

Demand-Based Option Pricing*

Nicolae Gârleanu[†]
Lasse Heje Pedersen[‡]
Allen M. Poteshman[§]

This version: August, 2006

Abstract

We model demand-pressure effects on option prices. The model shows that demand pressure in one option contract increases its price by an amount proportional to the variance of the unhedgeable part of the option. Similarly, the demand pressure increases the price of any other option by an amount proportional to the covariance of the unhedgeable parts of the two options. Empirically, we identify aggregate positions of dealers and end users using a unique dataset, and show that demand-pressure effects contribute to well-known option-pricing puzzles. Indeed, time-series tests show that demand helps explain the overall expensiveness and skew patterns of both index options and single-stock options.

* We are grateful for helpful comments from David Bates, Nick Bollen, Oleg Bondarenko, Menachem Brenner, Andrea Buraschi, Josh Coval, Domenico Cuoco, Apoorva Koticha, Sophie Ni, Jun Pan, Neil Pearson, Mike Weisbach, Josh White, and especially from Steve Figlewski, as well as from seminar participants at the Caesarea Center Conference, Columbia University, Cornell University, Dartmouth University, Harvard University, HEC Lausanne, the 2005 Inquire Europe Conference, London Business School, the 2005 NBER Behavioral Finance Conference, the 2005 NBER Universities Research Conference, the NYU-ISE Conference on the Transformation of Option Trading, New York University, Oxford University, Stockholm Institute of Financial Research, Texas A&M University, the University of Chicago, University of Illinois at Urbana-Champaign, the University of Pennsylvania, University of Vienna, and the 2005 WFA Meetings.

[†]Wharton School of Business, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, email: garleanu@wharton.upenn.edu, URL: <http://finance.wharton.upenn.edu/~garleanu/>.

[‡]New York University, Centre of Economic Policy Research (CEPR), and National Bureau of Economic Research (NBER), 44 West Fourth Street, Suite 9-190, New York, NY 10012-1126, email: lpederse@stern.nyu.edu, URL: <http://pages.stern.nyu.edu/~lpederse/>.

[§]University of Illinois at Urbana-Champaign, 340 Wohlers Hall, 1206 South Sixth Street, Champaign, Illinois 61820, email: poteshma@uiuc.edu, URL: <http://www.business.uiuc.edu/poteshma>.

One of the major achievements of financial economics is the no-arbitrage theory that determines derivative prices independently of investor demand. Building on the seminal contributions of Black and Scholes (1973) and Merton (1973), a large literature develops various parametric implementations of the theory. This literature is surveyed by Bates (2003) who emphasizes that it cannot fully capture — much less explain — the empirical properties of option prices and concludes that there is a need for a new approach to pricing derivatives. He writes:

“To blithely attribute divergences between objective and risk-neutral probability measures to the free ‘risk premium’ parameters within an affine model is to abdicate one’s responsibilities as a financial economist. ... a renewed focus on the explicit financial intermediation of the underlying risks by option market makers is needed.”

We take on this challenge. Our model departs fundamentally from the no-arbitrage framework by recognizing that option market makers cannot perfectly hedge their inventories, and, consequently, option demand impacts option prices. We obtain explicit expressions for the effects of demand on option prices, provide empirical evidence consistent with the demand-pressure model using a unique dataset, and show that demand-pressure effects can help to resolve the main option-pricing puzzles.

The starting point of our analysis is that options are traded because they are useful and, therefore, options cannot be redundant for all investors (Hakansson (1979)). We denote the agents who have a fundamental need for option exposure as “end users.”

Intermediaries such as market makers and proprietary traders provide liquidity to end users by taking the other side of the end-user net demand. If competitive intermediaries can hedge perfectly — as in a Black-Scholes-Merton economy — then option prices are determined by no-arbitrage and demand pressure has no effect. In reality, however, even intermediaries cannot hedge options perfectly — that is, even they face incomplete markets — because of the impossibility of trading continuously, stochastic volatility, jumps in the underlying, and transaction costs (Figlewski (1989)).¹ In addition, intermediaries are sensitive to risk, e.g., because of capital constraints and agency problems (Shleifer and Vishny (1997)).

In light of these facts, we consider how options are priced by competitive risk-averse dealers who cannot hedge perfectly. In our model, dealers trade an arbitrary number of option contracts on the same underlying at discrete times. Since the dealers trade many option contracts, certain risks net out, while others do not. The dealers can hedge part of the remaining risk of their derivative positions by trading the underlying security and risk-free bonds. We consider a general class of distributions for the underlying, which can accommodate stochastic volatility and jumps. Dealers trade options with

¹Options may also be impossible to replicate due to asymmetric information (Back (1993) and Easley, O’Hara, and Srinivas (1998).)

end users. The model is agnostic about the end users' reasons for trade, which are irrelevant for our results and their empirical implementation.

We compute equilibrium prices as functions of demand pressure, that is, the prices that induce the utility-maximizing dealers to supply precisely the option quantities that the end users demand. We show explicitly how demand pressure enters into the pricing kernel. Intuitively, a positive demand pressure in an option increases the pricing kernel in the states of nature in which an optimally hedged position has a positive payoff. This pricing-kernel effect increases the price of the option, which entices the dealers to sell it. Specifically, a marginal change in the demand pressure in an option contract increases its price by an amount proportional to the variance of the unhedgeable part of the option, where the variance is computed under a certain probability measure depending on the demand. Similarly, demand pressure increases the price of any other option by an amount proportional to the covariance of their unhedgeable parts. Hence, while demand pressure in a particular option raises its price, it also raises the prices of other options on the same underlying.

Our model applies both when the underlying asset and the risk-free bond constitute a dynamically complete market and when these two securities leave the market incomplete. In the former case, demand pressure does not impact option prices, because the variance of the unhedgeable price change of any option and the covariance of the unhedgeable price changes of any two options are zero. Consequently, the interesting case is the one in which the underlying asset and the risk-free bond leave the market dynamically incomplete.² Our main theoretical results relating option price effects to the variance or covariance of the unhedgeable part of option price changes hold regardless of the source of market incompleteness. The magnitudes of the variances and covariances, and hence of the demand-based option price effects, depend upon the particular source of market incompleteness. Empirically, we test the specific predictions of the model under the assumptions that market incompleteness stems from discrete trading, stochastic volatility, or jumps.

Empirically, we use a unique dataset to identify aggregate daily positions of dealers and end users. In particular, we define dealers as market makers and end users as proprietary traders and customers of brokers.³ We are the first to document that end users have a net long position in S&P500 index options with large net positions in out-of-the-money puts. Since options are in zero net supply, this implies that dealers are short index options.⁴ We estimate that these large short dealer positions lead to

²It is not important whether options complete the market.

³The empirical results are robust to classifying proprietary traders as either dealers or end users.

⁴This fact and its relevance for pricing appear to be recognized by option traders. For instance, Vanessa Gray, director of global equity derivatives, Dresdner Kleinwort Benson, states that option implied volatility skew “is heavily influenced by supply and demand factors,” and Amine Belhadj-Soulami, head of equity derivatives trading for Europe, Paribas, remarks that the “number of players in the skew market is limited. [...] there’s a huge imbalance between what clients want and what professionals can provide.”

daily delta-hedged profits and losses varying between \$100 million and -\$100 million, and cumulative dealer profits of approximately \$800 million over our 6 year sample. Hence, consistent with our framework, dealers face significant unhedgeable risk and are compensated for bearing it.

The end-user demand for index options can help explain the two puzzles that index options appear to be expensive, and that low-moneyness options seem to be especially expensive (Longstaff (1995), Bates (2000), Coval and Shumway (2001), Bondarenko (2003), Amin, Coval, and Seyhun (2004)). In the time series, the model-based impact of demand for index options is positively related to their expensiveness, measured by the difference between their implied volatility and the volatility measure of Bates (2005). This link between demand and prices is stronger following recent dealer losses, as would be expected if dealers are more risk averse at such times. Likewise, the steepness of the smirk, measured by the difference between the implied volatilities of low-moneyness options and at-the-money options, is positively related to the skew of option demand.

Jackwerth (2000) finds that a representative investor's option-implied utility function is inconsistent with standard assumptions in economic theory.⁵ Since options are in zero net supply, a representative investor holds no options. We reconcile this finding for dealers who have significant short index option positions. Intuitively, a dealer will short index options, but only a finite number of options. Hence, while a standard-utility investor may not be marginal on options given a zero position, he is marginal given a certain negative position. We do not address why end users buy these options; their motives might be related to portfolio insurance and agency problems (e.g., between investors and fund managers) that are not well captured by standard utility theory.⁶

Another option-pricing puzzle is the significant difference between index-option prices and the prices of single-stock options, despite the relative similarity of the underlying distributions (e.g., Bakshi, Kapadia, and Madan (2003) and Bollen and Whaley (2004)). In particular, single-stock options appear cheaper and their smile is flatter. Consistently, we find that the demand pattern for single-stock options is very different from that of index options. For instance, end users are net short single-stock options — not long, as in the case of index options.

Demand patterns further help explain the time-series and cross-sectional pricing of single-stock options. Indeed, individual stock options are cheaper at times when end users sell more options, and, in the cross section, stocks with more negative demand for options, aggregated across contracts, tend to have relatively cheaper options.

The paper is related to several strands of literature. First, the literature on option pricing in the context of trading frictions and incomplete markets derives bounds on option prices. Arbitrage bounds are trivial with any transaction costs; for instance, the price of a call option can be as high as the price of the underlying stock (Soner, Shreve, and Cvitanic (1995)). This serious limitation of no-arbitrage pricing has led

⁵See also Driessen and Maenhout (2003).

⁶Leland (1980) characterizes circumstances in which a given investor may wish to buy index options.

Bernardo and Ledoit (2000) and Cochrane and Saa-Requejo (2000) to derive tighter option-pricing bounds by restricting the Sharpe ratio or gain/loss ratio to be below an arbitrary level, and stochastic dominance bounds for small option positions are derived by Constantinides and Perrakis (2002) and extended and implemented empirically by Constantinides, Jackwerth, and Perrakis (2005). Rather than deriving bounds, we compute explicit prices based on the demand pressure by end users. We further complement this literature by taking portfolio considerations into account, that is, the effect of demand for one option on the prices of other options.

Second, the literature on utility-based option pricing (“indifference pricing”) derives the option price that would make an agent (e.g., the representative agent) indifferent between buying the option and not buying it. See Rubinstein (1976), Brennan (1979), Stapleton and Subrahmanyam (1984), Hugonnier, Kramkov, and Schachermayer (2005), and references therein. While this literature computes the price of the first “marginal” option demanded, we show how option prices change when demand is non-trivial.

Third, Stein (1989) and Poteshman (2001) provide evidence that option investors misproject changes in the instantaneous volatility of underlying assets by examining the price changes of shorter and longer maturity options. Our paper shows how cognitive biases of option end users can translate (via their option demands) into option prices even if market makers are not subject to any behavioral biases. By contrast, under standard models like Black-Scholes-Merton, market makers who can hedge their positions perfectly will correct the mistakes of other option market participants before they affect option prices.

Fourth, the general idea of demand pressure effects goes back, at least, to Keynes (1923) and Hicks (1939) who considered futures markets. Our model is the first to apply this idea to option pricing and to incorporate the important features of option markets, namely dynamic trading of many assets, hedging using the underlying and bonds, stochastic volatility, and jumps. The generality of our model also makes it applicable to other markets. Consistent with our model’s predictions, Wurgler and Zhuravskaya (2002) extend Shleifer (1986) and find that stocks that are hard to hedge experience larger price jumps when included into the S&P 500 index. Greenwood (2005) considers a major redefinition of the Nikkei 225 index in Japan and finds that stocks that are not affected by demand shocks, but that are correlated with securities facing demand shocks, experience price changes. Similarly in the fixed income market, Newman and Rierson (2004) find that non-informative issues of telecom bonds depress the price of the issued bond as well as correlated telecom bonds, and Gabaix, Krishnamurthy, and Vigneron (2004) find related evidence for mortgage-backed securities. Further, de Roon, Nijman, and Veld (2000) find futures-market evidence consistent with the model’s predictions.

The most closely related paper is Bollen and Whaley (2004), which demonstrates that changes in implied volatility are correlated with signed option volume. These

empirical results set the stage for our analysis by showing that *changes* in option demand lead to *changes* in option prices while leaving open the question of whether the *level* of option demand impacts the overall *level* (i.e., expensiveness) of option prices or the overall shape of implied-volatility curves.⁷ We complement Bollen and Whaley (2004) by providing a theoretical model, by investigating empirically the relationship between the level of end user demand for options and the level and overall shape of implied volatility curves, and by testing precise quantitative implications of our model. In particular, we document that end users tend to have a net long SPX option position and a short equity-option position, thus helping to explain the relative expensiveness of index options. We also show that there is a strong downward skew in the net demand of index but not equity options which helps to explain the difference in the shapes of their overall implied volatility curves. In addition, we demonstrate that option prices are better explained by model-based rather than simple non-model based use of demand.

1 A Model of Demand Pressure

We consider a discrete-time infinite-horizon economy. There exists a risk-free asset paying interest at the rate of $R_f - 1$ per period (i.e., if the risk-free rate is 2% per period, then $R_f = 1.02$), and a risky security that we refer to as the “underlying” security. At time t , the underlying has an exogenous strictly positive price⁸ of S_t , dividend D_t , and an excess return of $R_t^e = (S_t + D_t)/S_{t-1} - R_f$. The distribution of future prices and returns is characterized by a stationary Markov state variable X_t , which can include the current level of volatility, the current jump intensity, etc. We assume that that (R_t^e, X_t) satisfies a Feller-type condition (made precise in the appendix) and that S_t and X_t are bounded for every t .

The economy further has a number of “derivative” securities, whose prices are to be determined endogenously. A derivative security is characterized by its index $i \in I$, where i collects the information that identifies the derivative and its payoffs. For a European option, for instance, the strike price, maturity date, and whether the option is a “call” or “put” suffice. The set of derivatives traded at time t is denoted by I_t , and the vector of prices of traded securities is $p_t = (p_t^i)_{i \in I_t}$.

The payoffs of the derivatives depend on S_t and X_t . We note that the theory is completely general and does not require that the “derivatives” have payoffs that depend on the underlying. In principle, the derivatives could be any securities — in particular, any securities whose prices are affected by demand pressure. Further, it

⁷Indeed, Bollen and Whaley (2004) find that a nontrivial part of the option price impact from day t signed option volume dissipates by day $t + 1$.

⁸All random variables are defined on a probability space $(\Omega, \mathcal{F}, Pr)$ with an associated filtration $\{\mathcal{F}_t : t \geq 0\}$ of sub- σ -algebras representing the resolution over time of information commonly available to agents.

is straightforward to extend our results to a model with any number of exogenously-priced securities. While we use the model to study options in particular, we think that the generality helps to illuminate the driving forces behind the results and, further, it allows future applications of the theory in other markets.

The economy is populated by two kinds of agents: “dealers” and “end users.” Dealers are competitive and there exists a representative dealer who has constant absolute risk aversion, that is, his utility for remaining life-time consumption is

$$U(C_t, C_{t+1}, \dots) = \mathbb{E}_t \left[\sum_{v=t}^{\infty} \rho^{v-t} u(C_v) \right], \quad (1)$$

where $u(c) = -\frac{1}{\gamma}e^{-\gamma c}$ and $\rho < 1$ is a discount factor. At any time t , the dealer must choose the consumption C_t , the dollar investment in the underlying θ_t , and the number of derivatives held $q_t = (q_t^i)_{i \in I_t}$, so as to maximize its utility while satisfying the transversality condition $\lim_{t \rightarrow \infty} \mathbb{E} [\rho^{-t} e^{-k W_t}] = 0$. The dealer’s wealth evolves as

$$W_{t+1} = (W_t - C_t)R_f + q_t(p_{t+1} - R_f p_t) + \theta_t R_{t+1}^e. \quad (2)$$

In the real world, end users trade options for a variety of reasons such as portfolio insurance, agency reasons, behavioral reasons, institutional reasons etc. Rather than trying to capture these various trading motives endogenously, we assume that end users have an exogenous aggregate demand for derivatives of $d_t = (d_t^i)_{i \in I_t}$ at time t . The distribution of future demand is characterized by X_t . We also assume, for technical reasons, that demand pressure is zero after some time \bar{T} , that is, $d_t = 0$ for $t > \bar{T}$.

Derivative prices are set through the interaction between dealers and end users in a competitive equilibrium.

Definition 1 *A price process $p_t = p_t(d_t, X_t)$ is a (competitive Markov) equilibrium if, given p , the representative dealer optimally chooses a derivative holding q such that derivative markets clear, i.e., $q + d = 0$.*

Our asset-pricing approach relies on the insight that, by observing the aggregate quantities held by dealers, one can determine the derivative prices consistent with the dealers’ utility maximization. Our goal is to determine how derivative prices depend on the demand pressure d coming from end users. We note that the assumption of inelastic end-user demand has no implications for the results we derive below. All that matters is that end users have demand curves that result in dealers choosing to hold, at the market prices, a position of $q = -d$.

To determine the representative dealer’s optimal behavior, we consider his value function $J(W; t, X)$, which depends on his wealth W , the state of nature X , and time t . Then, the dealer solves the following maximization problem:

$$\max_{C_t, q_t, \theta_t} -\frac{1}{\gamma}e^{-\gamma C_t} + \rho \mathbb{E}_t [J(W_{t+1}; t+1, X_{t+1})] \quad (3)$$

$$\text{s.t.} \quad W_{t+1} = (W_t - C_t)R_f + q_t(p_{t+1} - R_f p_t) + \theta_t R_{t+1}^e. \quad (4)$$

The value function is characterized in the following proposition.

Lemma 1 *If $p_t = p_t(d_t, X_t)$ is the equilibrium price process and $k = \frac{\gamma(R_f - 1)}{R_f}$, then the dealer's value function and optimal consumption are given by*

$$J(W_t; t, X_t) = -\frac{1}{k} e^{-k(W_t + G_t(d_t, X_t))} \quad (5)$$

$$C_t = \frac{R_f - 1}{R_f} (W_t + G_t(d_t, X_t)) \quad (6)$$

and the stock and derivative holdings are characterized by the first-order conditions

$$0 = E_t \left[e^{-k(\theta_t R_{t+1}^e + q_t(p_{t+1} - R_f p_t) + G_{t+1}(d_{t+1}, X_{t+1}))} R_{t+1}^e \right] \quad (7)$$

$$0 = E_t \left[e^{-k(\theta_t R_{t+1}^e + q_t(p_{t+1} - R_f p_t) + G_{t+1}(d_{t+1}, X_{t+1}))} (p_{t+1} - R_f p_t) \right], \quad (8)$$

where, for $t \leq T$, $G_t(d_t, X_t)$ is derived recursively using (7), (8), and

$$e^{-k R_f G_t(d_t, X_t)} = R_f \rho E_t \left[e^{-k(q_t(p_{t+1} - R_f p_t) + \theta_t R_{t+1}^e + G_{t+1}(d_{t+1}, X_{t+1}))} \right] \quad (9)$$

and for $t > T$, the function $G_t(d_t, X_t) = \bar{G}(X_t)$ where $(\bar{G}(X_t), \bar{\theta}(X_t))$ solves

$$e^{-k R_f \bar{G}(X_t)} = R_f \rho E_t \left[e^{-k(\bar{\theta}_t R_{t+1}^e + \bar{G}(X_{t+1}))} \right] \quad (10)$$

$$0 = E_t \left[e^{-k(\bar{\theta}_t R_{t+1}^e + \bar{G}(X_{t+1}))} R_{t+1}^e \right]. \quad (11)$$

The optimal consumption is unique and the optimal security holdings are unique provided their payoffs are linearly independent.

While dealers compute optimal positions given prices, we are interested in inverting this mapping and compute the prices that make a given position optimal. The following proposition ensures that this inversion is possible.

Proposition 1 *Given any demand pressure process d for end users, there exists a unique equilibrium p .*

Before considering explicitly the effect of demand pressure, we make a couple of simple ‘‘parity’’ observations that show how to treat derivatives that are linearly dependent such as puts and calls with the same strike and maturity. For simplicity, we do this only in the case of a non-dividend paying underlying, but the results can easily be extended. We consider two derivatives, i and j such that a non-trivial linear combination of their payoffs lies in the span of exogenously-priced securities, i.e., the

underlying and the bond. In other words, suppose that at the common maturity date T ,

$$p_T^i = p_T^j + \alpha + \beta S_T \quad (12)$$

for some constants α and β . Then it is easily seen that, if positions $(q_t^i, q_t^j, b_t, \theta_t)$ in the two derivatives, the bond,⁹ and the underlying, respectively, are optimal given the prices, then so are positions $(q_t^j + a, q_t^j - a, b_t - a\alpha R_f^{-(T-t)}, \theta_t - a\beta S_t^{-1})$. This has the following implications for equilibrium prices:

Proposition 2 *Suppose that $D_t = 0$ and $p_T^i = p_T^j + \alpha + \beta S_T$. Then:*

(i) *For any demand pressure, d , the equilibrium prices of the two derivatives are related by*

$$p_t^i = p_t^j + \alpha R_f^{-(T-t)} + \beta S_t. \quad (13)$$

(ii) *Changing the end user demand from (d_t^i, d_t^j) to $(d_t^i + a, d_t^j - a)$, for any $a \in \mathbb{R}$, has no effect on equilibrium prices.*

The first part of the proposition is a general version of the well-known put-call parity. It shows that if payoffs are linearly dependent then so are prices.

The second part of the proposition shows that linearly dependent derivatives have the same demand-pressure effects on prices. Hence, in our empirical exercise, we can aggregate the demand of calls and puts with the same strike and maturity. That is, a demand pressure of d^i calls and d^j puts is the same as a demand pressure of $d^i + d^j$ calls and 0 puts (or vice versa).

2 Price Effects of Demand Pressure

To see where we are going with the theory (and why we need it!), consider the empirical problem that we ultimately face: On any given day, around 120 SPX option contracts of various maturities and strike prices are traded. The demands for all these different options potentially affect the price of, say, the 1-month at-the-money SPX option because all of these options expose the market makers to unhedgeable risk. What is the aggregate effect of all these demands?

The model answers this question by showing how to compute the impact of demand d_t^j for any one derivative on the price p_t^i of the 1-month at-the-money option. The aggregate effect is then the sum of all of the individual demand effects, that is, the sum of all the demands weighted by their model-implied price impacts $\partial p_t^i / \partial d_t^j$.

We first characterize $\partial p_t^i / \partial d_t^j$ in complete generality, as well as other general demand effects on prices (Section 2.1). We then show how to compute $\partial p_t^i / \partial d_t^j$ specifically when unhedgeable risk arises from, respectively, discrete-time hedging, jumps in the underlying asset price, and stochastic volatility risk (Sections 2.2–2.4).

⁹This is a dollar amount; equivalently, we may assume that the price of the bond is always 1.

2.1 General Results

We think of the price p , the hedge position θ_t in the underlying, and the consumption function G as functions of d_t^j and X_t . Alternatively, we can think of the dependent variables as functions of the dealer holding q_t^j and X_t , keeping in mind the equilibrium relation that $q = -d$. For now we use this latter notation.

At maturity date T , an option has a known price p_T . At any prior date t , the price p_t can be found recursively by “inverting” (8) to get

$$p_t = \frac{E_t \left[e^{-k(\theta_t R_{t+1}^e + q_t p_{t+1} + G_{t+1})} p_{t+1} \right]}{R_f E_t \left[e^{-k(\theta_t R_{t+1}^e + q_t p_{t+1} + G_{t+1})} \right]}, \quad (14)$$

where the hedge position in the underlying, θ_t , solves

$$0 = E_t \left[e^{-k(\theta_t R_{t+1}^e + q_t p_{t+1} + G_{t+1})} R_{t+1}^e \right] \quad (15)$$

and where G is computed recursively as described in Lemma 1. Equations (14) and (15) can be written in terms of a demand-based pricing kernel:

Theorem 1 *Prices p and the hedge position θ satisfy*

$$p_t = E_t(m_{t+1}^d p_{t+1}) = \frac{1}{R_f} E_t^d(p_{t+1}) \quad (16)$$

$$0 = E_t(m_{t+1}^d R_{t+1}^e) = \frac{1}{R_f} E_t^d(R_{t+1}^e) \quad (17)$$

where the pricing kernel m^d is a function of demand pressure d :

$$m_{t+1}^d = \frac{e^{-k(\theta_t R_{t+1}^e + q_t p_{t+1} + G_{t+1})}}{R_f E_t \left[e^{-k(\theta_t R_{t+1}^e + q_t p_{t+1} + G_{t+1})} \right]} \quad (18)$$

$$= \frac{e^{-k(\theta_t R_{t+1}^e - d_t p_{t+1} + G_{t+1})}}{R_f E_t \left[e^{-k(\theta_t R_{t+1}^e - d_t p_{t+1} + G_{t+1})} \right]}, \quad (19)$$

and E_t^d is expected value with respect to the corresponding risk-neutral measure, i.e. the measure with a Radon-Nykodim derivative of $R_f m_{t+1}^d$.

To understand this pricing kernel, suppose for instance that end users want to sell derivative i such that $d_t^i < 0$, and that this is the only demand pressure. In equilibrium, dealers take the other side of the trade, buying $q_t^i = -d_t^i > 0$ units of this derivative, while hedging their derivative holding using a position θ_t in the underlying. The pricing kernel is small whenever the “unhedgeable” part $q_t p_{t+1} + \theta_t R_{t+1}^e$ is large. Hence, the

pricing kernel assigns a low value to states of nature in which a hedged position in the derivative pays off profitably, and it assigns a high value to states in which a hedged position in the derivative has a negative payoff. This pricing kernel-effect decreases the price of this derivative, which is what entices the dealers to buy it.

It is interesting to consider the first-order effect of demand pressure on prices. Hence, we calculate explicitly the sensitivity of the prices of a derivative p_t^i with respect to the demand pressure of another derivative d_t^j . We can initially differentiate with respect to q rather than d since $q^i = -d_t^i$.

For this, we first differentiate the pricing kernel¹⁰

$$\frac{\partial m_{t+1}^d}{\partial q_t^j} = -k m_{t+1}^d \left(p_{t+1}^j - R_f p_t^j + \frac{\partial \theta_t}{\partial q_t^j} R_{t+1}^e \right) \quad (20)$$

using the facts that $\frac{\partial G(t+1, X_{t+1}; q)}{\partial q_t^j} = 0$ and $\frac{\partial p_{t+1}}{\partial q_t^j} = 0$. With this result, it is straightforward to differentiate (17) to get

$$0 = E_t \left(m_{t+1}^d \left(p_{t+1}^j - R_f p_t^j + \frac{\partial \theta_t}{\partial q_t^j} R_{t+1}^e \right) R_{t+1}^e \right), \quad (21)$$

which implies that the marginal hedge position is

$$\frac{\partial \theta_t}{\partial q_t^j} = - \frac{E_t \left(m_{t+1}^d \left(p_{t+1}^j - R_f p_t^j \right) R_{t+1}^e \right)}{E_t \left(m_{t+1}^d \left(R_{t+1}^e \right)^2 \right)} = - \frac{\text{Cov}_t^d \left(p_{t+1}^j, R_{t+1}^e \right)}{\text{Var}_t^d \left(R_{t+1}^e \right)}. \quad (22)$$

Similarly, we derive the price sensitivity by differentiating (16)

$$\begin{aligned} \frac{\partial p_t^i}{\partial q_t^j} &= -k E_t \left[m_{t+1}^d \left(p_{t+1}^j - R_f p_t^j + \frac{\partial \theta_t}{\partial q_t^j} R_{t+1}^e \right) p_{t+1}^i \right] \\ &= -\frac{k}{R_f} E_t^d \left[\left(p_{t+1}^j - R_f p_t^j - \frac{\text{Cov}_t^d \left(p_{t+1}^j, R_{t+1}^e \right)}{\text{Var}_t^d \left(R_{t+1}^e \right)} R_{t+1}^e \right) p_{t+1}^i \right] \\ &= -\gamma (R_f - 1) E_t^d \left[\bar{p}_{t+1}^j \bar{p}_{t+1}^i \right] \\ &= -\gamma (R_f - 1) \text{Cov}_t^d \left[\bar{p}_{t+1}^j, \bar{p}_{t+1}^i \right] \end{aligned} \quad (23)$$

where \bar{p}_{t+1}^i and \bar{p}_{t+1}^j are the *unhedgeable* parts of the price changes as defined in:

Definition 2 The *unhedgeable price change* \bar{p}_{t+1}^k of any security k is defined as its excess return $p_{t+1}^k - R_f p_t^k$ optimally hedged with the stock position $\frac{\text{Cov}_t^d \left(p_{t+1}^k, R_{t+1}^e \right)}{\text{Var}_t^d \left(R_{t+1}^e \right)}$:

$$\bar{p}_{t+1}^k = R_f^{-1} \left(p_{t+1}^k - R_f p_t^k - \frac{\text{Cov}_t^d \left(p_{t+1}^k, R_{t+1}^e \right)}{\text{Var}_t^d \left(R_{t+1}^e \right)} R_{t+1}^e \right). \quad (24)$$

¹⁰We suppress the arguments of functions. We note that p_t , θ_t , and G_t are functions of (d_t, X_t, t) , and m_{t+1}^d is a function of $(d_t, X_t, d_{t+1}, X_{t+1}, y_{t+1}, R_{t+1}^e, t)$.

Equation (23) can also be written in terms of the demand pressure, d , by using the equilibrium relation $d = -q$:

Theorem 2 *The sensitivity of the price of security i to demand pressure in security j is proportional to the covariance of their unhedgeable risks:*

$$\frac{\partial p_t^i}{\partial d_t^j} = \gamma(R_f - 1)E_t^d(\bar{p}_{t+1}^i \bar{p}_{t+1}^j) = \gamma(R_f - 1)\text{Cov}_t^d(\bar{p}_{t+1}^i, \bar{p}_{t+1}^j). \quad (25)$$

This result is intuitive: it says that the demand pressure in an option j increases the option's own price by an amount proportional to the variance of the unhedgeable part of the option and the aggregate risk aversion of dealers. We note that since a variance is always positive, the demand-pressure effect on the security itself is naturally always positive. Further, this demand pressure affects another option i by an amount proportional to the covariance of their unhedgeable parts. Under the condition stated below, we can show that this covariance is positive, and therefore that demand pressure in one option also increases the price of other options on the same underlying.

Proposition 3 *Demand pressure in any security j :*

- (i) *increases its own price, that is, $\frac{\partial p_t^j}{\partial d_t^j} \geq 0$;*
- (ii) *increases the price of another security i , that is, $\frac{\partial p_t^i}{\partial d_t^j} \geq 0$, provided that $E_t^d[p_{t+1}^i | S_{t+1}]$ and $E_t^d[p_{t+1}^j | S_{t+1}]$ are convex functions of S_{t+1} and $\text{Cov}_t^d(p_{t+1}^i, p_{t+1}^j | S_{t+1}) \geq 0$.*

The conditions imposed in part (ii) are natural. First, we require that prices inherit the convexity property of the option payoffs in the underlying price. Convexity lies at the heart of this result, which, informally speaking, states that higher demand for convexity (or gamma, in option-trader lingo) increases its price, and therefore those of all options. Second, we require that $\text{Cov}_t^d(p_{t+1}^i, p_{t+1}^j | S_{t+1}) \geq 0$, that is, changes in the other variables have a similar impact on both option prices — for instance, both prices are increasing in the volatility or demand level. Note that both conditions hold if both options mature after one period. The second condition also holds if option prices are homogenous (of degree 1) in (S, K) , where K is the strike, and S_t is independent of X_t .

It is interesting to consider the total price that end users pay for their demand d_t at time t . Vectorizing the derivatives from Theorem 2, we can first-order approximate the price around a zero demand as

$$p_t \approx p_t(d_t = 0) + \gamma(R_f - 1)E_t^d(\bar{p}_{t+1} \bar{p}'_{t+1}) d_t. \quad (26)$$

Hence, the total price paid for the d_t derivatives is

$$d_t' p_t = d_t' p_t(d_t = 0) + \gamma(R_f - 1)d_t' E_t^d(\bar{p}_{t+1} \bar{p}'_{t+1}) d_t \quad (27)$$

$$= d_t' p_t(d_t = 0) + \gamma(R_f - 1)\text{Var}_t^d(d_t' \bar{p}_{t+1}). \quad (28)$$

The first term $d_t' p_t(d_t = 0)$ is the price that end users would pay if their demand pressure did not affect prices. The second term is total variance of the unhedgeable part of all of the end users' positions.

While Proposition 3 shows that demand for an option increases the prices of all options, the size of the price effect is, of course, not the same for all options. Nor is the effect on implied volatilities the same. Under certain conditions, demand pressure in low-strike options has a larger impact on the implied volatility of low-strike options, and conversely for high strike options. The following proposition makes this intuitively appealing result precise. For simplicity, the proposition relies on unnecessarily restrictive assumptions. We let $p(p, K, d)$, respectively $p(c, K, d)$, denote the price of a put, respectively a call, with strike price K and one period to maturity, where d is the demand pressure. It is natural to compare low-strike and high-strike options that are "equally far out of the money." We do this by considering an out-of-the-money put with the same price as an out-of-the-money call.

Proposition 4 *Assume that the one-period risk-neutral distribution of the underlying return is symmetric and consider demand pressure $d > 0$ in an option with strike $K < R_f S_t$ that matures after one trading period. Then there exists a value \bar{K} such that, for all $K' \leq \bar{K}$ and K'' such that $p(p, K', 0) = p(c, K'', 0)$, it holds that $p(p, K', d) > p(c, K'', d)$. That is, the price of the out-of-the-money put $p(p, K', \cdot)$ is more affected by the demand pressure than the price of out-of-the-money call $p(c, K'', \cdot)$. The reverse conclusion applies if there is demand for a high-strike option.*

Future demand pressure in a derivative j also affects the current price of derivative i . As above, we consider the first-order price effect. This is slightly more complicated, however, since we cannot differentiate with respect to the unknown future demand pressure. Instead, we "scale down" the future demand pressure, that is, we consider future demand pressures $\tilde{d}_s^j = \epsilon d_s^j$ for fixed d (equivalently, $\tilde{q}_s^j = \epsilon q_s^j$) for some $\epsilon \in \mathbb{R}$, $\forall s > t$, and $\forall j$.

Theorem 3 *Let $p_t(0)$ denote the equilibrium derivative prices with 0 demand pressure. Fixing a process d with $d_t = 0$ for all $t > T$ and a given T , the equilibrium prices p with a demand pressure of ϵd is*

$$p_t = p_t(0) + \gamma(R_f - 1) \left[E_t^0 (\bar{p}_{t+1} \bar{p}'_{t+1}) d_t + \sum_{s>t} R_f^{-(s-t)} E_t^0 (\bar{p}_{s+1} \bar{p}'_{s+1} d_s) \right] \epsilon + O(\epsilon^2). \quad (29)$$

This theorem shows that the impact of current demand pressure d_t on the price of a derivative i is given by the amount of hedging risk that a marginal position in security i would add to the dealer's portfolio, that is, it is the sum of the covariances of its unhedgeable part with the unhedgeable part of all the other securities, multiplied by their respective demand pressures. Further, the impact of future demand pressures d_s

is given by the expected future hedging risks. Of course, the impact increases with the dealers' risk aversion.

Next, we discuss how demand is priced in connection with three specific sources of unhedgeable risk for the dealers: discrete-time hedging, jumps in the underlying stock, and stochastic volatility risk. We focus on small hedging periods Δ_t and derive the results informally while relegating a more rigorous treatment to the appendix. The continuously compounded risk-free interest rate is denoted by r , i.e. the risk-free return over one Δ_t time period is $R_f = e^{r\Delta_t}$.

2.2 Price Effect of Risk due to Discrete-Time Hedging

To focus on the specific risk due discrete-time trading (rather than continuous trading), we consider a stock price that is a diffusion process driven by a Brownian motion with no other state variables. In this case, markets would be complete with continuous trading, and, hence, the dealer's hedging risk arises solely from his trading only at discrete times, spaced Δ_t time units apart.

We are interested in the price of option i as a function of the stock price S_t and demand pressure d_t , $p_t^i = p_t^i(S_t, d_t)$. We denote the price without demand pressure by f , that is, $f^i(t, S_t) := p_t^i(S_t, d = 0)$ and assume throughout that f is smooth for $t < T$. The change in the option price evolves approximately according to

$$p_{t+1}^i \cong f^i + f_S^i \Delta S + \frac{1}{2} f_{SS}^i (\Delta S)^2 + f_t^i \Delta_t \quad (30)$$

where $f^i = f^i(t, S_t)$, $f_t^i = \frac{\partial}{\partial t} f^i(t, S_t)$, $f_S^i = \frac{\partial}{\partial S} f^i(t, S_t)$, $f_{SS}^i = \frac{\partial^2}{\partial S^2} f^i(t, S_t)$, and $\Delta S = S_{t+1} - S_t$. The unhedgeable option price change is

$$e^{r\Delta_t} \bar{p}_{t+1}^i = p_{t+1}^i - e^{r\Delta_t} p_t^i - f_S^i (S_{t+1} - e^{r\Delta_t} S_t) \quad (31)$$

$$\cong -r\Delta_t f^i + f_t^i \Delta_t + r\Delta_t f_S^i S_t + \frac{1}{2} f_{SS}^i (\Delta S)^2 \quad (32)$$

where we expand p_{t+1} and use $e^{r\Delta_t} \cong 1 + r\Delta_t$. To consider the impact of demand d_t^j in option j on the price of option i , we need the covariance of their unhedgeable parts:

$$\text{Cov}_t(e^{r\Delta_t} \bar{p}_{t+1}^i, e^{r\Delta_t} \bar{p}_{t+1}^j) \cong \frac{1}{4} f_{SS}^i f_{SS}^j \text{Var}_t((\Delta S)^2) \quad (33)$$

Hence, by Theorem 2, we get the following result. (Details of the proof are in the appendix.)

Proposition 5 *If the underlying asset price follows a Markov diffusion and the period length is Δ_t , the effect on the price of demand at $d = 0$ is*

$$\frac{\partial p_t^i}{\partial d_t^j} = \frac{\gamma r \text{Var}_t((\Delta S)^2)}{4} f_{SS}^i f_{SS}^j + o(\Delta_t^2) \quad (34)$$

and the effect on the Black-Scholes implied volatility $\hat{\sigma}_t^i$ is:

$$\frac{\partial \hat{\sigma}_t^i}{\partial d_t^j} = \frac{\gamma r \text{Var}_t((\Delta S)^2)}{4} \frac{f_{SS}^i}{\nu^i} f_{SS}^j + o(\Delta_t^2), \quad (35)$$

where ν^i is the Black-Scholes vega.¹¹

Interestingly, the ratio of the Black-Scholes gamma to the Black-Scholes vega, f_{SS}^i/ν^i , does not depend on moneyness, so the first-order effect of demand with discrete trading risk is to change the level, but not the slope, of the implied-volatility curves.

Intuitively, the impact of the demand for options of type j depends on the gamma of these options, f_{SS}^j , since the dealers cannot hedge the non-linearity of the payoff.

The effect of discrete-time trading is small if hedging is frequent. More precisely, the effect is of the order of $\text{Var}_t((\Delta S)^2)$, namely Δ_t^2 . Hence, if we add up T/Δ_t terms of this magnitude — corresponding to demand in each period between time 0 and maturity T — then the total effect is order Δ_t , which approaches zero as the Δ_t approaches zero. This is consistent with the Black-Scholes-Merton result of perfect hedging in continuous time. As we show next, the risks of jumps and stochastic volatility do not vanish for small Δ_t (specifically, they are of order Δ_t).

2.3 Jumps in the Underlying

To study the effect of jumps in the underlying, we suppose next that S is a discretely traded jump diffusion with i.i.d. bounded jump size, independent of the state variables, and jump intensity π (i.e. jump probability over a period of $\pi\Delta_t$).

The unhedgeable price change is

$$e^{r\Delta_t} \bar{p}_{t+1}^i \cong -r\Delta_t f^i + f_t^i \Delta_t + r\Delta_t f_S^i S_t + (f_S^i S_t - \theta^i) \Delta S 1_{(\text{no jump})} + \kappa^i 1_{(\text{jump})} \quad (36)$$

where

$$\kappa^i = f^i(S_t + \Delta S) - f^i - \theta^i \Delta S. \quad (37)$$

is the unhedgeable risk in case of a jump of size ΔS .

Proposition 6 *If the underlying asset price can jump, the effect on the price of demand at $d = 0$ is*

$$\frac{\partial p_t^i}{\partial d_t^j} = \gamma r [(f_S^i S_t - \theta^i)(f_S^j S_t - \theta^j) \text{Var}_t(\Delta S) + \pi \Delta_t E_t(\kappa^i \kappa^j)] + o(\Delta_t) \quad (38)$$

¹¹Even though the volatility is constant within the Black-Scholes model, we follow the standard convention that defines the Black-Scholes implied volatility as the volatility that, when fed into the Black-Scholes model, makes the model price equal to the option price, and the Black-Scholes vega as the partial derivative measuring the change in the option price when the volatility fed into the Black-Scholes model changes.

and the effect on the Black-Scholes implied volatility $\hat{\sigma}_t^i$ is:

$$\frac{\partial \hat{\sigma}_t^i}{\partial d_t^j} = \frac{\gamma r [(f_S^i S_t - \theta^i)(f_S^j S_t - \theta^j) \text{Var}_t(\Delta S) + \pi \Delta_t E_t(\kappa^i \kappa^j)]}{\nu^i} + o(\Delta_t), \quad (39)$$

where ν^i is the Black-Scholes vega.

The terms of the form $f_S^i S_t - \theta^i$ arise because the optimal hedge θ differs from the optimal hedge without jumps, $f_S^i S_t$, which means that some of the local noise is being hedged imperfectly. If the jump probability is small, however, then this effect is small (i.e., it is second order in π). In this case, the main effect comes from the jump risk κ . We note that while conventional wisdom holds that Black-Scholes gamma is a measure of “jump risk,” this is true only for the small local jumps considered in Section 2.2. Large jumps have qualitatively different implications captured by κ . For instance, a far-out-of-the-money put may have little gamma risk, but, if a large jump can bring the option in the money, the option may have κ risk. It can be shown that this jump-risk effect (39) means that demand can affect the slope of the implied-volatility curve to the first order and generate a smile.¹²

2.4 Stochastic-Volatility Risk

To consider stochastic volatility, we let the the state variable be $X_t = (S_t, \sigma_t)$, where the stock price S is a diffusion with volatility σ_t , which is also a diffusion, driven by an independent Brownian motion. The option price $p_t^i = f^i(t, S_t, \sigma_t)$ has unhedgeable risk given by

$$e^{r\Delta_t} \bar{p}_{t+1}^i = p_{t+1}^i - e^{r\Delta_t} p_t^i - \theta^i R_{t+1}^e \quad (40)$$

$$\cong -r\Delta_t f^i + f_t^i \Delta_t + f_S^i S_t r \Delta_t + f_\sigma^i \Delta \sigma_{t+1} \quad (41)$$

Proposition 7 *With stochastic volatility, the effect on the price of demand at $d = 0$ is*

$$\frac{\partial p_t^i}{\partial d_t^j} = \gamma r \text{Var}(\Delta \sigma) f_\sigma^i f_\sigma^j + o(\Delta_t) \quad (42)$$

and the effect on the Black-Scholes implied volatility $\hat{\sigma}_t^i$ is:

$$\frac{\partial \hat{\sigma}_t^i}{\partial d_t^j} = \gamma r \text{Var}(\Delta \sigma) \frac{f_\sigma^i}{\nu^i} f_\sigma^j + o(\Delta_t), \quad (43)$$

where ν^i is the Black-Scholes vega.

¹²Of course, the jump risk also generates smiles without demand-pressure effects; the result is that demand can exacerbate these.

Intuitively, volatility risk is captured to the first order by f_σ . This derivative is not exactly the same as Black-Scholes vega, since vega is the price sensitivity to a permanent volatility change whereas f_σ measures the price sensitivity to a volatility change that mean reverts at the rate of ϕ . For an option with maturity at time $t + T$, we have

$$f_\sigma^i \cong \nu^i \frac{\partial}{\partial \sigma_t} E \left(\frac{\int_t^{t+T} \sigma_s ds}{T} \mid \sigma_0 \right) \cong \nu^i \frac{1 - e^{-\phi T}}{\phi T}. \quad (44)$$

Hence, combining (44) with (43) shows that stochastic volatility risk affects the level, but not the slope, of the implied volatility curves to the first order.

3 Descriptive Statistics

The main focus of this paper is the impact of net end-user option demand on option prices. We explore this impact both for S&P 500 index options and for equity (i.e., individual stock) options. Consequently, we employ data on SPX and equity option demand and prices. Our data period extends from the beginning of 1996 through the end of 2001. For the equity options, we limit the underlying stocks to those with strictly positive option volume on at least 80% of the trade days over the 1996 to 2001 period. This restriction yields 303 underlying stocks.

We acquire the data from two different sources. Data for computing net option demand were obtained directly from the Chicago Board Options Exchange (CBOE). These data consist of a daily record of closing short and long open interest on all SPX and equity options for public customers and firm proprietary traders.¹³ The SPX options trade only at the CBOE while the equity options sometimes are cross-listed at other option markets. Our open interest data, however, include activity from all markets at which CBOE listed options trade. The entire option market is comprised of public customers, firm proprietary traders, and market makers. Hence, our data cover all non-market-maker option open interest.

We compute the net demand for options under the assumption that public customers and firm proprietary traders — that is, all non-market makers — are end users. Consequently, the net demand for an option is defined as the sum of the public-customer and firm proprietary-trader long open interest minus the sum of the public-customer and firm proprietary-trader short open interest.

We have also computed the net demand for options under the assumption that only public customers are end users, i.e., that both firm proprietary traders and market makers are liquidity suppliers. Since proprietary traders constitute a relatively small

¹³The total long open interest for any option always equals the total short open interest. For a given investor type (e.g., public customers), however, the long open interest is not equal to the short open interest in general.

group in our data, none of the main features of the descriptive statistics presented in this section or the results presented in the next section changes under this alternative assumption. In fact, our results are slightly stronger with this alternative definition of end users. This may be because a non-trivial part of the firm proprietary trading is associated with supplying liquidity, and, indeed, proprietary traders' and market makers' net positions roughly move with one another (the time-series correlation is 0.44). However, to avoid datamining, we present the more conservative results.

Even though the SPX and individual equity option market have been the subject of extensive empirical research, there is no systematic information on end-user demand in these markets. Consequently, we provide a somewhat detailed description of net demand for SPX and equity options. Over the 1996-2001 period the average daily non-market-maker net demand for SPX options is 103,254 contracts. For puts (calls), the average daily net demand from non-market makers is 124,345 (−21,091) contracts. These numbers indicate that most net option demand comes from puts. Indeed, end users tend to be net suppliers of on the order of 20,000 call contracts.

For the equity options, the average daily non-market-maker net demand per underlying stock is −2717 contracts. Hence, in the equity option market, unlike the index-option market, end users are net suppliers of options. This fact suggests that if demand for options has a first order impact on option prices, index options should on average be more expensive than individual equity options. Another interesting contrast with the index option market is that in the equity option market the net end-user demand for puts and calls is similar. For puts (calls), the average daily non-market-maker net demand is −1103 (−1614) contracts.

Panel A of Table 1 reports the average daily non-market-maker net demand for SPX options broken down by option maturity and moneyness (defined as the strike price divided by the underlying index level.) Since our theoretical results indicate that the demand from a put or a call with the same strike price and maturity should have identical price impact, this table is constructed from the demands for puts and calls of all moneyness and maturity. For instance, the moneyness range 0.95-1 consists of put options that are up to 5% in-the-money and call options that are up to 5% out-of-the-money. Panel A indicates that 39 percent of the net demand comes from contracts with fewer than 30 calendar days to expiration. Consistent with conventional wisdom, the good majority of this net demand is concentrated at moneyness where puts are out-of-the-money (OTM) (i.e., *moneyness* < 1.) Panel B of Table 1 reports the average option net demand per underlying stock for individual equity options from non-market makers. With the exception of some long maturity option categories (i.e, those with more than one year to expiration and in one case with more than six months to expiration), the non-market-maker net demand for all of the moneyness/maturity categories is negative. That is, non-market makers are net suppliers of options in all of these categories. This stands in stark contrast to the index option market in Panel A where non-market makers are net demanders of options in almost every moneyness/maturity category.

Table 1: Average non-market-maker net demand for put and call option contracts for SPX and individual equity options by moneyness and maturity, 1996-2001. Equity-option demand is per underlying stock.

Maturity Range (Calendar Days)	Moneyness Range (K/S)								All
	0-0.85	0.85-0.90	0.90-0.95	0.95-1.00	1.00-1.05	1.05-1.10	1.10-1.15	1.15-2.00	
Panel A: SPX Option Non-Market Maker Net Demand									
1-9	6,014	1,780	1,841	2,357	2,255	1,638	524	367	16,776
10-29	7,953	1,300	1,115	6,427	2,883	2,055	946	676	23,356
30-59	5,792	745	2,679	7,296	1,619	-136	1,038	1,092	20,127
60-89	2,536	1,108	2,287	2,420	1,569	-56	118	464	10,447
90-179	7,011	2,813	2,689	2,083	201	1,015	4	2,406	18,223
180-364	2,630	3,096	2,335	-1,393	386	1,125	-117	437	8,501
365-999	583	942	1,673	1,340	1,074	816	560	-1,158	5,831
All	32,519	11,785	14,621	20,530	9,987	6,457	3,074	4,286	103,260
Panel B: Equity Option Non-Market Maker Net Demand									
1-9	-51	-25	-40	-45	-47	-31	-23	-34	-295
10-29	-64	-35	-57	-79	-102	-80	-55	-103	-576
30-59	-55	-31	-39	-55	-88	-90	-72	-144	-574
60-89	-47	-29	-37	-47	-60	-60	-55	-133	-469
90-179	-85	-60	-73	-84	-105	-111	-101	-321	-941
180-364	53	-19	-23	-24	-36	-35	-33	-109	-225
365-999	319	33	25	14	12	7	9	-56	363
All	70	-168	-244	-320	-426	-400	-331	-899	-2717

The other main source of data for this paper is the Ivy DB data set from OptionMetrics LLC. The OptionMetrics data include end-of-day volatilities implied from option prices, and we use the volatilities implied from SPX and CBOE listed equity options from the beginning of 1996 through the end of 2001. SPX options have European style exercise, and OptionMetrics computes implied volatilities by inverting the Black-Scholes formula. When performing this inversion, the option price is set to the midpoint of the best closing bid and offer prices, the interest rate is interpolated from available LIBOR rates so that its maturity is equal to the expiration of the option, and the index dividend yield is determined from put-call parity. The equity options have American style exercise, and OptionMetrics computes their implied volatilities using binomial trees that account for the early exercise feature and the timing and amount of the dividends expected to be paid by the underlying stock over the life of the options.

One of the central questions we are investigating is whether net demand pressure pushes option implied volatilities away from the volatilities that are expected to be realized over the remainder of the options' lives. We refer to the difference between implied volatility and a reference volatility estimated from the underlying security as *excess implied volatility*.

The reference volatility that we use for SPX options is the filtered volatility from the state-of-the-art model by Bates (2005), which accounts for jumps, stochastic volatility, and the risk premium implied by the equity market, but does not add extra risk premia

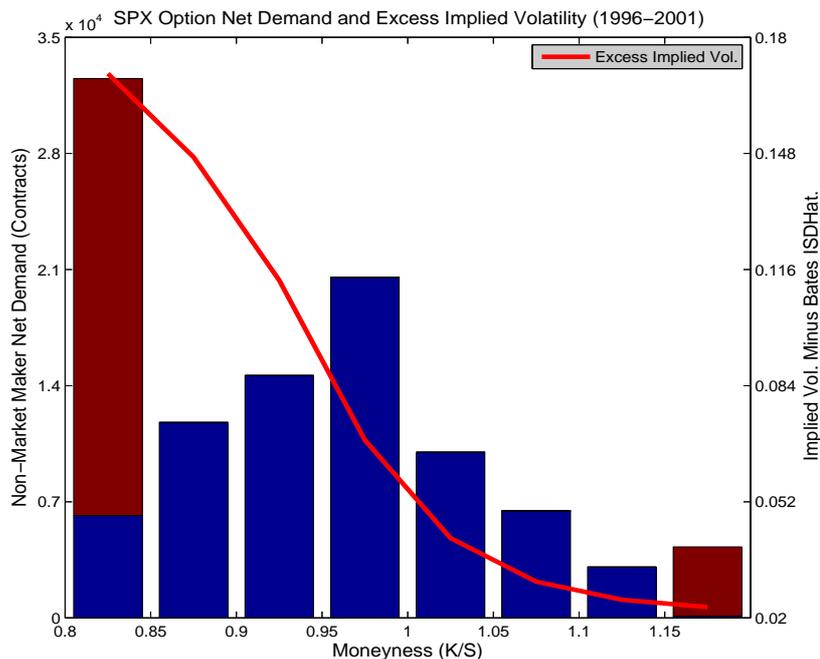


Figure 1: The bars show the average daily net demand for puts and calls from non-market makers for SPX options in the different moneyness categories (left axis). The top part of the leftmost (rightmost) bar shows the net demand for all options with moneyness less than 0.8 (greater than 1.2). The line is the average SPX excess implied volatility, that is, implied volatility minus the volatility from the underlying security filtered using Bates (2005), for each moneyness category (right axis). The data cover 1996-2001.

to (over-)fit option prices.¹⁴

The reference volatility that we use for equity options is the predicted volatility over their lives from a GARCH(1,1) model estimated from five years of daily underlying stock returns leading up to the day of option observations.¹⁵ (Alternative measures using historical or realized volatility lead to similar results.) The daily returns on the underlying index or stocks are obtained from the Center for Research in Security Prices

¹⁴We are grateful to David Bates for providing this measure.

¹⁵In particular, we use the GARCH(1,1) parameter estimates for the trade day (estimated on a rolling basis from the past five years of daily data) to compute the minimum mean square error volatility forecast for the number of trade days left in the life of the option. We annualize the volatility forecast (which is for the number of trade days left until the option matures) by multiplying by the square root of 252 and dividing by the square root of the number of trade days remaining in the life of the option.

(CRSP).

The daily average excess implied volatility for SPX options is 8.7%. To compute this number, on each trade day we average the implied volatilities on all SPX options that have at least 25 contracts of trading volume and then subtract the proxy for expected volatility. Consistent with previous research, on average the SPX options in our sample are expensive.¹⁶ For the equity options, the daily average excess implied volatility per underlying stock is -0.3%, which suggests that on average individual equity options are just slightly inexpensive. We required that an option trade at least 5 contracts and have a closing bid price of at least 37.5 cents in order to include its implied volatility in the calculation.

Figure 1 compares SPX option expensiveness to net demands across moneyness categories. The line in the figure plots the average SPX excess implied volatility for eight moneyness intervals over the 1996-2001 period. In particular, on each trade date the average excess implied volatility is computed for all puts and calls in a moneyness interval. The line depicts the means of these daily averages. The excess implied volatility inherits the familiar downward sloping smirk in SPX option implied volatilities. The bars in Figure 1 represent the average daily net demand from non-market maker for SPX options in the moneyness categories, where the top part of the leftmost (rightmost) bar shows the net demand for all options with moneyness less than 0.8 (greater than 1.2).

The first main feature of Figure 1 is that index options are expensive (i.e. have a large risk premium), consistent with what is found in the literature, and that end users are net buyers of index options. This is consistent with our main hypothesis: end users buy index options and market makers require a premium to deliver them.

The second main feature of Figure 1 is that the net demand for low-strike options is greater than the demand for high-strike options. This could help explain the fact that low-strike options are more expensive than high-strike options (Proposition 4).

The shape of the demand across moneyness is clearly different from the shape of the expensiveness curve. This is expected for two reasons. First, our theory implies that demand pressure in one moneyness category impacts the implied volatility of options in other categories, thus “smoothing” the implied volatility curve and changing its shape. Second, our theory implies that demands (weighted by the variance of the unhedgeable risks) affect prices, and the price effect must then be translated into volatility terms. It follows that a left-skewed hump-shaped price effect typically translates into a downward sloping volatility effect, consistent with the data. In fact, the observed average demands can give rise to a pattern of expensiveness similar to the one observed empirically when using a version of the model with jump risk. It is helpful to link these demands more

¹⁶Short maturity close to the money SPX options are about half as expensive. For example, if we restrict the SPX options to those that have moneyness between 0.98 and 1.02, maturity between 15 and 45 calendar days, and at least 25 contracts of trading volume the average excess implied volatility is 4.6%.

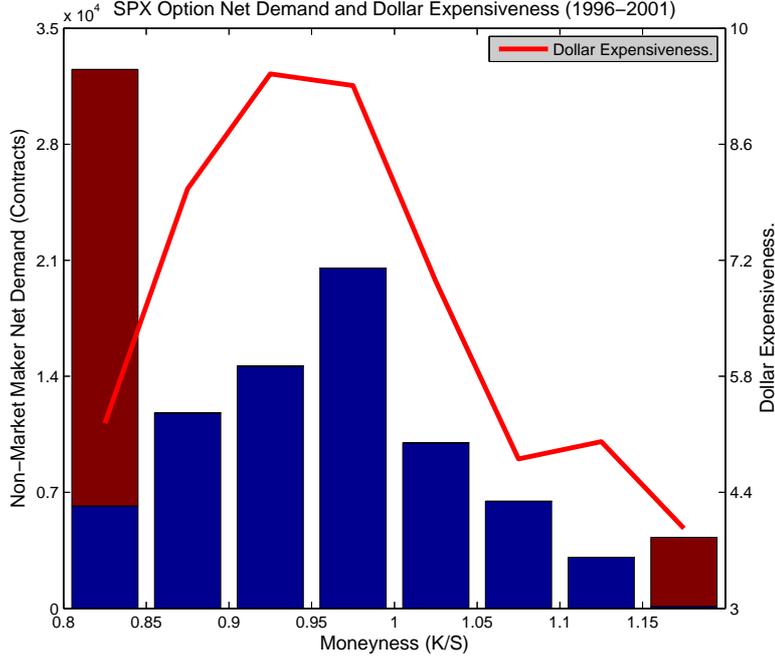


Figure 2: The bars show the average daily net demand for puts and calls from non-market makers for SPX options in the different moneyness categories (left axis). The top part of the leftmost (rightmost) bar shows the net demand for all options with moneyness less than 0.8 (greater than 1.2). The line is the average SPX price expensiveness, that is, the option price minus the “fair price” implied by Bates (2005) using filtered volatility from the underlying security and equity risk premia, for each moneyness category (right axis). The data cover 1996-2001.

directly to the predictions of our theory. Our model shows that every option contract demanded leads to an increase in its price — in dollar terms — proportional to the variance of its unhedgeable part (and an increase in any other option prices proportional to the covariance of the unhedgeable parts of the two options). Hence, the relationship between raw demands (that is, demands not weighted according to the model) and expensiveness is more directly visible when expensiveness is measured in dollar terms, rather than in terms of implied volatility. This fact is confirmed by Figure 2. Indeed, the price expensiveness has a similar shape to the demand pattern. Because of the cross-option effects and the absence of the weighting factor (the variance terms), we do not expect the shapes to be identical.¹⁷

¹⁷Even if the model-implied relationship involving dollar expensiveness is more direct, we follow the literature and concentrate on expensiveness expressed in terms of implied volatility. This can be

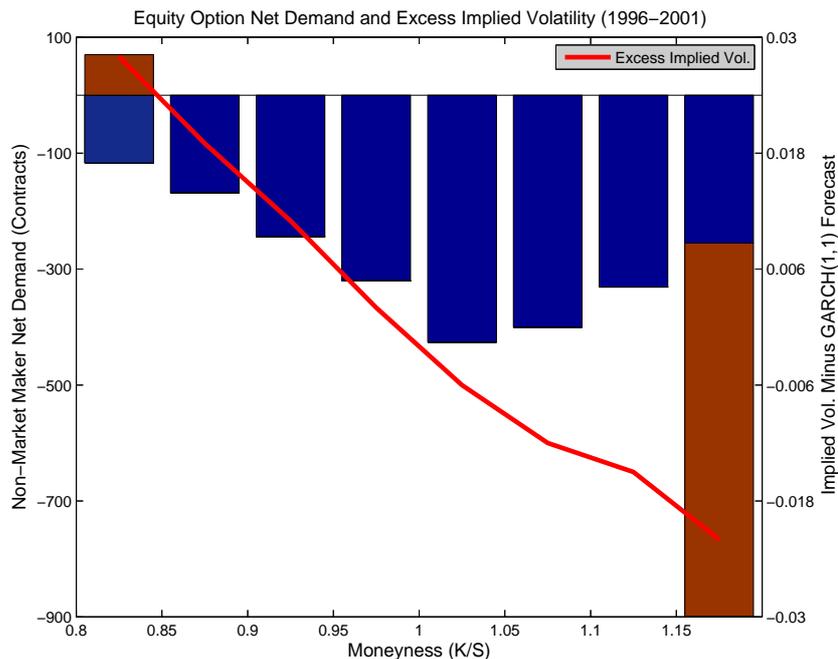


Figure 3: The bars show the average daily net demand per underlying stock from non-market makers for equity options in the different moneyness categories (left axis). The top part of the leftmost (rightmost) bar shows the net demand for all options with moneyness less than 0.8 (greater than 1.2). The line is the average equity option excess implied volatility, that is, implied volatility minus the GARCH(1,1) expected volatility, for each moneyness category (right axis). The data cover 1996-2001.

Figure 3 compares equity option expensiveness to net demands across moneyness categories. The line in the figure plots the average equity option excess implied volatility (with respect to the GARCH(1,1) volatility forecast) per underlying stock for eight moneyness intervals over the 1996-2001 period. In particular, on each trade date for each underlying stock the average excess implied volatility is computed for all puts and calls in a moneyness interval. These excess implied volatilities are averaged across underlying stocks on each trade day for each moneyness interval. The line depicts the means of these daily averages. The excess implied volatility line is downward sloping but only varies by about 5% across the moneyness categories. By contrast, for the SPX options the excess implied volatility line varies by 15% across the corresponding moneyness categories. The bars in the figure represent the average daily net demand

thought of as a normalization that eliminates the need for explicit controls for the price level of the underlying asset.

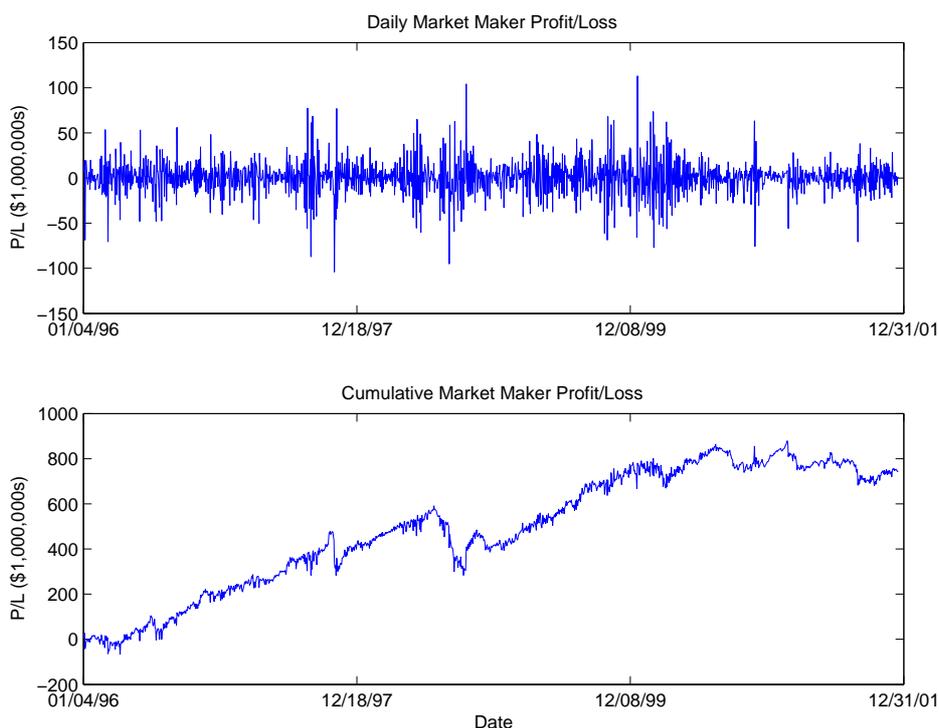


Figure 4: The top panel shows the market makers' daily profits and losses (P&L) assuming they delta-hedge their option positions once per day. The bottom panel shows the corresponding cumulative P&Ls.

per underlying stock from non-market makers for equity options in the moneyness categories. The figure shows that non-market makers are net sellers of equity options on average, consistent with these options being cheap. Further, the figure shows that non-market makers sell mostly high-strike options, consistent with these options being especially cheap.

To illustrate the magnitude of the net demands, we compute approximate daily profits and losses (P&Ls) for the market makers' hedged positions assuming daily delta-hedging. The daily and cumulative P&Ls are illustrated in Figure 4, which shows that the group of market makers faces substantial risk that cannot be delta-hedged, with daily P&L varying between ca. \$100M and \$-100M. Further, the market makers make cumulative profits of ca. \$800M over the 6-year period on their position taking.¹⁸ With just over a hundred SPX market makers on the CBOE, this corresponds to a profit of approximately \$1M per year per market maker. Hence, consistent with the premise of

¹⁸This number does not take into account the costs of market making or the profits from the bid-ask spread on round-trip trades. A substantial part of market makers' profit may come from the latter.

our model, market makers face substantial risk and are compensated on average for the risk that they take.

4 Empirical Results

Theorem 2 relates the demand for any option to a price impact on any option. Since our data contain both option demands and prices, we can test these theoretical results directly. Doing so requires that we choose a reason for the underlying asset and risk-free bond to form a dynamically incomplete market, and, hence, for the weighting factors $\text{Cov}_t(\bar{p}_{t+1}^i, \bar{p}_{t+1}^j)$ to be non-zero. As source of market incompleteness, we consider discrete-time trading, jumps, and stochastic volatility using Propositions 5–7.

We test the model’s ability to help reconcile the two main puzzles in the option literature, namely the drivers of the overall level of implied volatility and its skew across option moneyness. The first set of tests investigates whether the overall excess implied volatility is higher on trade dates where the demand for options — aggregated according to the model — is higher. The second set of tests investigates whether the excess implied volatility skew is steeper on trade dates where the model-implied demand-based skew is steeper.

4.1 Excess Implied Volatility and Net Demand

We investigate first the time-series evidence for Theorem 2 by regressing a measure of excess implied volatility on one of various demand-based explanatory variables:

$$\text{ExcessImplVol}_t = a + b \text{DemandVar}_t + \epsilon_t \quad (45)$$

SPX:

We consider first this time-series relationship for SPX options, for which we define *ExcessImplVol* as the average implied volatility of approximately one month to maturity and at-the-money SPX options minus the corresponding volatility of Bates (2005). When computing this variable, the SPX options included are those that have at least 25 contracts of trading volume, between 15 and 45 calendar days to expiration,¹⁹ and moneyness between 0.99 and 1.01. (We compute the excess implied volatility variable only from reasonably liquid options in order to make it less noisy in light of the fact that it is computed using only one trade date.)²⁰ By subtracting the volatility from

¹⁹On any give trade day, these are the options with maturity closest to one month. Alternatively, for each month we could include in our test only the day that is precisely one month before expiration. This approach yields similar results.

²⁰By contrast, in the previous section of the paper, when implied volatility statistics were computed from less liquid options or options with more extreme moneyness or maturity, they were averaged over the entire sample period.

the Bates (2005) model, we account for the direct effects of jumps, stochastic volatility, and the risk premium implied by the equity market.

The independent variable, *DemandVar*, is based on the aggregate net non-market-maker demand for SPX options that have 10–180 calendar days to expiration and moneyness between 0.8 and 1.20. We employ, separately, four different independent variables. The first is simply the sum of all net demands. Although the model does not yield this specification, we use it as a simple atheoretical alternative. According to Proposition 3, there should be a positive, albeit noisy, relationship between total demand and expensiveness. The other three independent variables correspond to “weighting” the net demands using the models based on the market maker risks associated with, respectively, discrete trading, jumps in the underlying, and stochastic volatility (Sections 2.2–2.4). Specifically, the net demands are weighted by the Black-Scholes gamma in the discrete-hedging model, by kappa computed using equally likely up and down moves of relative sizes 0.05 and 0.2 in the jump model, and by maturity-adjusted Black-Scholes vega in the stochastic volatility model. The appendix provides more details on the computation of the model-based weighting factors.

We run the regression on a monthly basis by averaging demand and expensiveness over each month. We do this because there are certain day-of-the-month effects for SPX options. (Our results are similar in an unreported daily regression.)

The results are shown in Table 2. We report the results over two subsamples because there are reasons to suspect a structural change in 1997. The change, apparent also in the time series of open interest and market-maker and public-customer positions (not shown here), stems from several events that altered the market for index options in the period from late 1996 to October 1997, such as the introduction of S&P500 e-mini futures and futures options on the competing Chicago Mercantile Exchange (CME), the introduction of Dow Jones options on the CBOE, and changes in margin requirements. Our results are robust to the choice of these sample periods as discussed below.

We see that the estimate of the demand effect b is positive but insignificant over the first subsample, and positive and statistically significant over the second, longer, subsample for all three model-based explanatory variables.²¹

The expensiveness and the fitted values from the jump model are plotted in Figure 5, which clearly shows their comovement over the later sample. The fact that the b coefficient is positive indicates that, on average, when SPX net demand is higher (lower), SPX excess implied volatilities are also higher (lower). For the most successful model, the one based on jumps, changing the dependent variable from its lowest to its highest values over the late sub-sample would change the excess implied volatility by about 5.6 percentage points. A one-standard-deviation change in the jump-based demand variable results in a one-half standard deviation change in excess implied

²¹The model-based explanatory variables work better than just adding all contracts (*#Contracts*), because they give greater weight to near-the-money options. If we just count contracts using a more narrow band of moneyness, then the *#Contracts* variable also becomes significant.

Table 2: The relationship between the SPX Excess Implied Volatility (i.e. observed implied volatility minus volatility from the Bates (2005) model) and the SPX non-market-maker demand pressure weighted using either: (i) equal weights, (ii) weights based on discrete-time trading risk, (iii) weights based on jump risk, or (iv) weights based on stochastic-volatility risk. T-statistics computed using Newey-West are in parentheses.

	Before Structural Changes 1996/01–1996/10				After Structural Changes 1997/10–2001/12			
Constant	0.0001 (0.004)	0.0065 (0.30)	0.005 (0.17)	0.020 (0.93)	0.04 (7.28)	0.033 (4.67)	0.032 (7.7)	0.038 (7.4)
#Contracts	2.1E-7 (0.87)				3.8E-7 (1.55)			
Disc. Trade		6.9E-10 (0.91)				2.8E-9 (3.85)		
Jump Risk			6.4E-6 (0.79)				3.2E-5 (3.68)	
Stoch. Vol.				8.7E-7 (0.27)				1.1E-5 (2.74)
Adj. R^2 (%)	9.6	6.1	8.1	0.5	7.0	18.5	25.9	15.7
N	10	10	10	10	50	50	50	50

volatility (the corresponding R^2 is 26%). The model is also successful in explaining a significant proportion of the level of the excess implied volatility. Over the late sub-sample, the average level is 4.9%, of which approximately one third – specifically, 1.7% – corresponds to the average level of demand, given the regression coefficient.

We note that, in addition to the CBOE demand pressure observed in our data, there is over-the-counter demand for index options, for instance, products for individual investors such as index-linked bonds. These securities give end users essentially a riskfree security in combination with a call option on the index (or the index plus a put option), which leaves Wall Street short index options. Of course, this demand also contributes to the excess implied volatility. Hence, our estimate that demand pressure explains on average 1.7 percentage points (and 26% of the variation) of the excess implied volatility can be viewed as a conservative lower bound.

Further support for the hypothesis that the supply for options is upwardly sloping comes from the comparison between the estimated supply-curve slopes following market

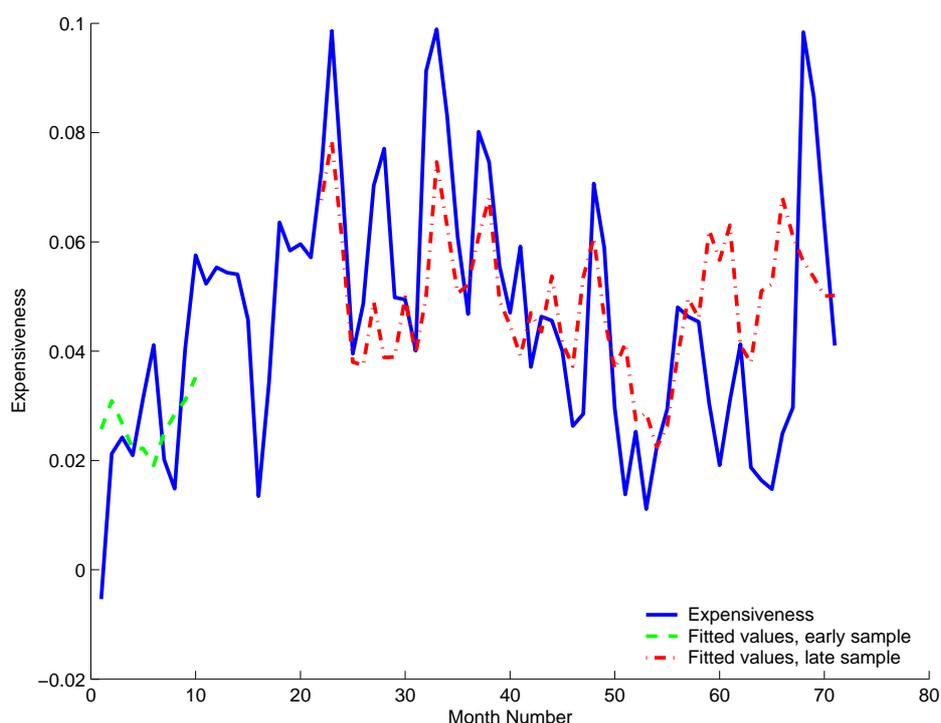


Figure 5: The solid line shows the expensiveness of SPX options, that is, implied volatility of 1-month at-the-money options minus the volatility measure of Bates (2005) which takes into account jumps, stochastic volatility, and the risk premium from the equity market. The dashed lines are, respectively, the fitted values of demand-based expensiveness using a model with underlying jumps, before and after certain structural changes (1996/01–1996/10 and 1997/10–2001/12).

maker losses, respectively, gains. If market maker risk aversion plays an important role in pricing options, then one would expect prices to be less sensitive to demand when market makers are well funded following profitable periods. This is exactly what we find. Breaking the daily sample²² in two subsamples depending on whether the hedged market maker profits over the previous 20 trading days is positive or negative,²³ we estimate the regression (45) for each subsample and find that, following losses, the b coefficient is approximately twice as large as the coefficient obtained in the other subsample. For instance, in the jump model, the regression coefficient following losses is $2.6\text{E-}05$ with a t -statistic of 3.7 (330 observations), compared to a value of $1.1\text{E-}5$ with t -statistic of 6.1 in the complementary subsample (646 observations).

²²Because of the structural changes discussed above, we restrict our attention to the period starting on October 1, 1997.

²³Similar results obtain if the breaking point is the mean or median daily profit.

As explained above, we split our sample into earlier and later subsamples due to exogenous events leading to non-stationarities. We check for robustness of the specific choice of sample periods by redoing the analysis with the sample periods lengthened or shortened by 2 or 4 months. This does not change our results qualitatively; indeed, the demand effect in the later subsample remains positive and significant, with the explanatory power virtually unchanged. Due to the structural break, no particular results are expected in the full sample. It turns out that some of our results do hold over the full sample, albeit less robustly to model specification and with smaller explanatory power, while others do not. Of course, we must entertain the possibility that the model's limited ability to jointly explain the full sample is due to problems with the theory. Also, our results are robust to our definition of end users as public customers plus proprietary traders. Similar, in fact stronger, results obtain when end user demand is defined simply as public-customer demand.

Equity Options:

We consider next the time-series relationship between demand and expensiveness for equity options. In particular, we run the time-series regression (45) for each stock, and average the coefficients across stocks. The results are shown in Table 3. We consider separately the subsamples before and after the summer of 1999. This is because most options were listed only on one exchange before the summer of 1999, but many were listed on multiple exchanges after this summer. Hence, there was potentially a larger total capacity for risk taking by market makers after the cross listing. See for instance De Fontnouvelle, Fishe, and Harris (2003) for a detailed discussion of this well-known structural break. We follow Fontnouvelle et al. in the specific choice of the starting and ending dates for our subsamples, but we have confirmed that our results are robust to variations in this choice. For example, if the breakpoint for either subsample is moved by two months in either direction, there are only very small changes to the numbers reported in Table 3.

We run the time-series regression separately for the 303 underlying stocks with strictly positive option volume on at least 80% of the trade days from the beginning of 1996 through the end of 2001. We compute the excess implied volatility as the average implied volatility of selected options²⁴ minus the GARCH(1,1) volatility. The net demand as the total net non-market-maker demand for options with moneyness between 0.8 and 1.2 and maturity between 10 and 180 calendar days, which for simplicity we aggregate without weighting. We run the regression using monthly data on underlying stocks that have at least 12 months of data available. (Daily regressions give stronger results.)

The average coefficient b measuring the effect of demand on expensiveness is positive and significant in both subsamples. This means that when the demand for equity

²⁴That is, options with moneyness between 0.95 and 1.05, maturity between 15 and 45 calendar days, at least 5 contracts of trading volume, and implied volatilities available on OptionMetrics.

Table 3: The relationship between option expensiveness — i.e. implied volatility minus GARCH volatility — and non-market-maker net demand for equity options on 303 different underlying stock (Equation 45). We run time-series regressions for each underlying and report the average coefficients. The number p-val is the p-value of the binomial test that the coefficients are equally likely to be positive and negative.

	Before Cross-Listing of Options 04/1996–06/1999	After Cross-Listing of Options 10/1999–12/2001
Constant	-0.01	0.02
<i>NetDemand</i>	12E-6	8.5E-6
Average Adj. R^2 (%)	6	5
# positive	227	213
# negative	76	86
p-val	0.00	0.00

options is larger their implied volatility is higher. In the first subsample, the coefficient estimate is positive for 227 underlying stocks and negative for only 76.²⁵ In the second subsample, the coefficient estimate is positive for 213 underlying stocks and negative for only 86.²⁶ The results are illustrated in Figure 6, which shows the expensiveness and fitted values of the demand effect on a monthly basis. The positive correlation is apparent. We note that the relation between average demand and average expensiveness is more striking if we do a single regression for these averaged variables. Also, the coefficient b is significant in both subsamples for the regression of average demand on average expensiveness. It is comforting that the relation also holds both when considering averages and when considering each stock separately.

Finally, we investigate the cross-sectional relationship between excess implied volatility and net demand in the equity option market. We multiply the net demand variable by the price volatility of the underlying stock (defined as the sample return volatility just described multiplied by the day’s closing price of the stock.) We scale the net demand in this way because market makers are likely to be more concerned about holding net demand in their inventory when the underlying stock’s price volatility is greater.

We run the cross-sectional regression on each day and then employ the Fama-MacBeth method to compute point estimates and standard errors. We also use the Newey-West procedure to control for serial-correlation in the slope estimates. The slope coefficient is 5.9E-8 with a t-statistic of 6.44.

²⁵Of these coefficients, 92, respectively 11, are significant at a 5% level.

²⁶Of these coefficients, 75, respectively 7, are significant at a 5% level.

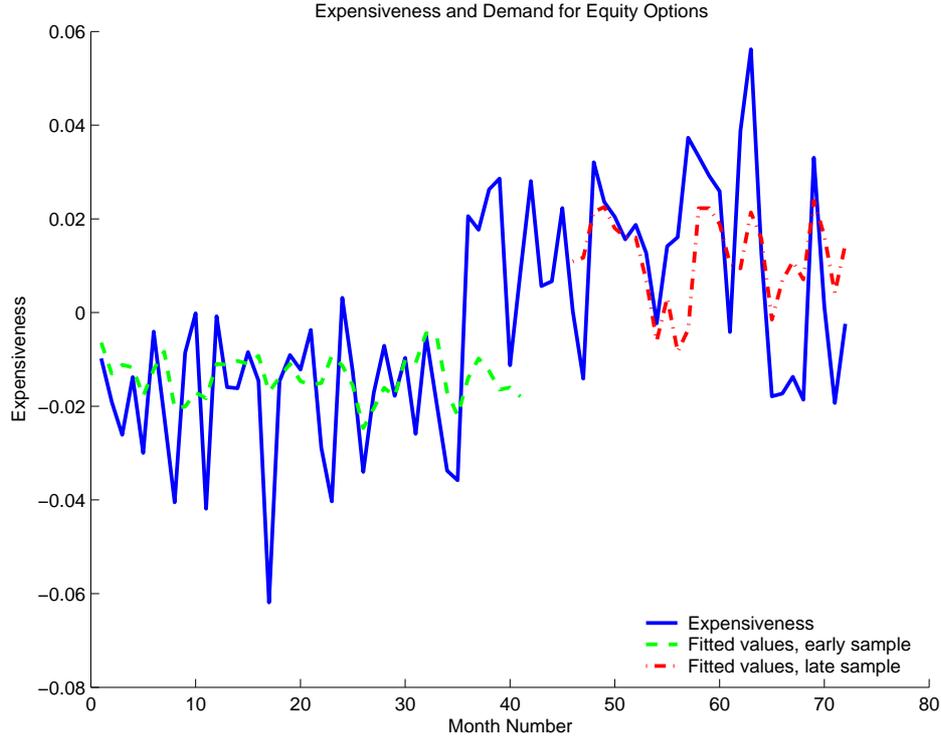


Figure 6: The solid line is the expensiveness of equity options, averaged across stocks. The dashed lines are, respectively, the fitted values of demand-based expensiveness before and after the cross-listing of options (1996/04–1999/05 and 1999/10–2001/12) using the average regression coefficients from stock-specific regressions and the average demand.

4.2 Implied Volatility Skew and Net Demand Skew

Our second set of tests investigates the explanatory power of the model for the skew of the implied-volatility curve. Specifically, we regress a measure of the steepness of the excess implied volatility skew on one of two demand-based explanatory variables:

$$ExcessImplVolSkew_t = a + b DemandVarSkew_t + \epsilon_t. \quad (46)$$

SPX:

For the SPX analysis, $ExcessImplVolSkew_t$ is the date- t implied volatility skew over and above the skew predicted by the jumps and stochastic volatility of the underlying index. The implied volatility skew is defined as the average implied volatility of options with moneyness between 0.93 and 0.95 that trade at least 25 contracts on trade date t and have more than 15 and fewer than 45 calendar days to expiration, minus the average implied volatility of options with moneyness between 0.99 and 1.01 that meet

Table 4: The relationships between SPX implied-volatility skew and two potential explanatory variables: (i) the SPX non-market-maker demand skew and (ii) the jump-risk-model demand-based implied-volatility skew. T-statistics computed using Newey-West are in parentheses.

	Before Structural Changes 1996/01–1996/10		After Structural Changes 1997/10–2001/12	
Constant	0.032 (20)	0.04 (5.4)	0.04 (29)	0.032 (14)
#Contracts	-1.2E-7 (-0.8)		5.4E-7 (3.7)	
Jump Risk		-3.0E-6 (-1.1)		1.8E-5 (4.3)
Adj. R^2 (%)	7.6	13.8	22.1	28.2
N	10	10	50	50

the same volume and maturity criteria. In order to eliminate the skew that is due to jumps and stochastic volatility of the underlying, we consider the implied volatility skew net of the similarly defined volatility skew implied by the objective distribution of Broadie, Chernov, and Johannes (2005) where the underlying volatility is that filtered from the Bates (2005) model.²⁷

As explanatory variable, we use two measures. The first is the skew in net option demand, defined as the net non-market-maker demand for options with moneyness between 0.93 and 0.95 minus the net non-market-maker demand for options with moneyness between 0.99 and 1.01, using options with maturity between 10 and 180 calendar days. This measure is not generated as such by the model, but is supported by Proposition 4. The second measure is the excess implied-volatility skew from the model with underlying jumps described in Section 2.3. (We do not consider the models with discrete trading and stochastic volatility, since they do not have first-order skew implications, as explained in Sections 2.2 and 2.4.)

Table 4 reports the monthly OLS estimates of the skewness regression and Figure 7 illustrates the effects. As discussed in Section 4.1, we divide the sample into two subsamples because of structural changes. The slope coefficient is significantly positive in the late subsample using both the simple demand variable and the variable based on jump risk.²⁸ Similar results obtain when we compute the dependent variable

²⁷The model-implied skew is evaluated for one-month options with moneyness of, respectively, 0.94 and 1. We thank Mikhail Chernov for providing this time series.

²⁸The slope coefficient is also significant over the full sample; the demand skewness is less non-

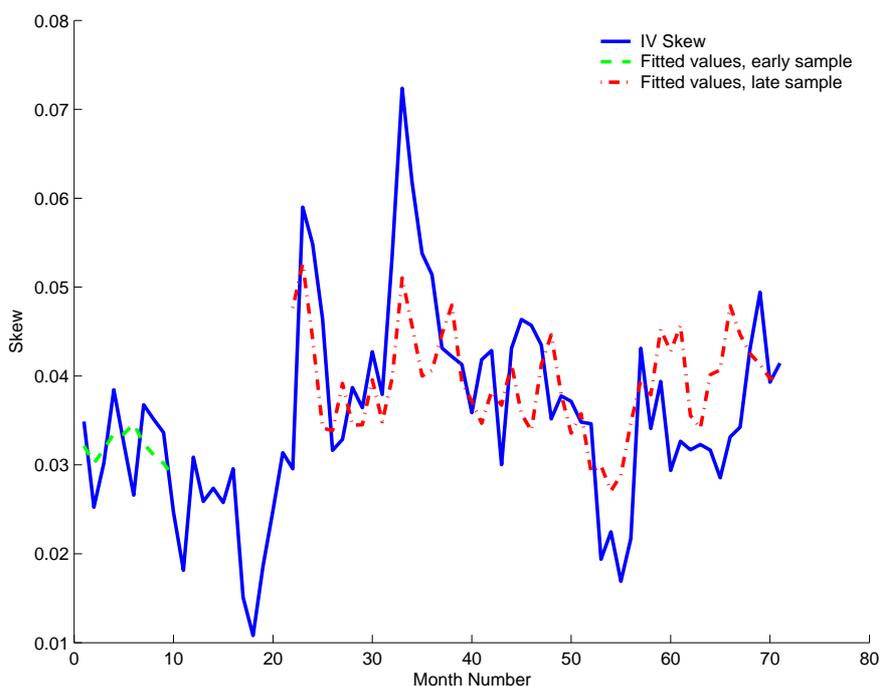


Figure 7: The solid line shows the implied volatility skew for SPX options. The dashed lines are, respectively, the fitted values from the skew in demand before and after certain structural changes (1996/01–1996/10 and 1997/10–2001/12).

only from the implied volatility of puts which ensures that the implied volatilities of in-the-money options are not used in the test. Using the jump-risk model, variation between the minimum and maximum levels of the independent variable results in a skew change of about 3 percentage points. Further, a one standard deviation move in the independent variable results in a change in the dependent variable of 0.53 standard deviations.

Equity Options:

We consider next the time-series relationship for equity options between skew in demand and skew in implied volatility. In particular, we run the time-series regression (46) for each stock, and average the coefficients across stocks.²⁹ Once again, we consider separately the subsamples before and after the summer of 1999, because of the widespread move toward cross-listing in the summer of 1999.

The results are shown in Table 5. The time-series regression is run separately for stationary than the level of demand over the full sample.

²⁹Now the dependent variable is simply the time t implied volatility from low moneyness options minus that from near-the-money options.

Table 5: The relationship between implied volatility skew (i.e., implied volatility of low moneyness minus near-the-money options) and non-market-maker net demand skew (i.e., non-market-maker net demand for low moneyness minus high moneyness options) on 303 different underlying stocks (Equation 46). We run time-series regressions for each underlying and report the average coefficients. The number p-val is the p-value from the binomial test that the coefficients are equally likely to be positive and negative.

	Before Cross-Listing of Options 04/1996–06/1999	After Cross-Listing of Options 10/1999–12/2001
Constant	0.03	0.03
<i>NetDemand</i>	4.0E-6	4.2E-6
Average Adj. R^2 (%)	3	8
# positive	200	163
# negative	70	61
p-val	0.00	0.00

the same 303 underlying stocks as above. We compute the implied volatility skew as the average implied volatility of selected low moneyness options minus that of near-the-money options.³⁰ The skew in demand is the total net non-market-maker demand for options with moneyness between 0.80 and 1.00 minus that for options with moneyness between 1.00 and 1.20.³¹ We run the regression using monthly data on underlying stocks that have at least 12 months of data available.

The average coefficient b measuring the effect of skew in demand on skew in expensiveness is positive and significant in both subsamples. In the first subsample, the coefficient estimate is positive for 200 underlying stocks and negative for only 70.³² In the second subsample, the coefficient estimate is positive for 163 underlying stocks and negative for only 61.³³ These results imply that when the demand for equity options is more skewed, their implied volatility is more skewed. If the tests are run with the dependent variable computed only from the implied volatility of puts, so that implied volatilities of in-the-money options are not used, similar results obtain.

³⁰That is, the average implied volatility from options with moneyness between 0.85 and 0.95 minus the average implied volatility from option with moneyness between 0.97 and 1.03. These options must have maturity between 15 and 45 calendar days, at least 1 contract of trading volume, and implied volatilities available on OptionMetrics.

³¹Options with maturity of 10 to 180 calendar day are included. The results are robust to variation in the definition in these demand categories.

³²Of these coefficients, 64, respectively 9, are significant at a 5% level.

³³Of these coefficients, 82, respectively 7, are significant at a 5% level.

5 Conclusion

Relative to the Black-Scholes-Merton benchmark, index and equity options display a number of robust pricing anomalies. A large body of research has attempted to address these anomalies, in large part by generalizing the Black-Scholes-Merton assumptions about the dynamics of the underlying asset. While these efforts have met with some success, non-trivial pricing puzzles remain. Further, it is not clear that this approach can yield a satisfactory description of option prices. For example, index and equity option prices display very different properties, although the dynamics of their underlying assets are quite similar.

This paper takes a different approach to option pricing. We recognize that, in contrast to the Black-Scholes-Merton framework, in the real world options cannot be hedged perfectly. Consequently, if intermediaries who take the other side of end-user option demand are risk-averse, end-user demand for options will impact option prices.

The theoretical part of the paper develops a model of competitive risk-averse intermediaries who cannot perfectly hedge their option positions. We compute equilibrium prices as a function of net end-user demand and show that demand for an option increases its price by an amount proportional to the variance of the unhedgeable part of the option and that it changes the prices of other options on the same underlying asset by an amount proportional to the covariance of their unhedgeable parts.

The empirical part of the paper measures the expensiveness of an option as its Black-Scholes implied volatility minus a proxy for the expected volatility over the life of the option. We show that on average index options are quite expensive by this measure, and that they have high positive end-user demand. Equity options, on the other hand, are on average slightly inexpensive and have a small negative end-user demand. In accordance with the predictions of our theory, we find that options are overall more expensive when there is more end-user demand for options and that the expensiveness skew across moneyness is positively related to the skew in end-user demand across moneyness. Finally, demand effects are stronger following recent market maker losses compared to times of recent market maker gains.

A Proofs

Proof of Lemma 1:

We start by imposing the following technical conditions. First d and X have compact supports. Second, for any continuous function f ,

$$\mathbb{E} [f(y_{t+1}, R_{t+1}^e, X_{t+1}) | X_t = x] \quad (47)$$

is continuous in x . This Feller-like is a weak requirement on the dependence of $(y_{t+1}, R_{t+1}^e, X_{t+1})$ on X_t , which should not be too irregular. Finally, we require that R_{t+1}^e be bounded, to so that all expectations are well defined for sure.

The Bellman equation is

$$\begin{aligned} J(W_t; t, X_t) &= -\frac{1}{k} e^{-k(W_t + G_t(d_t, X_t))} \\ &= \max_{C_t, q_t, \theta_t} \left\{ -\frac{1}{\gamma} e^{-\gamma C_t} + \rho \mathbb{E}_t [J(W_{t+1}; t+1, X_{t+1})] \right\} \end{aligned} \quad (48)$$

Given the strict concavity of the utility function, the maximum is characterized by the first-order conditions (FOC's). Using the proposed functional form for the value function, the FOC for C_t is

$$0 = e^{-\gamma C_t} + k R_f \rho \mathbb{E}_t [J(W_{t+1}; t+1, X_{t+1})], \quad (49)$$

which, together with (48), yields

$$0 = e^{-\gamma C_t} + k R_f \left[J(W_t; t, X_t) + \frac{1}{\gamma} e^{-\gamma C_t} \right], \quad (50)$$

that is,

$$e^{-\gamma C_t} = e^{-k(W_t + G_t(d_t, X_t))}, \quad (51)$$

implying (6). The FOC's for θ_t and q_t are (7) and (8). We derive G recursively as follows. First, we let $G(t+1, \cdot)$ be given. Then, θ_t and q_t are given as the unique solutions to Equations (7) and (8). Clearly, θ_t and q_t do not depend on the wealth W_t . Further, (50) implies that

$$0 = e^{-\gamma C_t} - R_f \rho \mathbb{E}_t \left[e^{-k((W_t - C_t) R_f + q_t (p_{t+1} - R_f p_t) + \theta_t R_{t+1}^e + G_{t+1}(d_{t+1}, X_{t+1}))} \right], \quad (52)$$

that is,

$$e^{-\gamma C_t - k R_f C_t + k R_f W_t} = R_f \rho \mathbb{E}_t \left[e^{-k(q_t (p_{t+1} - R_f p_t) + \theta_t R_{t+1}^e + G_{t+1}(d_{t+1}, X_{t+1}))} \right], \quad (53)$$

which, using (6), yields the equation that defines $G_t(d_t, X_t)$ (since X_t is Markov):

$$e^{-kR_f G_t(d_t, X_t)} = R_f \rho E_t \left[e^{-k(q_t(p_{t+1} - R_f p_t) + \theta_t R_{t+1}^e + G_{t+1}(d_{t+1}, X_{t+1}))} \right] \quad (54)$$

At $t = \bar{T}$, we want to show the existence of a stationary solution. First note that the operator A defined by

$$AF(w; x) = \max_{C, \theta} \left\{ -\frac{1}{\gamma} e^{-\gamma C} + \rho E_t [F(W_{t+1}, X_{t+1}) | W_t = w, X_t = x] \right\} \quad (55)$$

subject to

$$W_{t+1} = (W_t - C)R_f + \theta R_{t+1}^e, \quad (56)$$

satisfies the conditions of Blackwell's Theorem,³⁴ and is therefore a contraction.

Furthermore, A maps any function of the type

$$F(w; x) = -\frac{1}{k} e^{-kw} g(x) \quad (57)$$

into a function of the same type, implying that the restriction of A to g , denoted also by A , is a contraction as well.

We are looking for a non-zero fixed point g to

$$g(X_t) = \rho R_f E [g(X_{t+1})^{R_f} e^{-k\theta^* R_{t+1}^e} | X_t].$$

It suffices to show that there exists $m > 0$ such that A maps the set

$$G^m = \{g : \mathbb{X} \rightarrow \mathbb{R} : g \text{ is continuous, } g \geq m\} \quad (58)$$

into itself, where $\mathbb{X} \subset \mathbb{R}^n$ is the support of X_t , since then A has a (unique) fixed point in the complete set G^m .

Continuity holds by assumption (the Feller property). Let us look for $m > 0$. Since

$$Ag \geq \inf_x \inf_{\theta} R_f \rho E \left[e^{-k(\theta R_{t+1}^e)} g(X_{t+1})^{\frac{1}{R_f}} | X_t = x \right] \quad (59)$$

$$\geq \inf_x \inf_{\theta} R_f \rho E \left[e^{-k(\theta R_{t+1}^e)} | X_t = x \right] \left(\min_z g(z) \right)^{\frac{1}{R_f}} \quad (60)$$

$$\geq B \left(\min_z g(z) \right)^{\frac{1}{R_f}} \quad (61)$$

for a constant $B > 0$ (the inner infimum is a strictly positive, continuous function of $x \in \mathbb{X}$ compact), the assertion holds for any m not bigger than $B^{\frac{R_f}{R_f-1}}$.

³⁴See, for instance, Stokey, Lucas, and Prescott (1989).

It remains to prove that, given our candidate consumption and investment policy,

$$\lim_{t \rightarrow \infty} \mathbb{E} [\rho^{-t} e^{-kW_t}] = 0. \quad (62)$$

Start by noting that, for $t > T$,

$$W_{t+1} = W_t - (R_f - 1)G(X_t) + \theta_t R_{t+1}^e, \quad (63)$$

implying, by a repeated application of the iterated expectations, that

$$\mathbb{E}_T [e^{-k(W_t + G(X_t))}] = e^{-k(W_T + G(X_T))}, \quad (64)$$

which is bounded. Since $G(X_t)$ is bounded, it follows that $\lim_{t \rightarrow \infty} \mathbb{E} [\rho^{-t} e^{-kW_t}] = 0$.

The verification argument is standard, and particularly easy in this case given the boundedness of g . □

Proof of Proposition 1:

Given a position process from date t onwards and a price process from date $t + 1$ onward, the price at time t is determined by (8). It is immediate that p_t is measurable with respect to time- t information. □

Proof of Proposition 2:

Part (i) is immediate, since prices are linear. Part (ii) follows because, for any $a \in \mathbb{R}$, the pricing kernel is kept exactly the same by the offsetting change in (q, θ) . □

Proof of Proposition 3:

Part a) is immediately since a variance is always positive. The proof of b) is based on the following result, which is proved below.

Lemma 2 *Given h_1 and h_2 convex functions on \mathbb{R} , $\forall \beta < 0, \alpha, \gamma \in \mathbb{R}, \exists \alpha', \gamma' \in \mathbb{R}$ such that*

$$|h_1(x) - \alpha'x - \gamma'| \leq |h_1(x) - \alpha x - \beta h_2(x) - \gamma| \quad (65)$$

$\forall x \in \mathbb{R}$. Consequently, under any distribution, regressing h_1 on h_2 and the identity function results in a positive coefficient on h_2 .

Letting $\tilde{p}_{t+1} = p_{t+1} - \mathbb{E}_t^d [p_{t+1}]$ and suppressing subscripts, consider the expression

$$\Psi = \mathbb{E}^d [\tilde{p}^i \tilde{p}^j] \text{Var}(R^e) - \mathbb{E}^d [\tilde{p}^i R^e] \mathbb{E}^d [\tilde{p}^j R^e], \quad (66)$$

which we want to show to be positive. Letting $\hat{p}^i = E^d[\tilde{p}^i | S]$ and $\hat{p}^j = E^d[\tilde{p}^j | S]$, we write

$$\begin{aligned}\Psi &= E^d [\text{Cov}(\tilde{p}^i, \tilde{p}^j | S) \text{Var}(R^e) + \hat{p}^i \hat{p}^j \text{Var}(R^e) - E^d[\hat{p}^i R^e] E^d[\hat{p}^j R^e]] \\ &= E^d [\text{Cov}(\tilde{p}^i, \tilde{p}^j | S) \text{Var}(R^e)] + E^d [\hat{p}^i \hat{p}^j \text{Var}(R^e) - E^d[\hat{p}^i R^e] E^d[\hat{p}^j R^e]].\end{aligned}\quad (67)$$

The first term is positive by assumption, while the second is positive because \hat{p}^i and \hat{p}^j are convex and then using Lemma 2 . □

Proof of Lemma 2: We consider three cases, following whether the intersection of the graphs of h_1 and $g \equiv \gamma + \alpha Id + \beta h_2$ (i) is empty; (ii) is a singleton; (iii) contains more than one point.

(i) Since the graphs do not intersect, there exists a hyperplane that separates the convex sets $\{(x, y) : h_1(x) \leq y\}$ and $\{(x, y) : g(x) \geq y\}$.

(ii) The same is true if the two graphs have a tangency point. If they intersect in one point, \hat{x} , and are not tangent, then assume that, for $x < \hat{x}$, $h_1(x) > g(x)$ (a similar argument settles the complementary case). The convex set generated by $\{(\hat{x}, h_1(\hat{x}))\} \vee \{(x, h_1(x)) : x < \hat{x}\} \vee \{(x, g(x)) : x > \hat{x}\}$ and the one generated by $\{(\hat{x}, h_1(\hat{x}))\} \vee \{(x, h_1(x)) : x > \hat{x}\} \vee \{(x, g(x)) : x < \hat{x}\}$ have only one point in common, and therefore can be separated with a hyperplane.

(iii) Consider the line generated by the intersection. If there existed a point x at which the ordinate of the line was higher than $h_1(x)$ and than $g(x)$, then it would follow that at least one of the intersection points is actually interior to $\{(x, y) : h_1(x) \leq y\}$, which would be a contradiction. Similarly if the line was too low. □

Proof of Theorem 3:

We compute the sensitivity of current prices to a deviation in future positions from 0 in the direction of demand $\tilde{d}_s = \epsilon_s d_s$ at time s by differentiating with respect to $\epsilon_s = \epsilon$ (evaluated at $\epsilon = 0$). We then aggregate the demands at all times to compute the total effect:

$$\frac{\partial p_t}{\partial \epsilon} = \sum_{s \geq 0} \frac{\partial p_t}{\partial \epsilon_s} \frac{\partial \epsilon_s}{\partial \epsilon} = \sum_{s \geq 0} \frac{\partial p_t}{\partial \epsilon_s} \quad (68)$$

To compute the price effect of expected demand at any time s , we note that it follows from the dealer's problem that

$$p_t = E_t [\rho^{s-t} e^{-\gamma(C_s - C_t)} p_s] \quad (69)$$

which implies

$$\begin{aligned}
\frac{\partial p_t}{\partial \epsilon_s} &= \mathbb{E}_t \left[\rho^{s-t} e^{-\gamma(C_s - C_t)} \frac{\partial p_s}{\partial \epsilon_s} \right] \\
&= R_f^{-(s-t)} \mathbb{E}_t^* \left[\frac{\partial p_s}{\partial \epsilon_s} \right] \\
&= R_f^{-(s-t)} \mathbb{E}_t^* \left[\frac{\partial p_s}{\partial q_s^j} q_s^j \right].
\end{aligned} \tag{70}$$

where we use that $\frac{\partial C_t}{\partial \epsilon_s} = \frac{\partial C_s}{\partial \epsilon_s} = 0$ at $q = 0$. The equality $\frac{\partial C_s}{\partial \epsilon_s} = 0$ follows from

$$\frac{\partial C_s}{\partial q_s^j} = \frac{k}{\gamma} \frac{\partial G(s, X_s; q)}{\partial q_s^j} = -\frac{k^2 R_f \rho}{\gamma} \mathbb{E}_s^* \left[p_{s+1}^j - R_f p_s^j + \frac{\partial \theta_s}{\partial q_s^j} R_{s+1}^e \right] = 0 \tag{71}$$

and the other equality follows from differentiating the condition that marginal rates of substitution are equal

$$e^{-\gamma C_t} = e^{-\rho(s-t)} \mathbb{E}_t [e^{-\gamma C_s}], \tag{72}$$

which gives that

$$e^{-\gamma C_t} \frac{\partial C_t}{\partial \epsilon_s} = e^{-\rho(s-t)} \mathbb{E}_u \left[e^{-\gamma C_s} \frac{\partial C_s}{\partial \epsilon_s} \right] = 0 \tag{73}$$

It remains to show that the price is a smooth (C^∞) function of ϵ . Consider a process d_t characterized by $d_t = 0$ for all $t > T$, and let the demand process be given by $\tilde{d}_t = \epsilon_t d_t = -q_t$. At time t , the following optimality conditions must hold:

$$e^{-\gamma C_t} = (R_f \rho)^{T+1-t} \mathbb{E}_t [e^{-k(W_{T+1} + G_{T+1})}] \tag{74}$$

$$0 = \mathbb{E}_t [e^{-k(W_{T+1} + G_{T+1})} R_{t+1}^e] \tag{75}$$

$$0 = \mathbb{E}_t [e^{-k(W_{T+1} + G_{T+1})} (p_{T+1} - R_f^{T+1-t} p_t)], \tag{76}$$

with

$$W_{T+1} = (W_t - C_t) R_f^{T+1-t} + \sum_{s=t}^T (y_{s+1} - C_{s+1} + \theta_s R_{s+1}^e - \epsilon_s d_s (p_{s+1} - R_f p_s)) R_f^{T+1-s} + y_{T+1}.$$

We use the notation p_{T+1} for the time $T + 1$ -money value of the payoff of options expired by $T + 1$.

We show by induction that, given $X_s, (p_s, \theta_s, C_s)$ is a smooth function of $(\epsilon_s, \dots, \epsilon_{s+1})$. Note that the statement holds trivially for $s > T$.

Assume therefore the statement for all $s > t$. There are $n_t + 2$ equations in (74)–(76), with n_{t+1} being the number of derivatives priced at time t . Note that the equations

do not depend on ϵ_s for $s < t$ and that they are smooth in ϵ_s for all s , as well as in θ_t , C_t , and p_t . In order to prove the claim, we have to show that the derivative of the functions giving (74)–(76) with respect to (p_t, θ_t, C_t) is invertible (at $\epsilon = 0$ suffices), i.e., it has non-zero determinant. The implicit function theorem, then, proves the induction statement for $s = t$.

The non-zero determinant is shown as follows. If we let F_C , F_θ , respectively F_q denote the functions implicit in equations (74)–(76), it follows easily that

$$\begin{aligned}
D_C F_C &\neq 0 \\
D_C F_\theta &= 0 \\
D_C F_q &= 0 \\
D_\theta F_\theta &\neq 0 \\
D_p F_\theta &= 0 \\
\det(D_p F_q) &\neq 0,
\end{aligned} \tag{77}$$

implying that $D_{(C_t, \theta_t, p_t)} F$ has non-zero determinant. \square

Proof of Proposition 4:

Consider an optimally hedged short put position with strike price $K < R_f S_t$. With $x = S_{t+1} - R_f S_t$, the payoff from this position is

$$\Pi(x) = -d(K - R_f S_t - x)^+ + \theta x. \tag{78}$$

The optimality of the hedge means that, under the risk-neutral measure,

$$\mathbb{E}^d [e^{-k\Pi(x)} x] = 0. \tag{79}$$

Note that, since $K < R_f S_t$, $\Pi(x) < 0$ for $x > 0$ and $\Pi(x) > 0$ for $K - R_f S_t < x < 0$. Consequently, given the symmetry of x around 0 and the zero-expectation condition above, with ξ denoting the density of x ,

$$\int_K^\infty (e^{-k\Pi(x)} x - e^{-k\Pi(-x)} x) \xi(x) dx = - \int_0^{K-R_f S_t} (e^{-k\Pi(x)} x - e^{-k\Pi(-x)} x) \xi(x) dx < \mathbf{(80)}$$

It immediately follows that it cannot be true that $\Pi(-x) \geq \Pi(x)$ for all $x > |K - R_f S_t|$. In other words, for some value $x > |K - R_f S_t|$, $\Pi(-x) < \Pi(x)$, which then gives $d + \theta > -\theta$, or $|\theta| < \frac{1}{2}|d|$: the payoff is more sensitive to large downward movements in the underlying than to large upward movements. Thus, there exists \bar{K} such that, for all $S_{t+1} < \bar{K}$,

$$\Pi(S_{t+1} - R_f S_t) < \Pi(-(S_{t+1} - R_f S_t)), \tag{81}$$

implying that, whenever $K' < \bar{K}$ and $K'' = 2R_f S_t - K'$,

$$p(p, K', d) > p(c, K'', d) \quad (82)$$

$$p(p, K', 0) = p(c, K'', 0), \quad (83)$$

the second relation being the result of symmetry. \square

Arbitrary Number of Time Units Δ_t between Hedging Dates:

Most of our analysis relies on the assumption that there is 1 time unit per period (which simplifies notation). For the results on frequent hedging (i.e., $\Delta_t \rightarrow 0$) in Propositions 5-7, it is useful to see how to adapt the results for an arbitrary Δ_t .

For simplicity, let the stock price be $S = X^1$, with no dividends, and let X follow a jump-diffusion with time-independent coefficient functions:

$$dX_t = \mu(X_t) dt + v(X_t) dB_t + \eta(X_t) dN_t, \quad (84)$$

where N_t is a counting process with arrival intensity $\lambda(X_t)$ and where the jumps size is drawn from a uniformly bounded distribution also depending on X_t .

Let Δ_t be arbitrary, provided that all derivatives mature at times that are integral multiples of Δ_t . The dealer's problem is to maximize

$$U(C_t, C_{t+\Delta_t}, \dots) = \mathbb{E}_t \left[\sum_{l=0}^{\infty} \rho^{l\Delta_t} u(C_{t+l\Delta_t}) \Delta_t \right], \quad (85)$$

with $u(c) = -\frac{1}{\gamma} e^{-\gamma c}$. At any time t , the dealer must choose the consumption level C_t , the dollar investment θ_t in the underlying, and the number of derivatives held $q_t = (q_t^i)_{i \in I_t}$, while satisfying the transversality condition $\lim_{t \rightarrow \infty} \mathbb{E} [\rho^{-t} e^{-kW_t}] = 0$, where the dealer's wealth evolves as

$$W_{t+\Delta_t} = (W_t - C_t \Delta_t) R_f + q_t (p_{t+\Delta_t} - r p_t) + \theta_t R_{t+\Delta_t}^e. \quad (86)$$

Here, $R_f = r \Delta_t$ for a fixed r and C is the annualized consumption.

The results of Lemma 1 become

$$J(W_t; t, X_t) = -\frac{1}{k} e^{-k(W_t + G_t(d_t, X_t))} \quad (87)$$

$$C_t = \frac{k}{\gamma} (W_t + G_t(d_t, X_t)), \quad (88)$$

with $k = \frac{R_f - 1}{R_f \Delta_t} \gamma = \frac{e^{r\Delta_t} - 1}{e^{r\Delta_t} \Delta_t} \gamma$. The stock and derivative holdings are characterized by the first-order conditions

$$0 = \mathbb{E}_t \left[e^{-k(\theta_t R_{t+\Delta_t}^e + q_t(p_{t+\Delta_t} - r p_t) + G_{t+\Delta_t}(d_{t+\Delta_t}, X_{t+\Delta_t}))} R_{t+\Delta_t}^e \right] \quad (89)$$

$$0 = \mathbb{E}_t \left[e^{-k(\theta_t R_{t+\Delta_t}^e + q_t(p_{t+\Delta_t} - r p_t) + G_{t+\Delta_t}(d_{t+\Delta_t}, X_{t+\Delta_t}))} (p_{t+\Delta_t} - r p_t) \right], \quad (90)$$

where, for $t \leq T$, the function $G_t(d_t, X_t)$ is derived recursively using (89), (90), and

$$e^{-kR_f G_t(d_t, X_t)} = R_f \rho \mathbf{E}_t \left[e^{-k(p_{t+\Delta_t} - R_f p_t) + \theta_t R_{t+\Delta_t}^e + G_{t+\Delta_t}(d_{t+\Delta_t}, X_{t+\Delta_t})} \right], \quad (91)$$

and, for $t > T$, the function $G_t(d_t, X_t) = \bar{G}(X_t)$ where $(\bar{G}(X_t), \bar{\theta}(X_t))$ solves

$$\begin{aligned} e^{-kR_f \bar{G}(X_t)} &= R_f \rho \mathbf{E}_t \left[e^{-k(\bar{\theta}_t R_{t+\Delta_t}^e + \bar{G}(X_{t+\Delta_t}))} \right] \\ 0 &= \mathbf{E}_t \left[e^{-k(\bar{\theta}_t R_{t+\Delta_t}^e + \bar{G}(X_{t+\Delta_t}))} R_{t+\Delta_t}^e \right]. \end{aligned} \quad (92)$$

Replicating calculations in the body of the paper, with the obvious modification, the statement of Theorem 3 is generalized to

$$\begin{aligned} p_t &= p_t(0) + \gamma \frac{R_f - 1}{\Delta_t} \left[E_t^0 (\bar{p}_{t+\Delta_t} \bar{p}'_{t+\Delta_t}) dt + \sum_{l>0} e^{-rl\Delta_t} E_t^0 (\bar{p}_{t+(l+1)\Delta_t} \bar{p}'_{t+(l+1)\Delta_t}) d_{t+l\Delta_t} \right] \epsilon \\ &+ O(\epsilon^2). \end{aligned} \quad (93)$$

Proofs of Propositions 5-7:

In order to prove Propositions 5-7, we proceed with a few technical preliminaries. We assume that under the risk-neutral pricing measure given by $d = 0$ the process X_t is a jump-diffusion of the same functional form as given by Equation (84). For zero demand, an option price is defined by

$$p(S_t, t; X_t^{(1)}) = e^{-r(T-t)} \mathbf{E}^0 [(S_T - K)^+ | X_t], \quad (94)$$

regardless of the frequency of trading, where $X^{(1)} = (X^2, \dots, X^n)$.

For the computation of the covariances necessary, we shall be relying on applications of the following result. For sufficiently smooth functions $g, h : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$d(g(X_t, t)h(X_t, t)) = m_t^{gh} dt + dM_t^{gh}, \quad (95)$$

where M^{gh} is a martingale and

$$\begin{aligned} m_t^{gh} &= \left(g_X \mu + g_t + \frac{1}{2} \text{tr} (g_{XX} v v^\top) \right) h + \left(h_X \mu + h_t + \frac{1}{2} \text{tr} (h_{XX} v v^\top) \right) g + \\ &g_X v v^\top h_X^\top - \lambda (\mathbf{E}_t^0 [g(X_t + \eta)h(X_t + \eta)] - g(X_t)h(X_t)). \end{aligned} \quad (96)$$

It follows that

$$\mathbf{E}_t^0 [g(X_{t+\Delta_t}, t + \Delta_t)h(X_{t+\Delta_t}, t + \Delta_t)] = g(X_t, t)h(X_t, t) + m_t^{gh} \Delta_t + O(\Delta_t^2). \quad (97)$$

We now apply this approximation result to compute the unhedgeable price covariances. We start with the hedge ratio θ , which requires estimating

$$\begin{aligned} \mathbb{E}_t^0 [p_{t+\Delta_t} R_{t+\Delta_t}^e] &= \frac{1}{S_t} (v^1 \cdot v^\top p_X^\top + \lambda p \mathbb{E}_t^0 [\eta] - \lambda (\mathbb{E}_t^0 [p(X_t + \eta) \eta])) \Delta_t + O(\Delta_t^2) \quad (98) \\ &= \frac{1}{S_t} (v^1 \cdot v^\top p_X^\top - \lambda (\mathbb{E}_t^0 [(p(X_t + \eta) - p(X_t)) \eta])) \Delta_t + O(\Delta_t^2) \end{aligned}$$

and

$$\mathbb{E}_t^0 [(R_{t+\Delta_t}^e)^2] = \frac{1}{S_t^2} (v^1 \cdot (v^1)^\top - \lambda \mathbb{E}_t^0 [\eta^2]) \Delta_t + O(\Delta_t^2), \quad (99)$$

where v^1 denotes the first row of the matrix v . This gives θ up to order $O(\Delta_t)$.

Finally, compute the un-hedged payoff covariance using θ and applying (97) with appropriate choices for g and h :

$$\text{Cov}_t^0 [\bar{p}_{t+\Delta_t}^i, \bar{p}_{t+\Delta_t}^j] = \left(m_t^{p^i p^j} - \theta_t^i m_t^{p^j R^e} - \theta_t^j m_t^{p^i R^e} + \theta_t^i \theta_t^j m_t^{(R^e)^2} \right) \Delta_t + O(\Delta_t^2), \quad (100)$$

where the only quantity not computed above is

$$m_t^{p^i p^j} = 2r p^i p^j + p_X^j v (p_X^i v)^\top - \lambda \mathbb{E}_t^0 [(p^i(X_t + \eta) - p^i(X_t)) (p^j(X_t + \eta) - p^j(X_t))] \quad (101)$$

Here we used the fact that, since $e^{-rs} p(X_s, s)$ is a martingale, the fundamental pricing PDE holds, i.e., the drift of $p(X_s, s)$ equals $rp(X_s, s)$.

We now specialize the model to three choices, corresponding to each of Sections 2.2–2.4:

Proof of Proposition 5:

In the first model $S = X^1$ is a Markov diffusion, i.e., the jump component is 0. We obtain immediately that $\theta_t = S_t p_S(S_t, t) + O(\Delta_t) = S_t f_S(S_t, t) + O(\Delta_t)$, and therefore $\text{Cov}_t^0 [\bar{p}_{t+\Delta_t}^i, \bar{p}_{t+\Delta_t}^j] = 0$ up to terms in $O(\Delta_t^2)$.

To obtain a more nuanced answer we work directly with an exact third-order Taylor expansion of the function $p(S_{t+\Delta_t}, t + \Delta_t)$ around the point (S_t, t) . It follows that

$$\begin{aligned} &\text{Cov}_t^0 [\bar{p}_{t+\Delta_t}^i, \bar{p}_{t+\Delta_t}^j] \\ &= \text{Cov}_t^0 \left[\frac{1}{2} f_{SS}^i \Delta S^2 + \Delta S O(\Delta_t) + O(\Delta_t^2), \frac{1}{2} f_{SS}^j \Delta S^2 + \Delta S O(\Delta_t) + O(\Delta_t^2) \right] \\ &= \frac{1}{4} f_{SS}^i f_{SS}^j \text{Var}_t^0 [\Delta S^2] + O\left(\Delta_t^{\frac{5}{2}}\right) \quad (102) \\ &= \frac{1}{2} f_{SS}^i f_{SS}^j v^4 \Delta_t^2 + O\left(\Delta_t^{\frac{5}{2}}\right). \end{aligned}$$

□

Proof of Proposition 6:

In this model, S is a Markov jump-diffusion. Let $\bar{\eta} = S^{-1}\eta$ be the relative jump size and $\sigma = S^{-1}v$ be the relative volatility, giving

$$\theta_t = \frac{\sigma^2 f_S S_t - \lambda \mathbb{E}_t^0 [(p(S_t(1 + \bar{\eta})) - p(S_t)) \bar{\eta}]}{\sigma^2 - \lambda \mathbb{E}_t^0 [\bar{\eta}^2]} + O(\Delta_t) \quad (103)$$

and

$$\text{Cov}_t^0 [\bar{p}_{t+\Delta_t}^i, \bar{p}_{t+\Delta_t}^j] = ((f_S^i S_t - \theta^i)(f_S^j S_t - \theta^j) \sigma^2 + \lambda \mathbb{E}_t^0 (\kappa^i \kappa^j)) \Delta_t + O(\Delta_t^2), \quad (104)$$

with $\kappa^i = f^i(S_t(1 + \bar{\eta})) - f^i - \theta^i \bar{\eta}$.

□

Proof of Proposition 7:

Here we let $X = (S, \sigma)^\top$ with

$$dS_t = rS_t dt + \sigma_t S_t dB_{1,t} \quad (105)$$

$$d\sigma_t = \phi(\bar{\sigma} - \sigma_t) dt + \sigma_\sigma(S_t, \sigma_t) dB_{2,t}. \quad (106)$$

Since the Brownian motion driving the volatility, $B_{2,t}$ is independent from the one driving the underlying return, $B_{1,t}$, the hedge ratio is the same as in the 1-dimensional Markov-diffusion case,

$$\theta_t = f_S S_t + O(\Delta_t) \quad (107)$$

$$\text{Cov}_t^0 [\bar{p}_{t+\Delta_t}^i, \bar{p}_{t+\Delta_t}^j] = f_\sigma^i f_\sigma^j \sigma_\sigma^2 \Delta_t + O(\Delta_t^2), \quad (108)$$

where f_σ is the derivative of f with respect to the second argument. □

Empirical Implementation:

For the empirical implementation of the model, we do as follows. We weight the demand of the included options according to the model-implied covariances of unhedgeable parts, using the models of Sections 2.2–2.4. Specifically, for the diffusive risk we compute an option's Black-Scholes-Merton gamma (i.e. the second derivative with respect to the price of the underlying) evaluated at the option's moneyness and maturity, 20% volatility, $S = 1$, an interest rate of $r = 5\%$, dividend rate $q = 1.5\%$, and $\Delta_t = 0.01$. The results are robust to the choices of these parameters. We keep S , r , q , and σ constant throughout the sample to avoid biases arising from changes of these variables. Hence, changes in the model-implied demand effect is solely due to changes in demand patterns.

Similarly, we compute the demand effect with stochastic volatility risk by computing the Black-Scholes-Merton vega (i.e. the derivative with respect to the volatility) evaluated at the same parameters as above and adjusting for maturity. We use the maturity adjustment described in Equation (44), where the volatility mean-reversion parameter ϕ is set to 6 based on the estimate of Pan (2002).

We compute the covariance jump risk $E(\kappa^i \kappa^j)$ by computing for each option its unhedged profit/loss in the case of equally likely jumps of 5% and -20%, jump arrival rate of 5% per year, and the parameters above. (Again, the results are robust to the choice of these parameters.) For this, we compute a jump-adjusted option price function f without demand effects and without jump-risk premia using a straightforward adaptation of Merton (1976). We use the delta hedging given by $\theta = f_S$. Finally, we compute the variables κ and their covariance using the definitions (37) and (39).

References

- AMIN, K., J. D. COVAL, AND H. N. SEYHUN (2004): "Index Option Prices and Stock Market Momentum," *Journal of Business*, 77.
- BACK, K. (1993): "Asymmetric Information and Options," *Review of Financial Studies*, 6, 435–72.
- BAKSHI, G., N. KAPADIA, AND D. MADAN (2003): "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options," *Review of Financial Studies*, 16(1), 101–143.
- BATES, D. S. (2000): "Post '87 Crash Fears in the S&P 500 Futures Options Market," *Journal of Econometrics*, 94, 181–238.
- (2003): "Empirical Option Pricing: A Retrospection," *Journal of Econometrics*, 116, 387–404.
- (2005): "Maximum Likelihood Estimation of Latent Affine Processes," *Review of Financial Studies*, forthcoming.
- BERNARDO, A. E., AND O. LEDOIT (2000): "Gain, Loss and Asset Pricing," *The Journal of Political Economy*, 108(1), 144–172.
- BLACK, F., AND M. S. SCHOLES (1973): "The Pricing of Options and Corporate Liabilities," *The Journal of Political Economy*, 81, 637–654.
- BOLLEN, N. P., AND R. E. WHALEY (2004): "Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?," *Journal of Finance*, 59, 711–753.
- BONDARENKO, O. (2003): "Why are Put Options So Expensive?," working paper, University of Illinois at Chicago.
- BRENNAN, M. (1979): "The Pricing of Contingent Claims in Discrete Time Models," *Journal of Finance*, 34(1), 53–68.
- BROADIE, M., M. CHERNOV, AND M. JOHANNES (2005): "Model Specification and Risk Premia: Evidence from Futures Options," *Working Paper, Columbia*.
- COCHRANE, J. H., AND J. SAA-REQUEJO (2000): "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets," *The Journal of Political Economy*, 108(1), 79–119.
- CONSTANTINIDES, G. M., J. C. JACKWERTH, AND S. PERRAKIS (2005): "Mispricing of S&P 500 Index Options," Working Paper, University of Chicago.

- CONSTANTINIDES, G. M., AND S. PERRAKIS (2002): “Stochastic dominance bounds on derivatives prices in a multiperiod economy with proportional transaction costs,” *Journal of Economic Dynamics and Control*, 26, 1323–1352.
- COVAL, J. D., AND T. SHUMWAY (2001): “Expected Option Returns,” *Journal of Finance*, 56(3), 983–1009.
- DE FONTNOUELLE, P., R. P. H. FISHE, AND J. H. HARRIS (2003): “The Behavior of Bid-Ask Spreads and Volume in Options Markets during the Competition for Listings in 1999,” *Journal of Finance*, 58(6), 2437–2463.
- DE ROON, F., T. E. NIJMAN, AND C. VELD (2000): “Hedging Pressure Effects in Futures Markets,” *Journal of Finance*, 55(3), 1437–1456.
- DRIESSEN, J., AND P. MAENHOUT (2003): “A Portfolio Perspective on Option Pricing Anomalies,” *Working Paper, University of Amsterdam*.
- EASLEY, D., M. O’HARA, AND P. S. SRINIVAS (1998): “Option Volume and Stock Prices: Evidence on Where Informed Traders Trade,” *Journal of Finance*, 53, 432–65.
- FIGLEWSKI, S. (1989): “Options Arbitrage in Imperfect Markets,” *Journal of Finance*, 44(5), 1289–1311.
- GABAIX, X., A. KRISHNAMURTHY, AND O. VIGNERON (2004): “Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market,” Working Paper, MIT.
- GREENWOOD, R. (2005): “Short-and long-term demand curves for stocks: theory and evidence on the dynamics of arbitrage,” *Journal of Financial Economics*, forthcoming.
- HAKANSSON, N. (1979): “The Fantastic World of Finance: Progress and the Free Lunch,” *The Journal of Financial and Quantitative Analysis*, 14, 717–734.
- HICKS, J. R. (1939): *Value and Capital*. Cambridge: Oxford University Press.
- HUGONNIER, J., D. KRAMKOV, AND W. SCHACHERMAYER (2005): “On Utility Based Pricing of Contingent Claims in Incomplete Markets,” *Mathematical Finance*, 15(2), 203–212.
- JACKWERTH, J. (2000): “Recovering risk aversion from option prices and realized returns,” *Review of Financial Studies*, 13, 433–451.
- KEYNES, J. M. (1923): “Some Aspects of Commodity Markets,” *Manchester Guardian Commercial, European Reconstruction Series*, Section 13, 784–786.

- LELAND, H. E. (1980): “Who Should Buy Portfolio Insurance?,” *Journal of Finance*, 35(2), 581–594.
- LONGSTAFF, F. A. (1995): “Option Pricing and the Martingale Restriction,” *Review of Financial Studies*, 8(4), 1091–1124.
- MERTON, R. (1973): “The Theory of Rational Option Pricing,” *Bell Journal of Economics and Management Science*, 4, 141–183.
- (1976): “Option Pricing when the Underlying Stock Returns are Discontinuous,” *Journal of Financial Economics*, 5, 125–144.
- NEWMAN, Y., AND M. RIERSON (2004): “Illiquidity Spillovers: Theory and Evidence from European Telecom Bond Issuance,” Working Paper, Stanford Graduate School of Business.
- PAN, J. (2002): “The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study,” *Journal of Financial Economics*, 63, 3–50.
- POTESHMAN, A. M. (2001): “Underreaction, Overreaction, and Increasing Misreaction to Information in the Options Market,” *Journal of Finance*, 56(3), 851–876.
- RUBINSTEIN, M. (1976): “The Valuation of Uncertain Income Streams and Pricing of Options,” *Bell Journal of Economics*, 7, 407–425.
- SHLEIFER, A. (1986): “Do Demand Curves for Stocks Slope Down?,” *Journal of Finance*, 41, 579–590.
- SHLEIFER, A., AND R. W. VISHNY (1997): “The Limits of Arbitrage,” *Journal of Finance*, 52(1), 35–55.
- SONER, H. M., S. E. SHREVE, AND J. CVITANIC (1995): “There is No Nontrivial Hedging Portfolio for Option Pricing with Transaction Costs,” *The Annals of Applied Probability*, 5(2), 327–355.
- STAPLETON, R. C., AND M. G. SUBRAHMANYAM (1984): “The Valuation of Multivariate Contingent Claims in Discrete Time Models,” *Journal of Finance*, 39(1), 207–228.
- STEIN, J. (1989): “Overreactions in the Options Market,” *Journal of Finance*, 44, 1011–1023.
- STOKEY, N. L., R. E. LUCAS, AND E. C. PRESCOTT (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.
- WURGLER, J., AND E. V. ZHURAVSKAYA (2002): “Does Arbitrage Flatten Demand Curves for Stocks?,” *Journal of Business*, 75(4), 583–608.