Demand-Based Option Pricing*

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

This version: October 5, 2005

Preliminary and Incomplete

Abstract

We model the demand-pressure effect on prices when options cannot be perfectly hedged. 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 their unhedgeable parts. Empirically, we identify aggregate positions of dealers and end users using a unique dataset, and show that demand-pressure effects help explain well-known option-pricing puzzles. First, end users are net long index options, especially out-of-money puts, which helps explain their apparent expensiveness and the smirk. Second, demand patterns help explain the prices of single-stock options.

*We are grateful for helpful comments from Apoorva Koticha and especially from Steve Figlewski, as well as from seminar participants at New York University and University of Pennsylvania

†Wharton School of Business, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104-6367, email: garleanu@wharton.upenn.edu

‡New York University 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

1 Introduction

"Skew is heavily influenced by supply and demand factors"
— *Gray, director of global equity derivatives, Dresdner Kleinwort Benson*

"The number of players in the skew market is limited. ... there's a huge imbalance between what clients want and what professionals can provide."
— *Belhadj-Soulami, head of equity derivatives trading for Europe, Paribas*

"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."
— *Bates (2003)*

We take on this challenge by providing a model of option intermediation, and by showing empirically using a unique dataset that demand pressure can help explain 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 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 because of rebalancing over discrete time intervals, stochastic volatility, and transaction costs (Figlewski (1989)).[1]

To capture this effect, we consider a model of competitive risk-averse dealers who trade at discrete times. The dealers trade many option contracts on the same underlying such that certain risks net out, while others do not. Further, dealers can hedge derivative positions by trading the underlying security. We consider a general class of distributions for the underlying, which can accommodate stochastic volatility and jumps. Dealers trade with end users. The model is agnostic about the end users' reasons for trade.

We compute equilibrium prices as functions of demand pressure, that is, the prices that make dealers optimally choose to supply the options 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

[1] Also, traders have capital constraints as emphasized, e.g., by Shleifer and Vishny (1997).

effect increases the price of the option, which entices the dealers to sell it. Specifically, a marginal change in 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. Similarly, the 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 price of other options with the same underlying, especially similar contracts.

Empirically, we use a unique dataset to identify aggregate daily positions of dealers and end users. In particular, we define dealers as marketmakers and proprietary traders and end users as customers of brokers. We find that end users have a net long position in S&P500 index options with large net positions in out-of-money puts. Hence, since options are in zero net supply, dealers are short index options. While it is conventional wisdom among option traders that Wall Street is short index volatility, this paper is the first to demonstrate this fact using data on option holdings. This can help explain the puzzle 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), Amin, Coval, and Seyhun (2003), Bondarenko (2003)). In the time series, demand for index options is related to their expensiveness, measured by the difference between their implied volatility and the volatility measure of Bates (2005). Further, the steepness of the smirk, measured by the difference between the implied volatility of low-moneyness options and at-the-money options, is positively related to the skew of option demand, measured by the demand of low-moneyness options minus the demand of high-moneyness options.

Jackwerth (2000) finds that a representative investor's option-implied utility function is inconsistent with standard assumptions in economic theory.[2] Since options are in zero net supply, a representative investor holds no options. We reconcile this finding for dealers who have significantly short index options positions. Intuitively, an investor 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 could be related to portfolio insurance and agency problems (e.g. between investors and fund managers) that are not well captured by standard utility theory.

Another option-pricing puzzle is that index option prices are so different from the prices of single-stock options despite the fact that the distributions of the underlying appear relatively similar (e.g. 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

[2] See also Driessen and Maenhout (2003).

instance, end users are net short single stock options — not long as with 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 pricing. Arbitrage bounds are trivial with any transactions 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)). Cochrane and Saa-Requejo (2000) and Bernardo and Ledoit (2000) derive tighter option-pricing bounds by restricting the Sharpe ratio or gain/loss ratio to be below an arbitrary level, and Constantinides and Perrakis (2002) derive bounds using stochastic dominance for small option positions. 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 in one options on the price 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 or not. See Rubinstein (1976), Brennan (1979), Stapleton and Subrahmanyam (1984), Hugonnier, Kramkov, and Schachermayer (2004) 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, since our model is general and can in principle be applied to any market, our work is related to the broader literature on demand pressure effects. Consistent with our model's predictions, Wurgler and Zhuravskaya (2002) 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 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.[3] We complement Bollen and

[3]Indeed, Bollen and Whaley (2004) find that a nontrivial part of the option price impact from day

Whaley (2004) by providing a theoretical model and by investigating empirically the relationship between the level of end user demand for options and the level and overall shape of implied volatility curves. 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.

The rest of the paper is organized as follows. Section 2 describes the model, and Section 3 derives its pricing implications. Section 4 provides descriptive statistics on demand patterns for options, Section 5 tests of the effect of demand pressure on option prices, and Section 6 concludes. The appendix contains proofs.

2 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 $r-1$ per period, and a risky security that we refer to as the "underlying" security. At time t, the underlying has an exogenous price[4] of S_t, dividend D_t, and an excess return of $R^e_t = (S_t + D_t)/S_{t-1} - r$ and the distribution of future prices and returns is characterized by a stationary Markov state variable $X_t \in \mathbb{X} \subset \mathbb{R}^n$, with $\mathbb{X}$ compact.[5] (The state variable could include the current level of volatility, the current jump intensity, etc.) The only condition we impose on the transition function $\pi : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_+$ of X is that it have the Feller property.

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 $p_t = (p^i_t)_{i \in I_t}$.

We assume that 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 whose prices are affected be demand pressure.

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

t signed option volume dissipates by day $t+1$.

[4]All 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.

[5]This condition can be relaxed at the expense of further technical complexity.

risk aversion, that is, his utility for remaining life-time consumption is:

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

where $u(c) = -\frac{1}{\gamma} e^{-\gamma c}$ and $\rho \in \mathbb{R}$ 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 $q_t = (q_t^i)_{i \in I_t}$, while satisfying the transversality condition $\lim_{t\to\infty} \mathrm{E}\left[\rho^{-t} e^{-kW_t}\right] = 0$, where the dealer's wealth evolves as

$$W_{t+1} \;=\; y_{t+1} + (W_t - C_t) r + q_t (p_{t+1} - r p_t) + \theta_t R_{t+1}^e,$$

$k = \gamma(r-1)/r$, and y_t is the dealer's time-t endowment. We assume that the distribution of future endowments is characterized by X_t.

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. We assume that R_t^e, D_t/S_t, y_t, and d_t are continuous functions of X_t. Furthermore, for technical reasons, we assume that, after some time $\overline{T}$, demand pressure is zero, that is, $d_t = 0$ for $t > \overline{T}$.

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 of q such that derivative markets clear $q + d = 0$.*

Our asset pricing approach relies on the insight that, by observing the aggregate quantities held by dealers, we 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 it is not crucial that end users have inelastic demand. All that matters is that end users have demand curves that result in dealers holding a position of $q = -d$.

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

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

$$\text{s.t.} \quad W_{t+1} \;=\; y_{t+1} + (W_t - C_t) r + q_t (p_{t+1} - r p_t) + \theta_t R_{t+1}^e. \tag{2}$$

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 $k = \frac{\gamma(r-1)}{r}$, then the dealer's value function and optimal consumption are given by*

$$J(W_t; t, X_t) = -\frac{1}{k} e^{-k(W_t + f_t(d_t, X_t))} \quad (3)$$

$$C_t = \frac{k}{\gamma}(W_t + f_t(d_t, X_t)) \quad (4)$$

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

$$0 = E_t\left[e^{-k\left(y_{t+1}+\theta_t R^e_{t+1}+q_t(p_{t+1}-rp_t)+f_{t+1}(d_{t+1},X_{t+1})\right)} R^e_{t+1}\right] \quad (5)$$

$$0 = E_t\left[e^{-k\left(y_{t+1}+\theta_t R^e_{t+1}+q_t(p_{t+1}-rp_t)+f_{t+1}(d_{t+1},X_{t+1})\right)} (p_{t+1} - rp_t)\right], \quad (6)$$

where, for $t \leq T$, the function $f_t(d_t, X_t)$ is derived recursively using (5), (6), and

$$e^{-krf_t(d_t,X_t)} = r\rho E_t\left[e^{-k\left(y_{t+1}+q_t(p_{t+1}-rp_t)+\theta_t R^e_{t+1}+f_{t+1}(d_{t+1},X_{t+1})\right)}\right] \quad (7)$$

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

$$e^{-kr\bar{f}(X_t)} = r\rho E_t\left[e^{-k\left(y_{t+1}+\bar{\theta}_t R^e_{t+1}+\bar{f}(X_{t+1})\right)}\right] \quad (8)$$

$$0 = E_t\left[e^{-k\left(y_{t+1}+\bar{\theta}_t R^e_{t+1}+\bar{f}(X_{t+1})\right)} R^e_{t+1}\right]. \quad (9)$$

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,

$$p^i_T = \lambda p^j_T + \alpha + \beta S_T$$

for some constants α, β, and λ. Then it is easily seen that, if positions $\left(q_t^i, q_t^j, b_t, \theta_t\right)$ in the two derivatives, the bond,[6] and the underlying, respectively, are optimal given the prices, then so are positions $\left(q_t^j + a, q_t^j - \lambda a, b_t - a\alpha r^{-(T-t)}, \theta_t - a\beta S_t^{-1}\right)$. This has the following implications for equilibrium prices:

Proposition 2 *Suppose that $D_t = 0$ and $p_T^i = \lambda p_T^j + \alpha + \beta S_T$. Then:*
(i) For any demand pressure, d, the equilibrium prices of the two derivatives are related by

$$p_t^i = \lambda p_t^j + \alpha r^{-(T-t)} + \beta S_t.$$

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

The first part of the proposition is 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 $d^j + d^i$ calls and 0 puts (or vice versa).

3 Price Effects of Demand Pressure

We now consider the main implication of the theory, namely how demand pressure affects prices. Our goal is to compute security prices p_t^i as functions of the current demand pressure d_t^j and the state variable X_t (which incorporates beliefs about future demand pressure).

We think of the price p, the hedge position θ_t in the underlying and the consumption function f as functions of d_t^j and X_t. Alternatively, we can think of the dependent variables as function 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" (6) to get

$$p_t \quad = \quad \frac{\mathrm{E}_t\left[e^{-k\left(y_{t+1}+\theta_t R_{t+1}^e + q_t p_{t+1} + f_{t+1}\right)} p_{t+1}\right]}{r\mathrm{E}_t\left[e^{-k\left(y_{t+1}+\theta_t R_{t+1}^e + q_t p_{t+1} + f_{t+1}\right)}\right]} \tag{10}$$

where the hedge position in the underlying, θ_t, solves

$$0 \quad = \quad \mathrm{E}_t\left[e^{-k\left(y_{t+1}+\theta_t R_{t+1}^e + q_t p_{t+1} + f_{t+1}\right)} R_{t+1}^e\right] \tag{11}$$

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

and where f is computed recursively as described in Lemma 1. Equations (10) and (11) can be written in terms of a demand-based pricing kernel:

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

$$p_t = E_t(m^d_{t+1} p_{t+1}) = \frac{1}{r} E^d_t(p_{t+1}) \tag{12}$$

$$0 = E_t(m^d_{t+1} R^e_{t+1}) = \frac{1}{r} E^d_t(R^e_{t+1}) \tag{13}$$

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

$$m^d_{t+1} = \frac{e^{-k\left(y_{t+1}+\theta_t R^e_{t+1}+q_t p_{t+1}+f_{t+1}\right)}}{r E_t\left[e^{-k\left(y_{t+1}+\theta_t R^e_{t+1}+q_t p_{t+1}+f_{t+1}\right)}\right]} \tag{14}$$

$$= \frac{e^{-k\left(y_{t+1}+\theta_t R^e_{t+1}-d_t p_{t+1}+f_{t+1}\right)}}{r E_t\left[e^{-k\left(y_{t+1}+\theta_t R^e_{t+1}-d_t p_{t+1}+f_{t+1}\right)}\right]}, \tag{15}$$

and E^d_t is expected value with respect to the corresponding risk-neutral measure, i.e. the measure with a Radon-Nykodim derivative of rm^d_{t+1}.

To understand this pricing kernel, suppose for instance that end users want to sell derivative i such that $d^i_t < 0$, and that this is the only demand pressure. In equilibrium, dealers take the other side of the trade, buying $q^i_t = -d^i_t > 0$ shares of this derivative, while hedging their position using a position of θ_t in the underlying. The pricing kernel is small whenever the "unhedgeable" part $q_t p_{t+1} + \theta_t R^e_{t+1}$ 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^i_t with respect to the demand pressure of another derivative d^j_t. We can initially differentiate with respect to q rather than d since $q^i = -d^i_t$.

For this, we first differentiate the pricing kernel[7]

$$\frac{\partial m^d_{t+1}}{\partial q^j_t} = -k m^d_{t+1} \left(p^j_{t+1} - r p^j_t + \frac{\partial \theta_t}{\partial q^j_t} R^e_{t+1} \right) \tag{16}$$

[7] We suppress the arguments of functions. We note that p_t, θ_t, and f_t are functions of (d_t, X_t, t), and m^d_{t+1} is a function of $(d_t, X_t, d_{t+1}, X_{t+1}, y_{t+1}, R^e_{t+1}, t)$.

using that $\frac{\partial f(t+1,X_{t+1};q)}{\partial q_t^j} = 0$ and $\frac{\partial p_{t+1}}{\partial q_t^j} = 0$. With this result, it is straightforward to differentiate (13) to get

$$0 = \mathrm{E}_t \left(m_{t+1}^d \left(p_{t+1}^j - r p_t^j + \frac{\partial \theta_t}{\partial q_t^j} R_{t+1}^e \right) R_{t+1}^e \right) \tag{17}$$

which implies that the marginal hedge position is

$$\frac{\partial \theta_t}{\partial q_t^j} = -\frac{\mathrm{E}_t \left(m_{t+1}^d \left(p_{t+1}^j - r p_t^j \right) R_{t+1}^e \right)}{\mathrm{E}_t \left(m_{t+1}^d (R_{t+1}^e)^2 \right)} = -\frac{\mathrm{Cov}_t^d (p_{t+1}^j, R_{t+1}^e)}{\mathrm{Var}_t^d (R_{t+1}^e)} \tag{18}$$

Similarly, we derive the price sensitivity by differentiating (12)

$$\begin{aligned}
\frac{\partial p_t^i}{\partial q_t^j} &= -k \mathrm{E}_t \left[m_{t+1}^d \left(p_{t+1}^j - r p_t^j + \frac{\partial \theta_t}{\partial q_t^j} R_{t+1}^e \right) p_{t+1}^i \right] && (19)\\
&= -\frac{k}{r} \mathrm{E}_t^d \left[\left(p_{t+1}^j - r p_t^j - \frac{\mathrm{Cov}_t^d (p_{t+1}^j, R_{t+1}^e)}{\mathrm{Var}_t^d (R_{t+1}^e)} R_{t+1}^e \right) p_{t+1}^i \right] && (20)\\
&= -\gamma (r-1) \mathrm{E}_t^d \left[\bar{p}_{t+1}^j \bar{p}_{t+1}^i \right] && (21)\\
&= -\gamma (r-1) \mathrm{Cov}_t^d \left[\bar{p}_{t+1}^j, \bar{p}_{t+1}^i \right] && (22)
\end{aligned}$$

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 of any security k is*

$$\bar{p}_{t+1}^k = r^{-1} \left(p_{t+1}^k - r p_t^k - \frac{Cov_t^d (p_{t+1}^k, R_{t+1}^e)}{Var_t^d (R_{t+1}^e)} R_{t+1}^e \right). \tag{23}$$

Equation (22) can also be written in terms of the demand pressure, d, by using the equilibrium relation $d = -q$:

Theorem 2 *The price sensitivity to demand pressure is*

$$\frac{\partial p_t^i}{\partial d_t^j} = \gamma (r-1) E_t^d \left(\bar{p}_{t+1}^i \bar{p}_{t+1}^j \right) = \gamma (r-1) Cov_t^d \left(\bar{p}_{t+1}^i, \bar{p}_{t+1}^j \right)$$

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 covariation of their unhedgeable parts. For European options, we can show, under the condition stated below, that a 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* $E_t^d \left[p_{t+1}^i \middle| S_{t+1} \right]$ *and* $E_t^d \left[p_{t+1}^j \middle| S_{t+1} \right]$ *are convex functions of S_{t+1} and* $Cov_t^d \left(p_{t+1}^i, p_{t+1}^j \middle| S_{t+1} \right) \geq 0$.

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. Second, we require that $\text{Cov}_t^d \left(p_{t+1}^i, p_{t+1}^j \middle| S_{t+1} \right) \geq 0$, 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 follows

$$p_t \approx p_t(d_t = 0) + \gamma(r-1)\text{E}_t^d \left(\bar{p}_{t+1} \bar{p}_{t+1}' \right) d_t \tag{24}$$
$$\tag{25}$$

Hence, the total price paid for the d_t derivatives is

$$d_t' p_t = d_t' p_t(d_t = 0) + \gamma(r-1) d_t' \text{E}_t^d \left(\bar{p}_{t+1} \bar{p}_{t+1}' \right) d_t \tag{26}$$
$$= d_t' p_t(0) + \gamma(r-1)\text{Var}_t^d \left(d_t' \bar{p}_{t+1} \right) \tag{27}$$

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. 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 1 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 be 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*

$K < rS_t$ 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 the a high-strike option.

Future demand pressure in a derivative j also affects the current price of 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 $\tilde{d}_s^j = \epsilon d_s^j$ (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*

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

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 udhedgeable 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.

3.1 Price Effect of Risk due to Discrete-Time Hedging

In this section, we derive the first-order price impact of demand pressure in the case when hedging risk arises from small changes in the price of the underlying.

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, $f^i(t, S_t) := p_t^i(S_t, d = 0)$. We suppose that, over the next time period, the price can change by a small amount, $S_{t+1} = S_t + (r-1)S_t + S_t\varepsilon$, where $E(\varepsilon) = E(\varepsilon^3) = 0$, $E(\varepsilon^2) = \sigma^2\Delta_t$, and $E(\varepsilon^4) = O(\Delta_t^2)$. Hence, the change in the option price evolves approximately according to

$$p_{t+1}^i \cong f^i + f_S^i \Delta S + \frac{1}{2} f_{SS}^i (\Delta S)^2 + f_t^i \Delta_t \tag{28}$$

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

$$\begin{aligned} r\bar{p}_{t+1}^i &= p_{t+1}^i - rp_t^i - f_S^i(S_{t+1} - rS_t) & (29) \\ &\cong -(r-1)f^i + f_t^i \Delta_t + (r-1) f_S^i S_t + \frac{1}{2} f_{SS}^i (\Delta S)^2 & (30) \end{aligned}$$

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:

$$\mathrm{Cov}_t(r\bar{p}_{t+1}^i r\bar{p}_{t+1}^j) \cong \frac{1}{4} f_{SS}^i f_{SS}^j Var_t((\Delta S)^2)$$

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

Proposition 5 *With unhedgeable risk due to discrete trading over time periods Δ_t, the first-order effect on price of demand at $d = 0$ is*

$$\frac{\partial p_t^i}{\partial d_t^j} = \frac{\gamma(r-1)Var_t((\Delta S)^2)}{4r^2} f_{SS}^i f_{SS}^j + o(\Delta_t^2) \tag{31}$$

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

$$\frac{\partial \hat{\sigma}_t^i}{\partial d_t^j} = \frac{\gamma(r-1)Var_t((\Delta S)^2)}{4r^2} \frac{f_{SS}^i}{\nu^i} f_{SS}^j + o(\Delta_t^2) \tag{32}$$

where ν^i the Black-Scholes vega.

Interestingly, the Black-Scholes gamma over 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 $Var_t((\Delta S)^2)$ namely Δ_t^2. As we show next, the risks of jumps and stochastic volatility are more important for small Δ_t (specifically, they are of order Δ_t).

3.2 Jumps in the Underlying

To study the effect of jumps in the underlying, we suppose that $S_{t+1} = (r - \mu\pi\Delta_t)S_t + S_t\varepsilon + S_t(\eta - \varepsilon)1_{(jump)}$ where ε is the "local noise" as above, η is the "jump size" with mean $E_t(\eta) = \mu$, a jump happens with probability $E(1_{(jump)}) = \pi\Delta_t$, and ε, η, and $1_{(jump)}$ are independent.

The unhedgeable price change is

$$r\bar{p}^i_{t+1} \cong -(r-1)f^i + f^i_t\Delta_t + (r-1)f^i_S S_t + \frac{1}{2}f^i_{SS}S_t^2\varepsilon^2 + (f^i_S S_t - \theta^i)(\varepsilon - \mu\pi\Delta_t) + \kappa^i 1_{(jump)}$$

where

$$\kappa^i \;=\; f^i(S_t(r - \mu\pi\Delta_t + \eta)) - f^i - \theta^i\eta. \tag{33}$$

is the unhedgeable risk due to jumps.

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

$$\frac{\partial p^i_t}{\partial d^j_t} = \frac{\gamma(r-1)}{r^2}\left[(f^i_S S_t - \theta^i)(f^j_S S_t - \theta^j)\sigma^2\Delta_t + \pi\Delta_t E_t\left(\kappa^i\kappa^j\right)\right] + o(\Delta_t) \tag{34}$$

and the first-order effect on Black-Scholes implied volatility $\hat{\sigma}^i_t$ is:

$$\frac{\partial \hat{\sigma}^i_t}{\partial d^j_t} = \frac{\gamma(r-1))}{r^2}\frac{(f^i_S S_t - \theta^i)(f^j_S S_t - \theta^j)\sigma^2\Delta_t + \pi\Delta_t E_t\left(\kappa^i\kappa^j\right)}{\nu^i} + o(\Delta_t) \tag{35}$$

where ν^i the Black-Scholes vega.

The terms of the form $f^i_S S_t - \theta^i$ arise because the optimal hedge θ differs from the optimal hedge without jumps, $f^i_S S_t$, 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 3.1. 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 (35) means that demand can affect the slope of the implied-volatility curve and generate a smile.[8]

3.3 Stochastic Volatility Risk

To consider stochastic volatility, we let $S_{t+1} = rS_t + S_t\sigma_t\varepsilon$ where $E(\varepsilon^2) = \Delta_t$, and $\sigma_{t+1} = \sigma_t + \phi\Delta_t(\bar{\sigma} - \sigma_t) + \Upsilon_{t+1}$, where $E(\Upsilon) = 0$, $E(\Upsilon^2) = \Delta_t V^2$, and ε and Υ are independent. The price $p^i_t = f^i(t, S_t, \sigma_t)$ has unhedgeable risk given by

$$\begin{aligned} r\bar{p}^i_{t+1} &= p^i_{t+1} - rp^i_t - \theta^i R^e_{t+1} \\ &\cong f^i + f^i_S S_t(r - 1 + \sigma_t\varepsilon) + f^i_t\Delta_t + f^i_\sigma\Delta\sigma_{t+1} - rf^i - \theta^i R^e_{t+1} \\ &= -(r-1)f^i + f^i_t\Delta_t + f^i_S S_t(r-1) + (f^i_S S_t - \theta^i)\sigma_t\varepsilon + f^i_\sigma\Delta\sigma_{t+1} \end{aligned}$$

[8]Of course, the jump risk also generate smiles without demand-pressure effects; the results is that demand can exacerbate these.

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

$$\frac{\partial p_t^i}{\partial d_t^j} = \frac{\gamma(r-1)\,Var(\Upsilon)}{r^2} f_\sigma^i f_\sigma^j + o(\Delta_t) \tag{36}$$

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

$$\frac{\partial \hat{\sigma}_t^i}{\partial d_t^j} = \frac{\gamma(r-1))\,Var(\Upsilon)}{r^2} \frac{f_\sigma^i}{\nu^i} f_\sigma^j + o(\Delta_t) \tag{37}$$

where ν^i the Black-Scholes vega.

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

$$f_\sigma^i \;\cong\; \nu^i \frac{\partial}{\partial v_t} E\left(\frac{\int_t^{t+T} v_s ds}{T} \;\Big|\; v_0 \right) \cong \nu^i \frac{1 - e^{-\phi T}}{\phi T} \tag{38}$$

Hence, if we combine (38) with (37), we see that stochastic volatility risk affects the level, but not the slope, of the implied volatility curves to the first order.

4 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.[9] 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.[10] The SPX

[9]Options on the S&P 500 index have many different option symbols. In this paper, *SPX options* always refers to all options that have SPX as their underlying asset, not only to those with option symbol SPX.

[10]The 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.

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.

Firm proprietary traders sometimes are end-users of options and sometimes are liquidity suppliers. Consequently, we compute net end-user demands for an option in two different ways. First, we assume that firm proprietary traders are end-users and compute the net demand for an option 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 refer to net demand computed in this way as *non-market maker net demand*. Second, we assume that the firm proprietary traders are liquidity suppliers and compute the net demand for an option as the public customer long open interest minus the public customer short open interest. We refer to net demand computed in this second way as *public customer net demand*.

Even though the SPX and individual equity option market have been the subject of extensive empirical research, there is not much 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 105,890 contracts, and the average daily public customer net demand is 138,602 contracts. In other words, the typical end-user demand for SPX options during our data period is on the order of 125,000 SPX option contracts. For puts (calls), the average daily net demand from non-market makers is 126,514 ($-20,624$) contracts, while from public customers it is 184,429 ($-45,833$) 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 30,000 call contracts.

For the equity options, the average daily non-market maker net demand per underlying stock is -2717 contracts, and the average daily public customer net demand is -4873 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, while from public customers it is -2331 (-2543) contracts.

Panel A of Table 1 reports the average daily public customer net demand for SPX options broken down by option maturity and moneyness (defined as the strike price divided by the underlying index level.) Panel A indicates that 28 percent of the non-market maker 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.,

Maturity Range (Days)	Moneyness Range (K/S) 0–0.85	0.85–0.9	0.9–0.95	0.95–1	1–1.05	1.05–1.1	1.1–1.5	1.5–2	All
Panel A: SPX Option Public Customer Net Demand									
1– 9	6,198	1,957	2,059	2,320	1,161	933	341	333	15,302
10–29	8,159	1,829	1,898	7,653	1,537	909	461	643	23,089
30–59	6,215	1,678	4,235	10,381	1,807	-1,015	685	1,062	25,049
60–89	2,701	1,848	3,503	4,486	2,307	230	206	489	15,769
90–179	7,635	4,041	4,028	4,577	2,356	2,000	235	2,913	27,784
180–364	4,334	5,374	4,744	462	2,245	1,938	209	756	20,062
365–999	1,926	2,711	3,407	2,173	1,222	283	248	-423	11,546
All	37,167	19,439	23,875	32,052	12,635	5,277	2,385	5,772	138,602
Panel B: Equity Option Public Customer Net Demand									
1–9	-73	-35	-53	-58	-61	-44	-33	-67	-423
10-29	-106	-50	-79	-107	-130	-103	-73	-162	-811
30–59	-127	-56	-72	-93	-128	-121	-94	-220	-911
60–89	-111	-50	-64	-74	-89	-81	-69	-195	-733
90–179	-233	-102	-123	-131	-155	-150	-131	-459	-1,484
180–364	-97	-48	-55	-53	-66	-58	-53	-204	-634
365–999	165	14	8	1	2	-1	0	-66	123
All	-582	-327	-438	-515	-626	-558	-455	-1,372	-4,873

Table 1: Average daily public customer 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.

moneyness < 1.) Panel B of Table 1 reports the average option net demand per underlying stock for individual equity options from public customers. With the exception of the very long maturity options (i.e, those with more than one year to expiration) the public customer net demand for all of the moneyness/maturity categories is negative. That is, public customers are net suppliers of options in all of these categories. This stands in stark contrast to the index option market in Panel A where public customers are net demanders of options in almost every moneyness/maturity category.

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 which account for the early exercise feature and which assume that investors have perfect foresight of the timing and amount of the dividends paid by the underlying stock over the life of the options.

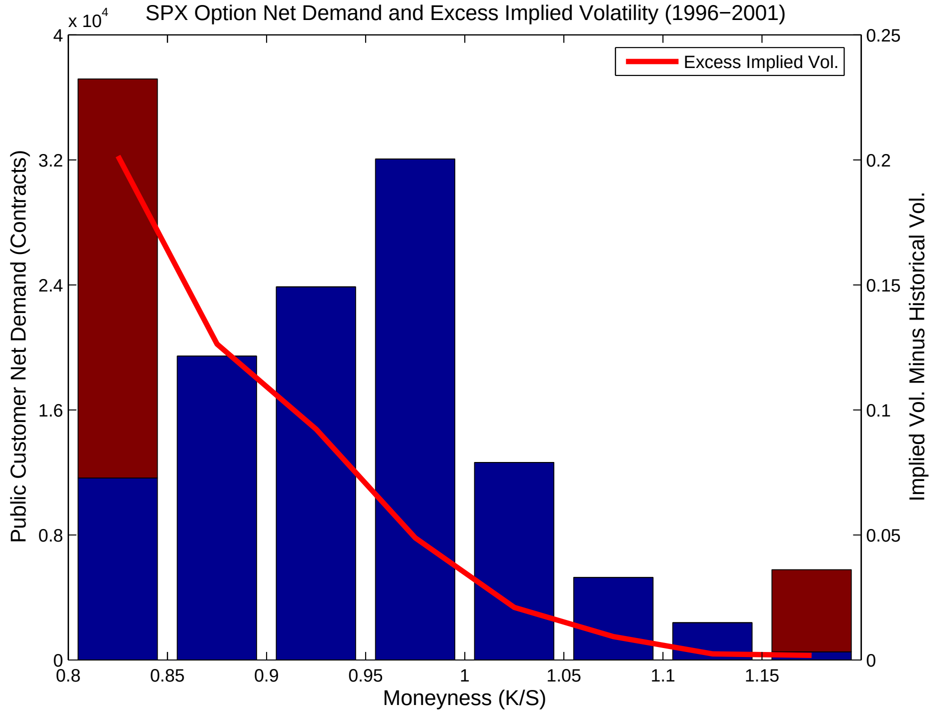

Figure 1: The bars show the average daily net demand for puts and calls from public customers for SPX options in the different moneyness categories (left axis). The top part of the leftmost (rightmost) bar shows the net demand for all options with moneyness less than 0.8 (greater than 1.2). The line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

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 and expected volatility as *excess implied volatility.*

We use different measures of the expected volatility and get similar results. We use the historical volatility for 60 trade days leading up to the observation of an option price as an ex-ante proxy for this expectation, a GARCH(1,1) estimate as an alternative proxy, and the volatility realized over the life of an option as an ex-post proxy for the expectation. For SPX options we also use the state-of-the-art measure of volatility from Bates (2005), which accounts for jumps, stochastic volatility, and the risk premium implied by the equity market, but does not add extra risk premia to (over-)fit option prices.[11]

To compute these volatility estimates directly from the daily returns on the underlying index or stock, we use daily SPX and stock returns from the Center for Research in Security Prices (CRSP). In particular, if we have N consecutive daily returns, we compute the annualized volatility of the SPX index or underlying stock over the N trade days as

$$\sigma = \sqrt{252\left(\frac{1}{N}\sum_{i=1}^{N} R_i^2\right)} \qquad (39)$$

Note that this method assumes that there are 252 trade days in a year and constrains the mean daily return to be zero as suggested by Figlewski (1997). The GARCH(1,1) estimates are computed in a standard way.

The daily average excess implied volatility for SPX options is 6.6%/ 6.5%/7.1%/ when historical/realized/GARCH(1,1) volatility is used as the proxy for the expected volatility. To compute these numbers, on each trade day we average the implied volatilities on all SPX options which 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. For the equity options, the daily average excess implied volatility per underlying stock is -0.5%/-0.4%/-0.3% when historical/realized/GARCH(1,1) volatility is used as the proxy for the expected volatility over the life of the option. These numbers suggest that on average individual equity options are just slightly inexpensive. Here we required that an option trade at least 5 contracts and have a closing bid price of at least 12.5 cents in order to includes 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 (with respect to historical volatility) for eight moneyness intervals over the 1996-2001 period.[12] 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 present the average daily net demand from public customers 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).[13]

[11] We are grateful to David Bates for providing this measure.

[12] The first (last) moneyness interval includes all options with moneyness less than 0.8 (greater than 1.2).

[13] We also constructed plots like Figure 1 with the net demand computed from non-market makers and/or with excess implied volatility defined as option implied volatility minus the realized volatility over the life of the option or minus a GARCH(1,1) prediction for the volatility over the life of the option. All of the variations shared the main features of Figure 1.

The first thing to notice in 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 thing to notice in Figure 1 is that the net demand for low-strike options is greater than the demand for high-strike options. This can potentially help explain 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. We note, however, that 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. In fact, the demand effect of these average demands can give rise to a pattern of expensiveness similar to the one observed in the context of a certain model of jump risk. [*new figure to be drawn*]

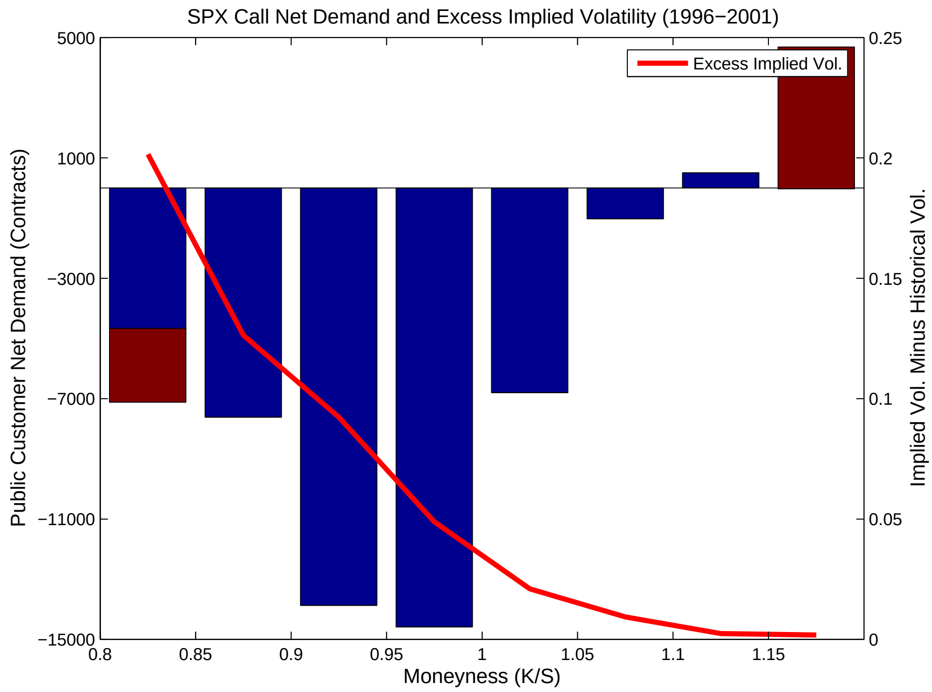

Figure 2: The bars show the average daily net demand for calls from public customers for SPX options in the different moneyness categories (left axis). The top part of the leftmost (rightmost) bar shows the net demand for all options with moneyness less than 0.8 (greater than 1.2). The line is the average SPX excess call implied volatility, that is, implied volatility minus historical volatility, for each moneyness category (right axis). The data covers 1996-2001.

We also constructed a figure like Figure 1 except that both the excess implied volatilities and the net demands were computed only from put data. Unsurprisingly, the plot looked much like Figure 1, because (as was shown above) SPX option net demands are dominated by put net demands and put-call parity ensures that (up to market frictions) put and call options with the same strike price and maturity have the same implied volatilities. For brevity, we omit this figure from the paper. Figure 2 is constructed like Figure 1 except that only calls are used to compute the excess implied volatilities and the net demands. For calls, there appears to be a negative relationship between excess implied volatilities and net demand. This relationship suggests that call net demand cannot explain the call excess implied volatilities. Proposition 2(ii) predicts, however, that it is the total demand pressure of calls and puts that matters as depicted in Figure 1. Intuitively, the large demand for puts increases the prices of puts, and, by put-call parity, this also increases the prices of calls. The relatively small negative demand for calls cannot overturn this effect.

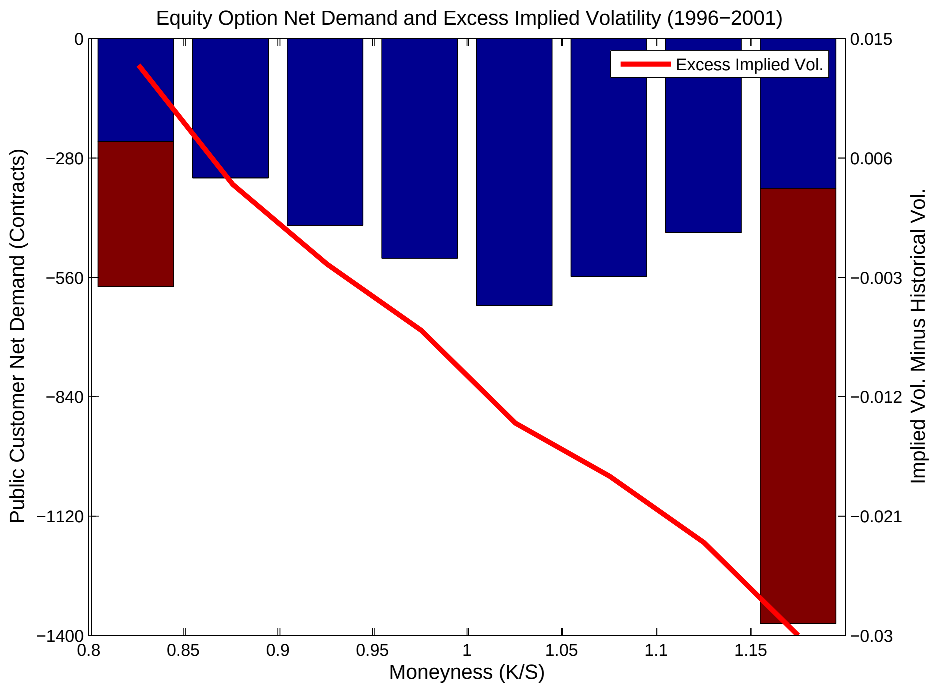

Figure 3: The bars show the average daily net demand per underlying stock from public customers for equity options in the different moneyness categories (left axis). The top part of the leftmost (rightmost) bar shows the net demand for all options with moneyness less than 0.8 (greater than 1.2). The line is the average equity option excess implied volatility, that is, implied volatility minus historical volatility, for each moneyness category (right axis). The data covers 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 historical volatility) per underlying stock for eight moneyness intervals over the 1996-2001 period.[14] 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 2.5% across the moneyness categories. By contrast, for the SPX options the excess implied volatility line varies by over 10% across the corresponding moneyness categories. The bars in the figure present the average daily net demand per underlying stock from public customers for equity options in the moneyness categories. The figure shows that public customers are net sellers of equity options on average, consistent with these options being cheap. Further, the figure shows that customers sell mostly high-strike options, consistent with these options being especially cheap. If the figure is constructed from only calls or only puts, it looks roughly the same (although the magnitudes of the bars are about half as large.)

Figure 4 plots the daily net positions (i.e., net demands) for SPX options aggregated across moneyness and maturity for public customers, firm proprietary traders, and market makers. The daily public customer net positions range from -1065 contracts to $+385,750$ contracts, and it tends to be larger over the first year or so of the sample. Although the public customer net position shows a good deal of variability, it is nearly always positive and never far from zero when negative. To a large extent, the market maker net option position is close to the public customer net position reflected across the horizontal axis. This is not surprising, because on each trade date the net positions of the three groups must sum to zero and the public customers constitute a much larger share of the market than the firm proprietary traders. The firm proprietary and market maker net positions roughly move with one another. In fact, the correlation between the two time-series is 0.44. This positive co-movement suggests that a non-trivial part of the firm proprietary option trading may be associated with supplying liquidity to the SPX option market. For this reason, in the empirical work below we usually define end users as public customers, although the results are similar when end users are defined as public customers plus firm proprietary traders. The correlations between the public customer time-series and those for firm proprietary traders and market makers on the other hand are, respectively, -0.78 and -0.90.

[14]The first (last) moneyness interval includes all options with moneyness less than 0.85 (greater than 1.15).

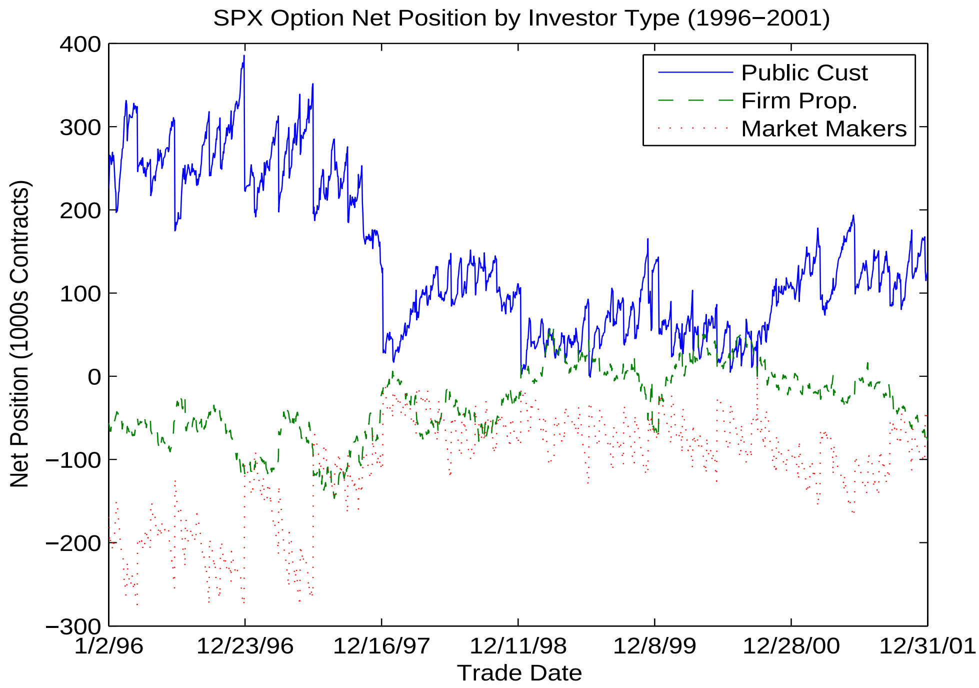

Figure 4: Time series of the daily net positions for SPX options aggregated across moneyness and maturity for public customers, firm proprietary traders, and market makers.

5 Empirical Results

Proposition 3 states that positive (negative) demand pressure on one option increases (decreases) the price of all options on the same underlying asset while Proposition 4 states that demand pressure on low (high) strike options has a greater price impact on low (high) strike options. The empirical work in this section of the paper examines these two predictions of the model by investigating both in the SPX and equity option markets whether overall excess implied volatility is higher on trade dates where net demand for options is higher and whether the excess implied volatility skew is steeper on trade dates where the skew in the net demand for options is steeper.

5.1 Excess Implied Volatility and Net Demand

We investigate first the time-series evidence for Proposition 3 by regressing a measure of excess implied volatility on a measure of option net demand:

$$ExcessImplVol_t = a + b\, NetDemand_t + \epsilon_t \qquad (40)$$

Table 2: The relationship between SPX public customer demand pressure and SPX Excess Implied Volatility relative to volatility implied by Bates (2005). T-statistics computed using Newey-West are in parentheses.

Panel A: Before Structural Changes, 1996/01–1996/10

a	b	Adj R^2	N
0.018	7.3E-8	0.01	10
(0.82)	(0.37)		

Panel B: After Structural Changes, 1997/10–2001/12

a	b	Adj R^2	N
0.038	7.0E-7	0.21	50
(7.9)	(4.7)		

We consider first the time series relationship for SPX options. We define $ExcessImplVol_t$ as the average implied volatility of 1-month at-the-money SPX options minus the corresponding volatility of Bates (2005). When computing this variable, the SPX options which are included are those which have at least 25 contracts of trading volume, more than 14 and fewer than 43 calendar days to expiration, 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.)[15] $NetDemand$ is the net public customer demand for all SPX options which have 10–180 calendar days to expiration and moneyness between 0.9 and 1.10.

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 stronger in a daily regression, not reported.)

The results are shown in Table 2. We report the results over two subsamples because, as seen in Figure 4, there appears to be a structural change in 1997. Also, a structural change happens in the time series of open interest (not shown). These changes may be related to several events that change 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 exchange Chicago Mercantile Exchange (CME), the introduction of Dow Jones options on CBOE, and changes in the margin requirements. Some of our results hold over the full sample, but their robustness and explanatory power are smaller. Of course, we must entertain the possibility that the

[15]By 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.

model's limited ability to jointly explain the full sample is due to problems with the theory.

We see that the estimate of the demand effect b is insignificant over the first subsample, but positive and statistically significant over the second longer subsample. The expensiveness and the fitted values are plotted in Figure 5, which clearly shows the relation between demand and expensiveness over the late sample. The fact that the b coefficient is positive indicates that on average when SPX net demand is higher (lower), excess SPX implied volatilities are also higher (lower). The point estimate of 7.0E-7 suggests that on average excess implied volatility increases by 7.0E-7 for each additional contract of SPX net demand. If the dependent variable changed from its lowest to highest value over the late sub-sample, the effect on implied volatility would be 6 percentage points. Another way to judge the magnitude is to consider the average net demand for the considered SPX options over the late subsample, ca. 16,000 contracts, which implies that demand pressure may increase option implied volatilities on a typical day by 1.1 percentage points. For the SPX options used to compute the dependent variable in this regression, the average expensiveness is 4.6%. Consequently, end user demand may explain about a quarter of the average expensiveness of SPX options and also a significant fraction of its time variation.

We consider next the time-series relationship between demand and expensiveness for equity options. In particular, we run the time-series the regression (40) for each stock, and average the coefficients across stocks. The results are shown in Table 3. We consider separately the subsample 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 ability 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.

The coefficient b measuring the effect of demand on expensiveness is positive and significant in both subsamples. This means that, the larger is demand for equity options, the higher is their implied volatility. The results are illustrated in Figure 6, which shows the expensiveness and fitted values of the demand effect on a monthly basis. The 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 variables. It is comforting, however, that the relation also holds when we consider each stock separately.

Finally, we investigate the cross-sectional relationship between excess implied volatility and net demand in the equity option market. We do this by performing daily regressions of Equation (40) for options on cross-sections of underlying stocks that meet a number of criteria. On each trade date the initial universe of underlying stocks are the 303 which have strictly positive option volume on at least 80% of the trade days from the beginning of 1996 through the end of 2001. We then identify those underlying

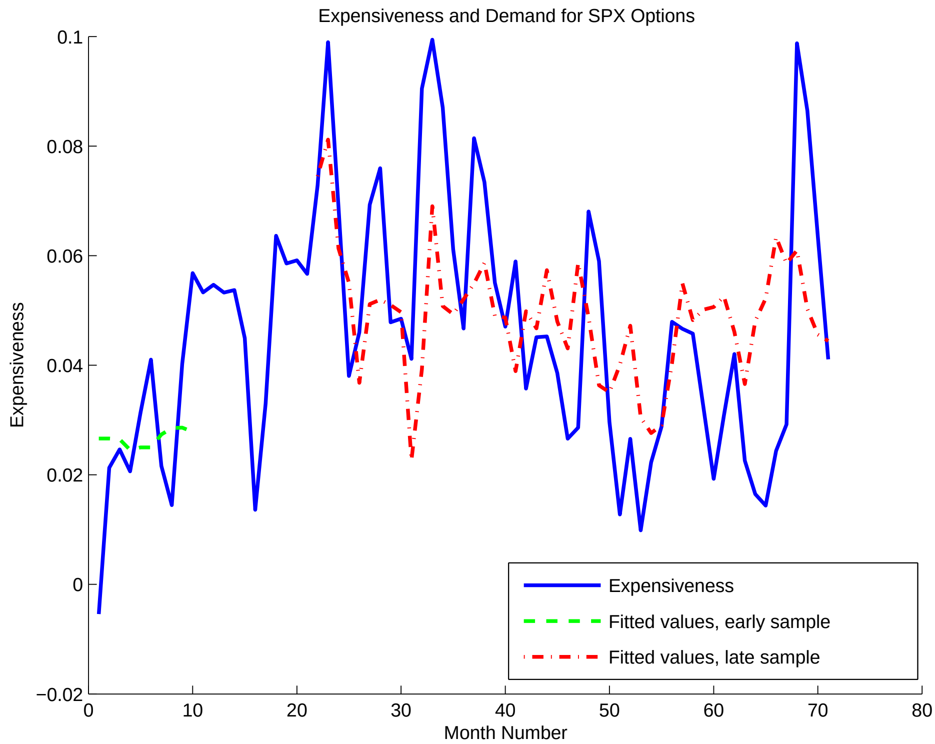

Figure 5: The solid line shows the expensiveness of SPX options, that is, implied volatility of 1-month at-the-money options minus the volatility measure of Bates (2005) which takes into account jumps, stochastic volatility, and the risk premium from the equity market. The dashed lines are, respectively,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

stocks that have two or more options with moneyness greater than or equal to 0.97 and less than or equal to 1.03, maturity greater than or equal to 15 calendar days and less than or equal to 45 calendar days, closing ask prices between 0 and 2 dollars greater than closing bid prices, at least 5 contracts of trading volume, and implied volatilities available on OptionMetrics. We also require that on at least 54 of the previous 60 trade days the underlying stocks have daily returns on CRSP and OptionMetrics that differ by less than one percent. For the underlying stocks that meet these criteria, we compute the excess implied volatility for the trade day as the average implied volatility of the options minus the sample volatility of the previous stock returns computed according to Equation (39). For both the implied volatilities and the stock returns we use the data which meet the criteria spelled out above. We multiply the net demand variable by the price volatility of the underlying stock (defined as the sample return

Table 3: The relationship between equity option public customer demand pressure and excess implied volatility relative to GARCH volatility. T-statistics are in parentheses.

Panel A: Before cross-listing of options, 04/1996–06/1999

a	b	Adj R^2	N
-0.011	9.26E-06	0.06	609
(-0.07)	(5.27)		

Panel B: After cross-listing of options, 10/1999–12/2001

a	b	Adj R^2	N
0.023	6.01E-06	0.07	422
(5.61)	(4.88)		

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. When public customer net demand is used, the slope coefficient is $4.08E-08$ with a t-statistic of 8.32. When non-market maker net demand is used, the slope coefficient is $5.91E-08$ with a t-statistic of 6.44. These findings provide empirical verification for Proposition 3.

Figure 7 plots a kernel regression estimate of the relationship between excess implied volatility and public customer net demand for underlying stocks across all trade days. The function, which is plotted from the 10th to the 90th percentile of the net demand data, clearly has a positive slope which indicates that option expensiveness is increasing in net option demand. The confidence intervals are not corrected for error correlation, so they should merely be viewed as a graphical illustration of where most of the data is concentrated. (The above t-statistics corrects for the correlation.)

5.2 Implied Volatility Skew and Net Demand Skew

In order to investigate Proposition 4, we regress a daily measure of the steepness of the excess implied volatility skew on the skewness in the option net demand:

$$ExcessImplVolSkew_t \quad = \quad a + b\, NetDemandSkew_t + \epsilon_t. \tag{41}$$

Here, $ExcessImplVolSkew_t$ is the date t skew in excess implied volatility, defined as the difference between the average implied volatility from "low moneyness" options and

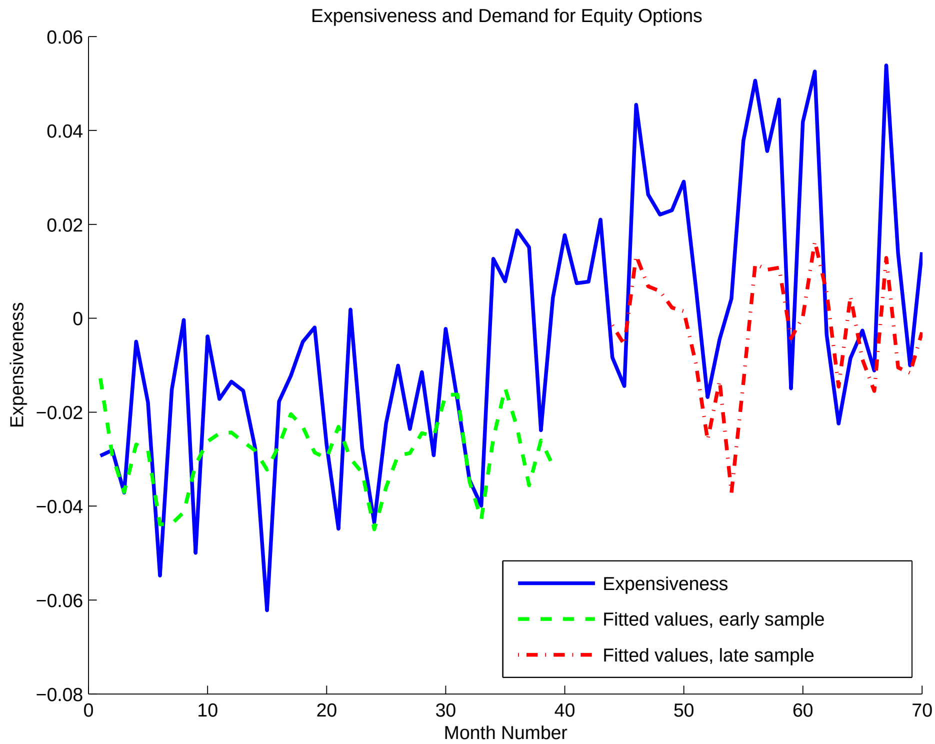

Figure 6: The solid line is the expensiveness of equity options, averaged across stocks. The dashed lines are, respectively, the fitted values of demand-based expensiveness before and after the cross-listing of options (1996/04–1999/05 and 1999/10–2001/12) using the average regression coefficients from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

the average implied volatility from options with moneyness "close to one."[16] For SPX options, the "low moneyness" options are defined as those with moneyness between 0.90 and 0.94 which trade at least 25 contracts on trade date t and have more than 14 and fewer than 46 calendar days to expiration, and the options with moneyness "close to one" are defined as those with moneyness between 0.98 and 1.02 which meet the same volume and maturity criteria. Similarly, $NetDemandSkew_t$ is the skew in net option demand, defined as the net public customer demand for options with moneyness between 0.90 and 0.98 minus the net public customer demand for options with moneyness between 0.98 and 1.05. Only options with more than 14 and fewer than 46 calendar days to expiration are included in the computation of net demand skew.

Table 4 reports the OLS estimates of the skewness regression for monthly data and

[16]Of course, subtracting the historical volatility from both groups of options would not change the value of this variable.

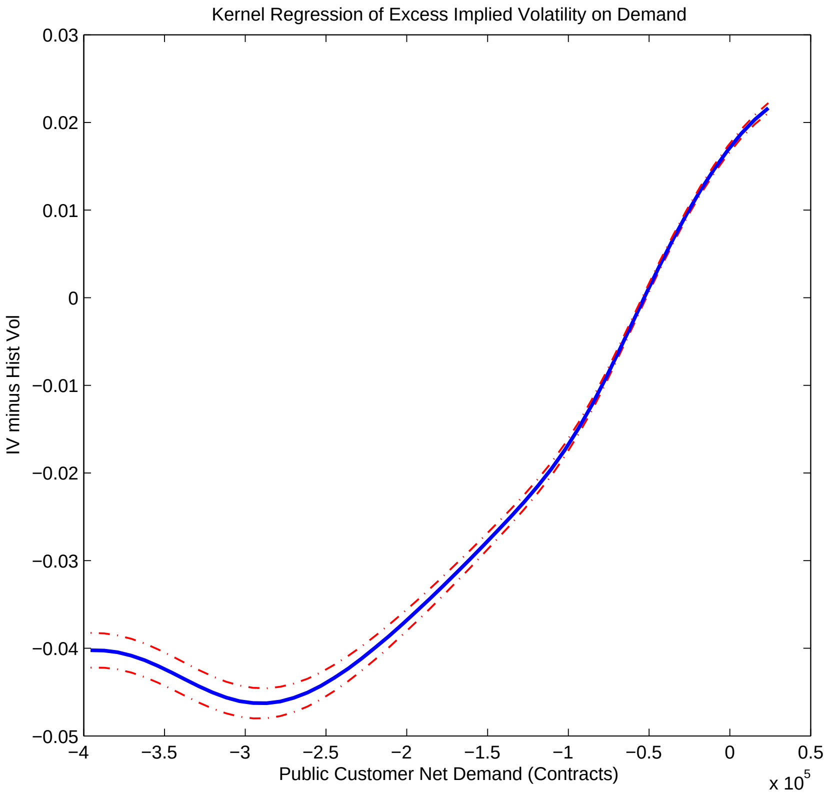

Figure 7: The solid line plots a kernel regression estimate of the relationship between excess implied volatility and public customer net demand for underlying stocks across all trade days. The dashed lines are 95% confidence intervals, not corrected for error correlation.

Figure 8 illustrates the effects. As discussed in Section 5.1, we divide the sample into two subsamples because of structural changes. The b coefficient is significantly positive in the late subsample.[17] The coefficient estimate of 3.5E-7 indicates that increasing the net demand for low moneyness options by one contract or decreasing the net demand for high moneyness options by one contracts is associated with a 3.5E-7 increase in the implied volatility of low moneyness short maturity options relative to the implied volatility of short maturity options with moneyness close to one. Consequently, if demand skew changes from its smallest to largest value from the late sub-sample, the implied change of the volatility skew is 3 percentage points.

[17]The b coefficient is also significant over the full sample; the demand skewness is less non-stationary than the level of demand over the full sample.

Table 4: Skewness in Implied Volatility versus Skewness in Net Demand. The relationship between SPX Implied Volatility Skew and SPX public customer demand pressure Skew. T-statistics computed using Newey-West are in parentheses.

Panel A: Before Structural Changes, 1996/01–1996/10

a	b	Adj R^2	N
0.072	-0.67E-7	0.08	10
(42)	(-0.86)		

Panel B: After Structural Changes, 1997/10–2001/12

a	b	Adj R^2	N
0.068	3.5E-7	0.14	50
(26)	(2.8)		

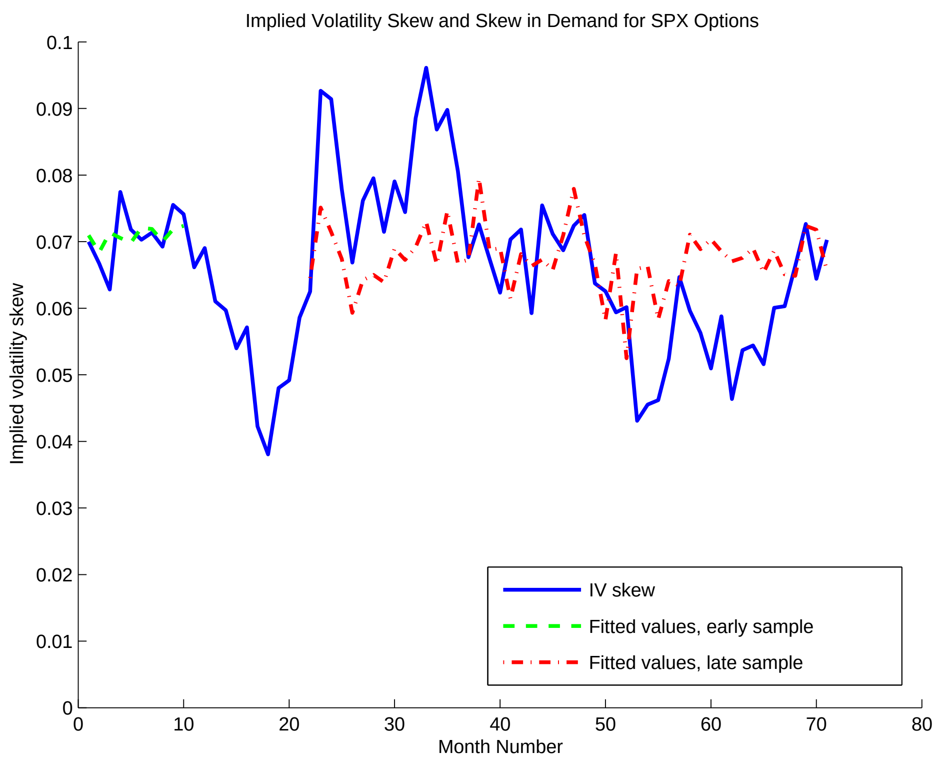

Figure 8: The solid line shows the implied volatility skew for SPX options. The dashed lines are, respectively, the fitted values from the skew in demand before and after certain structural changes (1996/01–1996/10 and 1997/10–2001/12).

6 Conclusion

Relative to the Black-Scholes-Merton benchmark, index and equity options display a number of robust pricing anomalies. A great deal of research has attempted to address these anomalies, in large part by generalizing the Black-Scholes-Merton assumption about the dynamics of the underlying asset. While these effort 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 anomalies, 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 perfectly hedged. Consequently, if intermediaries such as market makers and proprietary traders 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, on average are slightly inexpensive and have small negative end user demand. In accordance with our theory's predictions, we find that both in the index and equity option markets 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 net end-user demand across moneyness.

A Proofs

Proof of Lemma 1:

Note first that the boundedness of all the random variables considered (with the exception of S) ensures that all expectations are finite.

The Bellman equation is

$$\begin{aligned} J(W_t;t,X_t) &= -\frac{1}{k}e^{-k(W_t+f_t(d_t,X_t))} \\ &= \max_{C_t,q_t,\theta_t}\left\{-\frac{1}{\gamma}e^{-\gamma C_t}+\rho \mathrm{E}_t\left[J(W_{t+1};t+1,X_{t+1})\right]\right\} \end{aligned} \tag{42}$$

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

$$0 = e^{-\gamma C_t}+kr\rho \mathrm{E}_t\left[J(W_{t+1};t+1,X_{t+1})\right] \tag{43}$$

which together with (42) yields

$$0 = e^{-\gamma C_t}+kr\left[J(W_t;t,X_t)+\frac{1}{\gamma}e^{-\gamma C_t}\right] \tag{44}$$

that is,

$$e^{-\gamma C_t} = e^{-k(W_t+f_t(d_t,X_t))} \tag{45}$$

implying (4). The FOC's for θ_t and q_t are (5) and (6). We derive f recursively as follows. First, we let $f(t+1,\,\cdot\,)$ be given. Then, θ_t and q_t are given as the unique solutions to Equations (5) and (6). Clearly, θ_t and q_t do not depend on the wealth W_t. Further, (44) implies that

$$0 = e^{-\gamma C_t}-r\rho \mathrm{E}_t\left[e^{-k\left(y_{t+1}+(W_t-C_t)r+q_t(p_{t+1}-rp_t)+\theta_t R^e_{t+1}+f_{t+1}(d_{t+1},X_{t+1})\right)}\right] \tag{46}$$

that is,

$$e^{-\gamma C_t-krC_t+krW_t} = r\rho \mathrm{E}_t\left[e^{-k\left(y_{t+1}+q_t(p_{t+1}-rp_t)+\theta_t R^e_{t+1}+f_{t+1}(d_{t+1},X_{t+1})\right)}\right], \tag{47}$$

which, using (4), yields the equation that defines $f_t(d_t,X_t)$ (since X_t is Markov):

$$e^{-krf_t(d_t,X_t)} = r\rho \mathrm{E}_t\left[e^{-k\left(y_{t+1}+q_t(p_{t+1}-rp_t)+\theta_t R^e_{t+1}+f_{t+1}(d_{t+1},X_{t+1})\right)}\right] \tag{48}$$

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

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

subject to

$$W_{t+1} = y_{t+1} + (W_t - C)r + \theta R^e_{t+1},$$

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

Furthermore, AF maps any function of the type

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

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 now show that there exists $m > 0$ such that A maps the set

$$G^m = \{g : \mathbb{X} \to \mathbb{R} \,:\, g \text{ is continuous}, \, g \geq m\}$$

into itself.

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

$$\begin{aligned} Ag &\geq \inf_x \inf_\theta r\rho \mathrm{E}\left[e^{-k\left(y_{t+1}+\theta R^e_{t+1}\right)} g(X_{t+1})^{\frac{1}{r}} \right| X_t = x\Big] \\ &\geq \inf_x \inf_\theta r\rho \mathrm{E}\left[e^{-k\left(y_{t+1}+\theta R^e_{t+1}\right)} \right| X_t = x\Big]\left(\min_z g(z)\right)^{\frac{1}{r}} \\ &\geq B\left(\min_z g(z)\right)^{\frac{1}{r}} \end{aligned}$$

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

Since G^m is complete, we conclude that A has a (unique) fixed point in G^m (which, therefore, is not 0).

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

$$\lim_{t\to\infty} \mathrm{E}\left[\rho^{-t}e^{-kW_t}\right] = 0.$$

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

$$W_{t+1} = W_t - (r-1)f(X_t) + y_{t+1} + \theta_t R^e_{t+1},$$

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

$$\mathrm{E}_T\left[e^{-k(W_t+f(X_t))}\right] = e^{-k(W_T+f(X_T))},$$

which is bounded. Since $f(X_t)$ is bounded, it follows that $\lim_{t\to\infty} \mathrm{E}\left[\rho^{-t} e^{-kW_t}\right] = 0$.

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

$\square$

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 (6). It is immediate that p_t is measurable with respect to time-t information.

$\square$

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).

$\square$

Proof of Proposition 3:

Part a) follows immediately from the Cauchy-Schwarz inequality, so we offer a proof of part b). The proof is based on the following result.

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*

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

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

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

$$\Psi = \mathrm{E}^d\left[\tilde{p}^i \tilde{p}^j\right] \mathrm{Var}(R^e) - \mathrm{E}^d\left[\tilde{p}^i R^e\right] \mathrm{E}^d\left[\tilde{p}^j R^e\right],$$

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

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

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 .

$\square$

Proof of Lemma 2: We consider three cases: *(i)* the intersection of the graphs of h_1 and $g \equiv \gamma + \alpha Id + \beta h_2$ is empty; *(ii)* their intersection is a singleton; *(iii)* their intersection 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 d by differentiating with respect to ϵ (evaluated at $\epsilon = 0$). From the dealer's problem it follows that

$$p_t = \mathrm{E}_t \left[\rho^{s-t} e^{-\gamma(C_s - C_t)} p_s\right]$$

which implies

$$\begin{aligned}\frac{\partial p_t}{\partial \epsilon} &= \mathrm{E}_t \left[\rho^{s-t} e^{-\gamma(C_s - C_t)} \frac{\partial p_s}{\partial \epsilon}\right] \\ &= r^{-(s-t)} \mathrm{E}^*_t \left[\frac{\partial p_s}{\partial \epsilon}\right] \\ &= r^{-(s-t)} \mathrm{E}^*_t \left[\frac{\partial p_s}{\partial \tilde{q}^j_s} q^j_s\right].\end{aligned}$$

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

$$\frac{\partial C_s}{\partial q^j_s} = \frac{k}{\gamma} \frac{\partial f(s, X_s; q)}{\partial q^j_s} = -\frac{k^2 r \rho}{\gamma} \mathrm{E}^*_s \left[p^j_{s+1} - r p^j_s + \frac{\partial \theta_s}{\partial q^j_s} R^e_{s+1}\right] = 0 \qquad (49)$$

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

$$e^{-\gamma C_t} = e^{-\rho(s-t)} \mathrm{E}_t \left[e^{-\gamma C_s}\right],$$

which gives that

$$e^{-\gamma C_t}\frac{\partial C_t}{\partial \epsilon} \quad = \quad e^{-\rho(s-t)}\mathrm{E}_u\left[e^{-\gamma C_s}\frac{\partial C_s}{\partial \epsilon}\right] = 0$$

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:

$$e^{-\gamma C_t} = (r\rho)^{T+1-t}\,\mathrm{E}_t\left[e^{-k(W_{T+1}+f_{T+1})}\right] \tag{50}$$

$$0 = \mathrm{E}_t\left[e^{-k(W_{T+1}+f_{T+1})}R^e_{t+1}\right] \tag{51}$$

$$0 = \mathrm{E}_t\left[e^{-k(W_{T+1}+f_{T+1})}\left(p_{T+1} - r^{T+1-t}p_t\right)\right], \tag{52}$$

with

$$W_{T+1} = (W_t - C_t)r^{T+1-t} + \sum_{s=t}^{T}\left(y_{s+1} - C_{s+1} + \theta_s R^e_{s+1} - \epsilon_s d_s\left(p_{s+1} - rp_s\right)\right)r^{T+1-s} + y_{T+1}.$$

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, θ_s, C_s) is a smooth function of $(\epsilon_s, \ldots, \epsilon_{s+1})$. Note that the statement holds trivially for $s > T$.

Assume therefore the statement for all $s > t$. There are n_t+2 equations in (50)–(52), 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 (50)–(52) 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 (50)–(52), it follows easily that

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

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

Proof of Proposition 4:

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

$$\Pi(x) \quad = \quad -d(K - rS_t - x)^+ + \theta x.$$

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

$$\mathrm{E}^* \left[e^{-k\Pi(x)} x \right] \quad = \quad 0.$$

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

$$\int_K^\infty \left(e^{-k\Pi(x)} x - e^{-k\Pi(-x)} x \right) \xi(x)\, dx \quad = \quad -\int_0^K \left(e^{-k\Pi(x)} x - e^{-k\Pi(-x)} x \right) \xi(x)\, dx \quad < \quad 0.$$

It immediately follows that it cannot be true that $\Pi(-x) \geq \Pi(x)$ for all $x > |K - rS_t|$. In other words, for some value $x > |K - rS_t|$, $\Pi(-x) < \Pi(x)$, 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 $S_{t+1} < \bar{K}$,

$$\Pi(S_{t+1} - rS_t) < \Pi(-(S_{t+1} - rS_t)),$$

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

$$\begin{aligned} p(p, K', d) \quad &> \quad p(c, K'', d) \\ p(p, K', 0) \quad &= \quad p(c, K'', 0), \end{aligned}$$

the second relation being the result of symmetry.

□

References

Amin, K., J. D. Coval, and H. N. Seyhun (2003): "Demand for Portfolio Insurance and Index Option Prices," *Journal of Business*, forthcoming.

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.

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.

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., 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, 13231352.

Coval, J. D., and T. Shumway (2001): "Expected Option Returns," *Journal of Finance*, 56(3), 983–1009.

De Fontnouvelle, 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.*

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*, fortcoming.

Hakansson, N. (1979): "The Fantastic World of Finance: Progress and the Free Lunch," *The Journal of Financial and Quantitative Analysis*, 14, 717–734.

Hugonnier, J., D. Kramkov, and W. Schachermayer (2004): "On Utility Based Pricing of Contingent Claims in Incomplete Markets," *Mathematical Finance, forthcoming*.

Jackwerth, J. (2000): "Recovering risk aversion from option prices and realized returns," *Review of Financial Studies*, 13, 433–451.

Longstaff, F. A. (1995): "Option Pricing and the Martingale Restriction," *Review of Financial Studies*, 8(4), 1091–1124.

Newman, Y., and M. Rierson (2004): "Illiquidity Spillovers: Theory and Evidence from European Telecom Bond Issuance," Working Paper, Stanford Graduate School of Business.

Rubinstein, M. (1976): "The Valuation of Uncertain Income Streams and Pricing of Options," *Bell Journal of Economics*, 7, 407–425.

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.

Wurgler, J., and E. V. Zhuravskaya (2002): "Does Arbitrage Flatten Demand Curves for Stocks?," *Journal of Business*, 75(4), 583–608.