005), and, therefore, estimation risk could play an important role and be priced. In our context, estimation risk arises because it is difficult to estimate the parameters and spot volatility in our models. In particular, jump intensities, parameters of jump size distributions, long-run mean levels of volatility, and volatility mean reversion parameters are all notoriously difficult to estimate. Spot volatility is also not observed. While uncertainty in drift parameters in the stochastic volatility process will have a minor impact on short-dated options, the uncertainty in jump parameters can have a first order impact.²⁰

To see this, consider a standard Bayesian setting for learning about the parameters of the jump distribution.²¹ First, consider uncertainty over the jumps mean parameter: jump sizes are given by $Z_j = \mu_z + \sigma_z \varepsilon_j$ and $\mu_z \sim \mathcal{N}(\mu_0, \sigma_0^2)$. Then, the predictive distribution of Z_{k+1} upon observing k previous jumps is given by

$$p\left(Z_{k+1} | \{Z_j\}_{j=1}^k\right) \sim \mathcal{N}\left(\mu_k, \sigma_k^2\right),$$

where

$$\begin{aligned} \mu_k &= w_k \mu_0 + (1 - w_k) \bar{Z}_k, \quad \bar{Z}_k = k^{-1} \sum_{j=1}^k Z_j \\ \sigma_k^2 &= \left(\frac{k}{\sigma_z^2} + \frac{1}{\sigma_0^2}\right)^{-1} + \sigma_z^2, \quad w_k = \frac{\sigma_z^2/k}{\sigma_z^2/k + \sigma_0^2}. \end{aligned}$$

In addition to revising one's beliefs about the location, we also see that $\sigma_k^2 > \sigma_z^2$, implying that learning generates excess volatility. Quantitatively, its impact will be determined by prior beliefs and how many jumps have been observed. In practice, one would expect even more excess volatility, as jump sizes are not perfectly observed.

that market incompleteness generated by jumps or the inability to trade continuously, combined with exogenous demand pressure, qualitatively implies gaps between realized volatility and implied volatility.

²⁰Eraker, Johannes, and Polson (2003) provide examples of the estimation uncertainty impact on the implied volatility smiles.

²¹Benzoni, Collin-Dufresne, and Goldstein (2006) consider uncertainty over the mean parameters with normal priors. Johannes, Polson, and Stroud (2005) consider sequential learning about jump and stochastic volatility parameters in jump-diffusion setting using historical S&P 500 index returns.

The impact of uncertainty on σ_z is even greater. Assuming that μ_z is known, an inverse-gamma prior on the jump variance, $\sigma_z^2 \sim \mathcal{IG}$, and that jumps are observed without errors (which is true in continuous-time), the predictive distribution of the jump sizes is t -distributed:

$$p\left(Z_{k+1} - \mu_z \mid \{Z_j\}_{j=1}^k\right) \sim t_\nu,$$

where the degrees of freedom parameter ν depends on the prior parameters and sample size (Zellner, 1971, section 3.2.4). To compute prices, expectations of the form $E(\exp(Z_{k+1}) \mid \mathcal{F}_k)$ will have to be computed. However, if the jump sizes have a t -distribution, this expectation may not exist because the moment generating function of a t -random variable does not exist. Thus, parameter uncertainty can have a substantial impact on the conditional distribution of S_t , as the two examples demonstrate. A potentially even greater source of uncertainty is the functional form of the jump distribution.

Apart from the impact of parameter uncertainty, it is important to consider difficulties in estimating spot volatility. Even with high frequency data, there are dozens of different methods for estimating volatility, depending on the frequency of data assumed and whether or not jumps are present. Similarly, one could argue that it is possible to estimate V_t from options, but this requires an accurate model and parameter estimates. In practice, any estimate of V_t is a noisy measure because of all these factors.

To capture the spirit of estimation risk, without introducing a formal model for how investors calculate and price estimation risk, we consider the following intuitive approach. We assume that the parameters that we report in Table 3 represent the true data-generating process, that is, these parameters generated the observed S&P 500 index returns over our sample period. However, investors priced options using different parameters due to the estimation error. For simplicity, we assume that \mathbb{Q} -measure parameters were increased/decreased by one standard deviation from the \mathbb{P} -parameters reported in Table 3. Likewise, we assume that the spot volatility was adjusted by the posterior standard deviation. In our sample, the average posterior standard deviation of the spot volatility is 0.5%.

For example, denoting standard deviation by std , we set $\mu_z^{\mathbb{Q}} = \mu_z^{\mathbb{P}} - std(\mu_z^{\mathbb{P}})$ and $\lambda^{\mathbb{Q}} = \lambda^{\mathbb{P}} + std(\lambda^{\mathbb{P}})$. The spot volatility $\sqrt{V_t}$ was adjusted upward by 0.5%. The full set of assumed parameter values is reported in the second line of Table 6.

5.1.3 Peso problems

Another explanation for the observed option returns is Peso problems. In this scenario, the observed samples are unrepresentative. Potentially, this could mean that fewer large jumps were observed or that the realized stochastic volatility path had different characteristics than the true index return process and that investors were aware of this bias. Our finite sample simulations do not account for an unrepresentative sample because the parameter estimates are based on the observed sample. One way to think of the impact of unrepresentative samples is via Table 1. If, for example, one or two more periods like 2000 to 2003 were observed, than put and straddle returns would be far less negative.

Thus, general Peso problems could apply to multiple aspects of our model: the parameters of the jump distribution, the jump intensity, parameters of the volatility process, or volatility paths. For simplicity, we consider Peso problems through the lens of the jump intensity. This is the common way of analyzing the problem. In the context of model and parameter estimates, we increase the jump intensity threefold from 0.91 to 2.73. In our model, jump sizes are modest on average, thus our assumption implies that two additional modestly sized jumps were anticipated to arrive by investors. Alternatively, we could have assumed that jump intensity increases were smaller, but that the sizes (in terms of means or variances) were larger. We chose this case for parsimony.

Our assumption can be viewed as modest for two reasons. First, in reality, jump times and sizes are not observed, and from a statistical standpoint, it is difficult to disentangle our relatively modestly-sized jumps from time-varying volatility. Distinguishing between 0.91 or 2.73 small jumps per year is clearly a difficult statistical discrimination problem, which is even more difficult if the sample is unrepresentative. Second, previous evaluations of Peso problems in option prices have argued that quite severe assumptions are required to explain the observed option returns. For example, Bondarenko (2003) estimates that his 13-year sample would require an additional 18 crashes of the magnitude of the one in 1987, while Jackwerth (2000) comes up with a more modest, but still large, frequency of one 1987-size crash in four years. In contrast, even if we take the most conservative assumptions outlined in section 3.4 (a three-standard deviations diffusive move, spot volatility of 25% as on the day of crash of 1987) and combine them with our Peso quantification, the probability of a movement as large as the crash of 1987, is once every 47 years under \mathbb{Q} .

5.2 Results

Table 7 reports population values and p -values corresponding to factor risk premia, estimation risk, and Peso problem explanations corresponding to the parameter values in Table 6. First, note that all three mechanisms generate *expected* option returns that are similar in magnitude to the *average* option returns observed in the data. Compare these magnitudes to the ones in the zero-risk-premia case reported in Table 5. As we observed, the SVJ-based expected returns were lower (in absolute value) than the ones in the SV model. Thus, jump risk premia play a very important role in generating the return magnitudes. Large magnitudes of expected returns imply that observed returns should be insignificant based on finite sample distributions, which the p -values confirm. Similarly, all three explanations have no difficulty explaining CAPM alphas and Sharpe ratios.

Straddle returns are again the hardest to explain, but our explanations go a long way in understanding the magnitudes of these returns. The estimation risk story has the most difficult time explaining the returns with a small, but respectable, p -value of 5%. A less modest, but still reasonable parameter adjustment from one-standard deviation to two standard deviations generates much larger p -values and population straddle returns close to those observed. The p -values for the factor risk premium and the Peso explanation cases are about 8% and 15%, respectively, indicating insignificance or at least an absence of strong significance.

In reality, all three features are plausible explanations, and therefore a combined explanation based on modest risk aversion, modest estimation risk, and modest Peso problems will also be able to explain the observed returns. We also did not consider explanations based on learning, price jump variance risk premia, model misspecification, or more complicated models incorporating, for example, jumps in variance.

6 Conclusion

In this paper, we study the index option returns, and generate a number of new results and insights. We propose a new methodology to evaluate the significance of option returns. We argue that comparing observed option returns to those generated by standard models is a reasonable exercise. We do this by showing how to compute analytical EORs and using simulations to construct finite sample distributions.

	Moneyiness		0.94	0.96	0.98	1.00	Strdl
	Data, %		-56.8	-52.3	-44.7	-29.9	-15.7
Average Returns	RiskPrem	$E^{\text{P}},\%$	-57.2	-49.9	-39.0	-24.9	-11.0
		p -value,%	64.4	40.8	24.6	29.8	8.3
	EstRisk	$E^{\text{P}},\%$	-49.0	-42.4	-34.1	-23.8	-9.5
		p -value,%	46.9	32.2	20.0	29.8	5.3
	Peso	$E^{\text{P}},\%$	-53.8	-48.5	-39.6	-26.9	-13.4
		p -value,%	59.5	42.5	25.3	35.7	14.9
	Data, %		-48.3	-44.1	-36.8	-22.5	-15.4
CAPM α s	RiskPrem	$E^{\text{P}},\%$	-55.7	-48.3	-37.5	-23.6	-10.5
		p -value,%	84.3	69.5	47.2	64.0	13.9
	EstRisk	$E^{\text{P}},\%$	-47.2	-40.7	-32.5	-22.4	-9.2
		p -value,%	65.8	51.6	33.3	56.5	9.5
	Peso	$E^{\text{P}},\%$	-52.1	-46.9	-38.1	-25.6	-12.4
		p -value,%	78.8	67.8	48.3	75.2	20.2
	Data, %		-27.3	-29.0	-29.0	-23.4	-24.3
Sharpe Ratios	RiskPrem	$E^{\text{P}},\%$	-24.5	-24.1	-22.2	-18.2	-11.7
		p -value,%	60.4	46.8	32.7	37.0	8.5
	EstRisk	$E^{\text{P}},\%$	-16.5	-16.8	-17.5	-17.1	-10.0
		p -value,%	42.2	32.8	26.6	36.0	4.8
	Peso	$E^{\text{P}},\%$	-20.7	-22.4	-22.7	-20.2	-14.1
		p -value,%	53.4	43.8	34.9	43.6	14.0

Table 7: Finite sample distribution of options returns and risk adjustments. We report population values of expected options returns, CAPM α , and Sharpe ratios and the respective model-based p -values corresponding to these quantities observed in the data. We consider three risk premia in the SVJ model: RiskPrem refers to risk premia computed based on a general equilibrium model; EstRisk refers to estimation risk-based explanation; Peso refers to the Peso problem.

We document a number of surprising findings in the context of the Black-Scholes, Merton, and Heston models without priced jump risk or stochastic volatility risk. One of the biggest current puzzles, the very low returns to deep OTM options is, in fact, not inconsistent with the Black-Scholes or Heston models. We also document in Merton's model that a high slope of the implied volatility curve does not imply high absolute option returns, and could even generate less negative expected returns than the Black-Scholes model, if jump risk is not priced. Standard risk corrections such as CAPM alphas are strongly biased, even in the Black-Scholes model. We investigate explanations such as estimation risk, factor risk premia, and Peso problems, and find that these explanations are capable of matching the average returns of put options and straddles.

We conclude that there does not appear to be anything puzzling about put option returns. This is in strong contrast to the existing literature, and our finding is due to our new approach for evaluating the significance of option returns. The only potentially puzzling statistic was ATM straddle returns, but even these were not significant when accounting for jump risk premia, estimation risk and Peso problems.

We conclude by noting that our results are largely silent on the actual sources of the gaps between the \mathbb{P} and \mathbb{Q} measures. It would be interesting to test alternative potential explanations using models incorporating investor heterogeneity, discrete trading, model misspecification, or learning. For example, Garleanu, Pedersen, and Poteshman (2005) provide a theoretical model incorporating both investor heterogeneity and discrete trading. It would be interesting to study formal parameterizations of this model to see if it can quantitatively explain the observed straddle returns. We leave these for future research.

A Details of the options dataset

In this appendix, we provide a discussion of major steps taken to construct our options dataset.

There are two ways to construct a dataset of option prices for multiple strikes: using close prices or by sampling options over a window of time. Due to microstructure concerns with close prices, we followed the latter approach. For each trading day, we select put and call transactions that could be matched within one minute to a futures transaction, typically producing hundreds of matched options-futures transactions. With these matched pairs, we compute Black-Scholes implied volatilities using a binomial tree to account for the early exercise feature of futures options. Broadie, Chernov, and Johannes (2007) show that this produces accurate early exercise adjustments in models with stochastic volatility and jumps in prices.

To reduce the dimension of our dataset and to compute implied volatilities for specific strikes, we fit a piecewise quadratic function to the implied volatilities. This allows us to combine an entire days worth of information and compute implied volatilities for exact moneyness levels. Figure 4 shows a representative day, and Broadie, Chernov, and Johannes (2007) discuss the accuracy of the method. For each month, we select the day that is exactly one month to maturity (28 or 35 calendar days) and compute implied volatilities and option prices for fixed moneyness (in increments of 0.02), measured by strike divided by the underlying.

B Delta-hedging

In this appendix, we discuss delta-hedged returns and, more generally, returns on strategies with data- or model-based portfolio weights. Delta-hedging raises a number of issues that, in our view, make interpretation of the delta-hedged returns unclear. The main issue is that the deltas can be computed in multiple ways.

We see three ways how delta-hedging can be implemented. The first uses a formal option pricing model to computing the required hedging portfolio weights. Examples of this would include using the formal stochastic volatility models with jumps discussed earlier. The second uses a data-based approach that computes the hedge ratios using the shape of the current implied volatility smile (see Bates (2005)). The third, commonly used by prac-

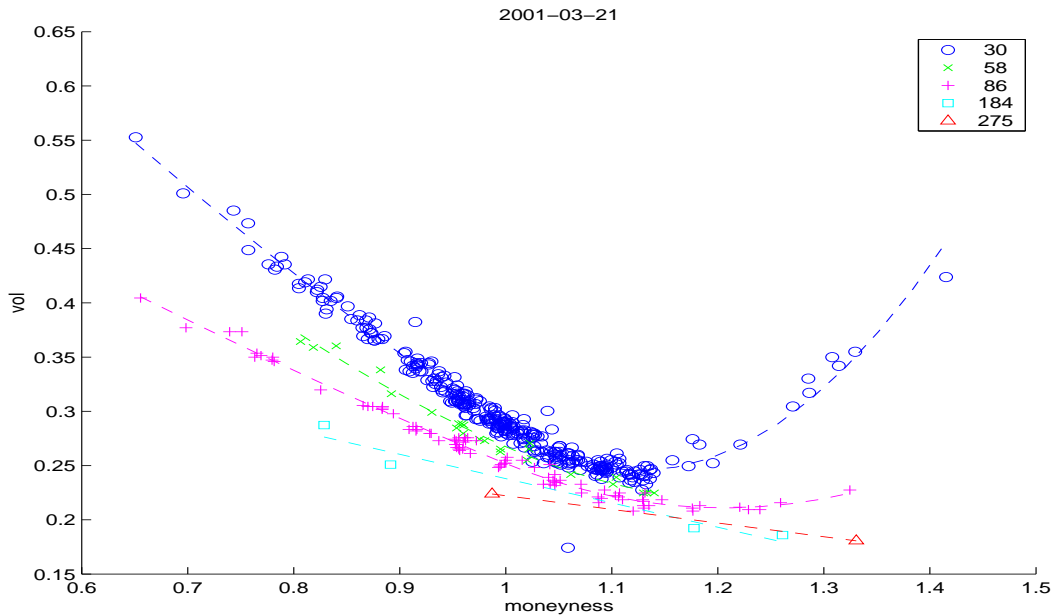


Figure 4: This figure shows representative implied volatility smiles that we construct. Circles represent the actual transactions. The solid line is the interpolated smile.

titioners, computes deltas from the Black-Scholes model and substitutes implied volatility as a constant volatility parameter. We discuss each of the approaches in turn.

Model-based hedging requires the knowledge of the spot variance, V_t , and parameters. On the one hand, to compute delta-hedged returns using the real data we have to estimate V_t in sample. To do this, we require a formal model, which leads to a joint hypothesis issue and introduces estimation noise. Moreover, estimates of spot volatilities and delta-hedged returns are highly sensitive to the model specification and, in particular, to the importance of jumps in prices.²² On the other hand, we have to compute deltas for our simulations. Here we observe V_t , but this V_t may be different from those estimated using the real data. Because of this, we will be comparing two portfolios (one in the data and one in simulations) with different portfolio weights.

As an alternative, Bates (2005) proposes an elegant model-free technique to establish delta-hedged weights. This approach circumvents the issues mentioned in the previous

²²For example, Branger and Schlag (2004) show that delta-hedged errors are not zero if the incorrect model is used or if rebalancing is discrete.

paragraph. However, the approach assumes that the markets price options correctly.²³ This concern is particularly relevant in the context of our paper, because we attempt to evaluate whether options are priced correctly.

The most practical approach is to use the Black-Scholes deltas evaluated at implied volatility. Still, the deltas computed from a model will be different from deltas computed from the data. As an extreme example, consider the Black-Scholes model where model-based implied volatility is constant across strikes and over time. In contrast, because of the well-known smile effect in the data, the 6% OTM delta in the Black-Scholes model will be evaluated at an implied volatility that is different from the one used for the ATM delta. They will also vary through time. Therefore, we will again be comparing portfolios with different weights.

Finally, delta-hedging requires rebalancing, which increases transaction costs and data requirements.²⁴ Thus, while less attractive from the theoretical perspective, the more practical static delta-hedging strategy should be evaluated. According to this strategy, a delta-hedged position is formed a month prior to an option's maturity and is not changed through the duration of the option contract.

We have evaluated a static delta-hedged strategy with the Black-Scholes deltas as described above. For ATM options, the variations in deltas across models and specifications are modest, although the variations for OTM options can be quite large (depending on the implied volatility smile). We find that the same models that can replicate returns on at-the-money straddles, can also replicate returns on delta-hedged portfolios. This should not be a surprise because ATM straddles are approximately delta-neutral. We believe that interpretation of the ATM straddle results are more clear precisely because the portfolio involves model- and data-independent weights.

²³Bates (2005) notes: "...while the proposed methodology may be able to infer the deltas ... perceived by the market, that does not mean the market is correct. If options are mispriced, it is probable that the implicit deltas ... are also erroneous."

²⁴Bollen and Whaley (2004) is the only paper that considers rebalancing. Because of data demands, they take a shortcut and use the volatility at the time the option position was opened and hold this constant until expiration.

C Instantaneous expected excess option returns

The pricing differential equation for a derivative price $f(S_t, V_t)$ in the SVJ model is

$$\begin{aligned} & \frac{\partial f}{\partial t} + \frac{\partial f}{\partial S_t} (r - \delta - \lambda^{\mathbb{Q}} \bar{\mu}^{\mathbb{Q}}) S_t + \frac{\partial f}{\partial V_t} \kappa (\theta_v^{\mathbb{Q}} - V_t) \\ & + \frac{1}{2} \frac{\partial^2 f}{\partial S_t^2} V_t S_t^2 + \frac{\partial^2 f}{\partial S_t \partial V_t} \rho \sigma_v V_t S_t + \frac{1}{2} \frac{\partial^2 f}{\partial V_t^2} \sigma_v^2 V_t \\ & + \lambda^{\mathbb{Q}} E_t^{\mathbb{Q}} [f(S_t e^Z, V_t) - f(S_t, V_t)] = r f, \end{aligned} \quad (13)$$

where Z is the jump size and the usual boundary conditions are determined by the type of derivative (e.g., Bates (1996)). We denote the change in the derivative's prices at a jump time, τ_j , as

$$\Delta f_{\tau_j} = f(S_{\tau_j-} e^{Z_j}, V_t) - f(S_{\tau_j-}, V_t)$$

and $F_t = \sum_{j=1}^{N_t} \Delta f_{\tau_j}$.

By Itô's lemma, the dynamics of derivative's price under the measure \mathbb{P} are given by

$$\begin{aligned} df = & \left[\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S_t^2} V_t S_t^2 + \frac{\partial^2 f}{\partial S_t \partial V_t} \rho \sigma_v V_t S_t + \frac{1}{2} \frac{\partial^2 f}{\partial V_t^2} \sigma_v^2 V_t \right] dt \\ & + \frac{\partial f}{\partial S_t} dS_t^c + \frac{\partial f}{\partial V_t} dV_t + d \left(\sum_{j=1}^{N_t} \Delta f_{\tau_j} \right), \end{aligned} \quad (14)$$

where S_t^c is the continuous portion of the index process:

$$\begin{aligned} dS_t^c = & (r + \mu - \delta) S_t dt + S_t \sqrt{V_t} dW_t^s - \lambda^{\mathbb{P}} \bar{\mu}^{\mathbb{P}} S_t dt \\ = & (r + \mu^c - \delta - \lambda^{\mathbb{Q}} \bar{\mu}^{\mathbb{Q}}) S_t dt + S_t \sqrt{V_t} dW_t^s. \end{aligned} \quad (15)$$

Substituting the pricing PDE into the drift, we see that

$$\begin{aligned} df = & \left[-\frac{\partial f}{\partial S_t} (r - \delta - \lambda^{\mathbb{Q}} \bar{\mu}^{\mathbb{Q}}) S_t - \frac{\partial f}{\partial V_t} \kappa (\theta_v^{\mathbb{Q}} - V_t) - \lambda^{\mathbb{Q}} E_t^{\mathbb{Q}} [f(S_t e^Z, V_t) - f(S_t, V_t)] + r f \right] dt \\ & + \frac{\partial f}{\partial S_t} dS_t^c + \frac{\partial f}{\partial V_t} dV_t + dF_t \\ = & [r f - \lambda^{\mathbb{Q}} E_t^{\mathbb{Q}} [f(S_t e^Z, V_t) - f(S_t, V_t)]] dt \\ & + \frac{\partial f}{\partial S_t} [dS_t^c - (r - \delta - \lambda^{\mathbb{Q}} \bar{\mu}^{\mathbb{Q}}) S_t dt] + \frac{\partial f}{\partial V_t} [dV_t - \kappa (\theta_v^{\mathbb{Q}} - V_t)] + dF_t. \end{aligned} \quad (16)$$

From this expression, we can compute instantaneous EORs. Taking objective measure expectations ,

$$\begin{aligned} \frac{1}{dt} E_t^{\mathbb{P}} [df] &= rf + \frac{\partial f}{\partial S_t} \mu^c S_t + \frac{\partial f}{\partial V_t} \kappa (\theta_v^{\mathbb{P}} - \theta_v^{\mathbb{Q}}) \\ &+ \{ \lambda^{\mathbb{P}} E_t^{\mathbb{P}} [f(S_t e^Z, V_t) - f(S_t, V_t)] - \lambda^{\mathbb{Q}} E_t^{\mathbb{Q}} [f(S_t e^Z, V_t) - f(S_t, V_t)] \}. \end{aligned} \quad (17)$$

Rearranging, instantaneous excess option returns are given by

$$\begin{aligned} \frac{1}{dt} E_t^{\mathbb{P}} \left[\frac{df(S_t, V_t)}{f(S_t, V_t)} - r dt \right] &= \frac{\partial \log [f(S_t, V_t)]}{\partial \log S_t} \mu^c + \frac{1}{f(S_t, V_t)} \frac{\partial f(S_t, V_t)}{\partial V_t} \kappa (\theta_v^{\mathbb{P}} - \theta_v^{\mathbb{Q}}) \\ &+ \frac{\lambda^{\mathbb{P}} E_t^{\mathbb{P}} [f(S_t e^Z, V_t) - f(S_t, V_t)] - \lambda^{\mathbb{Q}} E_t^{\mathbb{Q}} [f(S_t e^Z, V_t) - f(S_t, V_t)]}{f(S_t, V_t)}. \end{aligned} \quad (18)$$

References

- Aït-Sahalia, Yacine, Yubo Wang, and Francis Yared, 2001, Do option markets correctly price the probabilities of movement of the underlying asset?, *Journal of Econometrics* 102, 67–110.
- Andersen, Torben, Luca Benzoni, and Jesper Lund, 2002, Towards an empirical foundation for continuous-time equity return models, *Journal of Finance* 57, 1239–1284.
- Bakshi, Gurdip, and Dilip Madan, 2006, A theory of volatility spreads, *Management Science* 52, 1945–1956.
- Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.
- Bates, David, 1988, Pricing options under jump-diffusion processes, Working paper, University of Pennsylvania.
- , 1996, Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options, *Review of Financial Studies* 9, 69–107.
- , 2000, Post-'87 crash fears in S&P 500 futures options, *Journal of Econometrics* 94, 181–238.
- , 2005, Hedging the smirk, *Finance Research Letters* 2, 195–200.
- , 2006, The market for crash risk, working paper, University of Iowa.
- Benzoni, Luca, Pierre Collin-Dufresne, and Robert Goldstein, 2005, Can standard preferences explain the prices of out-of-the-money S&P 500 put options, working paper, University of Minnesota.
- Bollen, Nicholas, and Robert Whaley, 2004, Does net buying pressure affect the shape of implied volatility functions?, *Journal of Finance* 59, 711–753.
- Bondarenko, Oleg, 2003, Why are put options so expensive?, Working paper, University of Illinois, Chicago.
- Branger, Nicole, and Christian Schlag, 2004, Can tests based on option hedging errors correctly identify volatility risk premia?, Working paper, Goethe University.

- Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model specification and risk premia: Evidence from futures options, *Journal of Finance*.
- Buraschi, Andrea, and Alexei Jiltsov, 2006, Model uncertainty and option markets with heterogeneous agents, *Journal of Finance* 61.
- Cheredito, Patrick, Damir Filipovic, and Robert Kimmel, 2003, Market price of risk specifications for affine models: Theory and evidence, Working Paper, Princeton University.
- Chernov, Mikhail, 2007, On the role of risk premia in volatility forecasting, *Journal of Business and Economic Statistics*.
- , A. Ronald Gallant, Eric Ghysels, and George Tauchen, 2003, Alternative models for stock price dynamics, *Journal of Econometrics* 116, 225–257.
- David, Alexander, and Pietro Veronesi, 2006, Inflation and earnings uncertainty and volatility forecasts, working paper, University of Chicago.
- Driessen, Joost, and Pascal Maenhout, 2004a, A portfolio perspective on option pricing anomalies, Working paper, Insead and University of Amsterdam.
- , 2004b, The world price of jump and volatility risk, Working paper, Insead and University of Amsterdam.
- Eraker, Bjørn, 2004, Do equity prices and volatility jump? Reconciling evidence from spot and option prices, *Journal of Finance* 59, 1367–1403.
- , Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in equity index volatility and returns, *Journal of Finance* 58, 1269–1300.
- Garleanu, Nicolae, Lasse Pedersen, and Allen Poteshman, 2005, Demand-based option pricing, Working paper, University of Pennsylvania.
- Gastineau, G.L., and A. Madansky, 1979, Why simulations are an unreliable test of option strategies, *Financial Analysts Journal* 35, 61–76.
- Heston, Steven, 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327–343.

- Jackwerth, Jens, 2000, Recovering risk aversion from option prices and realized returns, *Review of Financial Studies* 13, 433–451.
- Johannes, Michael, Nicholas Polson, and Jonathan Stroud, 2005, Sequential parameter estimation in stochastic volatility models with jumps, Working paper, Columbia GSB.
- Jones, Christopher, 2006, A nonlinear factor analysis of S&P 500 index option returns, *Journal of Finance* 41, 2325–2363.
- Merton, Robert C., Myron S. Scholes, and Mathew L. Gladstein, 1978, The returns and risk of alternative call option portfolio investment strategies, *Journal of Business* 51, 183–242.
- , 1982, The returns and risks of alternative put-option portfolio investment strategies, *Journal of Business* 55, 1–55.
- Naik, Vasanttilak, and Moon Lee, 1990, General equilibrium pricing of options on the market portfolio with discontinuous returns, *Review of Financial Studies* 3, 493–521.
- Pan, Jun, 2002, The jump-risk premia implicit in options: Evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.
- Pastor, Lubos, and Robert F. Stambaugh, 2001, The equity premium and structural breaks, *Journal of Finance* 56, 1207–1239.
- Rubinstein, Mark, 1984, A simple formula for the expected rate of return of an option over a finite holding period, *Journal of Finance* 39, 1503–1509.
- Santa-Clara, Pedro, and Alessio Saretto, 2005, Option strategies: Good deals and margin calls, working paper, UCLA.
- Scott, Louis O., 1997, Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Applications of fourier inversion methods, *Mathematical Finance* 7, 413–426.
- Sheikh, Aamir M., and Ehud I. Ronn, 1994, A characterization of the daily and intraday behaviour of returns on options, *Journal of Finance* 49, 557–579.

Zellner, Arnold, 1971, *An Introduction to Bayesian Inference in Econometrics* (John Wiley & Sons: New York).