Journal of Banking & Finance 34 (2010) 2897–2916

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Trading activity and bid–ask spreads of individual equity options

Jason Wei [a,*], Jinguo Zheng [b]

[a] Joseph L. Rotman School of Management, University of Toronto, Toronto, Ontario, Canada
[b] Guanghua School of Management and Institute of Strategic Research, Peking University, Beijing, China

ARTICLE INFO

Article history:
Received 17 July 2009
Accepted 13 February 2010
Available online 19 March 2010

JEL classification:
G10
G12
G14

Keywords:
Trading activity
Bid–ask spread
Option expiration cycles
Liquidity
Liquidity determinants
Volume
Open interest

ABSTRACT

We empirically examine the impact of trading activities on the liquidity of individual equity options measured by the proportional bid–ask spread. There are three main findings. First, the option return volatility, defined as the option price elasticity times the stock return volatility, has a much higher power in explaining the spread variations than the commonly considered liquidity determinants such as the stock return volatility and option trading volume. Second, after controlling for all the liquidity determinants, we find a maturity-substitution effect due to expiration cycles. When medium-term options (60–90 days maturity) are not available, traders use short-term options as substitutes whose higher volume leads to a smaller bid–ask spread or better liquidity. Third, we also find a moneyness-substitution effect induced by the stock return volatility. When the stock return volatility goes up, trading shifts from in-the-money options to out-of-the-money options, causing the latter's spread to narrow.

1. Introduction

The literature on option liquidity is sparse. Among the few papers that study liquidity for various derivatives markets,[1] only Mayhew et al. (1999), Kalodera and Schlag (2004) and Cao and Wei (2010) empirically examine liquidity for individual equity options. Mayhew et al. (1999) link the options order flow to the characteristics of the underlying stock and find that options are more liquid for stocks with a higher price, greater volatility and higher trading volume. Kalodera and Schlag (2004) perform a similar analysis on German stocks and options and find that a higher stock trading volume positively impacts options' liquidity measured by volume and trading frequency. In a recent study, Cao and Wei (2010) examine liquidity commonality and other features of the overall equity option market.

The literature on the properties of equity-option liquidity per se is even thinner. George and Longstaff (1993), Chong et al. (2003) and Cao and Wei (2010) are three studies loosely related to the topic. George and Longstaff (1993) examine the bid–ask spread of the S&P 100 index options across exercise prices and maturities and find that the cross-sectional differences in liquidity are linked to market-making costs and trading frequency. They also find that call and put options tend to be substitutes – calls are traded more frequently when the bid–ask spread for puts is higher, and vice versa. Their study is largely cross-section in that they only examine data for 1989. Chong et al. (2003) find a negative relation between maturity and the bid–ask spread for at-the-money, OTC currency options. Cao and Wei (2010) mostly focus on the liquidity covariation among stock options. The literature is completely lacking on the time-series properties of liquidity for individual equity options.[2]

Our paper takes the first step toward filling this gap in the literature. We ask and attempt to answer the following two related questions: (1) What are the important liquidity determinants for stock options? (2) Is the time-series behavior of option liquidity affected by trading activities other than the identified liquidity determinants? The trading activities we focus on are related to option expiration cycles and the level of stock return volatility. Using the proportional bid–ask spread (PBA) to gauge liquidity for actively traded options retrieved from OptionMetrics from January 1996 to June 2007, we obtain several interesting findings.

* Corresponding author. Tel.: +1 416 978 3698; fax: +1 416 971 3048.
E-mail address: wei@rotman.utoronto.ca (J. Wei).

[1] These include Vijh (1990), Cho and Engle (1999), Mayhew et al. (1999), Brenner et al. (2001), Kalodera and Schlag (2004), Tang and Yan (2008), Deuskar et al. (2009) and Cao and Wei (2010).

[2] In contrast, the literature on stock market liquidity is rich. In this journal alone, several papers on this topic have been published recently (e.g. Chan et al., 2008; Frino et al., 2008; Rakowski and Beardsley, 2008; Mantecon and Poon, 2009; Wang, 2010).

0378-4266/$ - see front matter
doi:10.1016/j.jbankfin.2010.02.012

Regarding the first question, we find that time to maturity, moneyness, stock return volatility, option return volatility, option trading volume and option price all affect the level of PBA. Contrary to common beliefs, stock return volatility and option volume, though both negatively related to PBA, have the least impact compared with other determinants. Options with a shorter maturity, being out of the money and with lower prices tend to have a larger PBA. The above determinants together explain about 55% of the time-series variations in PBA. The most striking finding is the strong explanatory power of the option return volatility, denned as the option price elasticity times the stock return volatility, which alone explains over 45% of the PBA variations. In other words, option return volatility is by far the most significant determinant of option liquidity. That the option's bid–ask spread is inherently linked to its return volatility is consistent with the theoretical models (Stoll, 1978a; Ho and Stoll, 1981) and empirical evidence (Benston and Hagerman, 1974; Branch and Freed, 1977) for stocks.

Regarding the second question, supplementing the call-put substitution effect for index options (George and Longstaff, 1993), we uncover two substitution effects for equity options. The first, the maturity-substitution effect, establishes an intricate link between option liquidity and the expiration cycle. By convention, the expiration months follow a quarterly cycle and there must be options available for the current and the next months. This convention causes the third expiry month to not always situate right after the second expiry month. There can be a 1- or 2-month gap, and the width of this gap varies for the same stock depending on where the stock is with respect to its own expiration cycle. We show that this maturity gap causes traders to substitute the short-term options (i.e., options maturing in the current month or the next) for the distant maturity options. This maturity-substitution has a direct bearing on the liquidity of short-term options: the PBA of short-term options decreases when the trading volume is shifted from the distant maturity options to the short-term options. This substitution effect remains significant even after controlling for all the identified liquidity determinants. The average impact on PBA is about 20 basis points, a relatively modest impact in light of the average PBA of 8.6% in our sample.

The second substitution effect, called "moneyness-substitution effect," has to do with moneyness and stock return volatility. When the stock return volatility increases, trading is shifted from in-the-money options to out-of-the-money options, especially for short-term options. This shift in volume directly affects the relative liquidity of in-the-money and out-of-the-money options. Specifically, when the stock return volatility goes up, the PBA of out-of-the-money options decreases relative to in-the-money options. This moneyness-substitution effect is the strongest for the current month options. Here, when the stock return volatility goes up by 5% (e.g., from 30% to 35%), the PBA of out-of-the-money options goes down by about 40 basis points relative to that of in-the-money options.

Our study contributes to the literature in three aspects. First, we offer insights into the determination and time-series properties of liquidity for individual equity options, an area that is currently void in the literature. The finding that the option return volatility is the single most important determinant of proportional bid–ask spread will have direct implications for all future studies on option liquidity.

Second, the discovery of the two substitution effects sheds light on the unique feature of option liquidity. To the best of our knowledge, George and Longstaff (1993) is the only study documenting a substitution effect among options. Their focus is on the substitution between calls and puts while we uncover two additional types of substitution along the dimensions of moneyness and maturity. Besides, George and Longstaff (1993) only examine S&P 100 index options while we examine a cross-section of individual equity options. The economic significance of the two substitution effects is by no means large compared with the absolute level of the spread itself. However, our findings are best judged on contributing to the overall understanding of option liquidity, not on profitable trading strategies.

The third aspect in which we contribute to the literature is the implication for the breakdown of bid–ask spreads. The literature on spread decomposition for stocks is voluminous (examples include Stoll, 1978a,b; Ho and Stoll, 1981; Glosten and Milgrom, 1985; Glosten and Harris, 1988). The three commonly considered components/factors are order processing cost, inventory cost and adverse selection cost. No consensus has emerged as to the relative importance of the three components. Similar research for options does not yet exist. Our finding that more than 45% of the spread variations can be explained by the option return volatility alone speaks to the importance of inventory risk in the formation of option spreads. On a day-to-day basis, the option return volatility measures the inventory risk faced by market makers. The high explanatory power of this return volatility means that a large portion of the spread is due to inventory risk.

The rest of the paper is organized as follows. In Section 2, we describe the data and explain the option expiration cycles. Section 3 studies the option liquidity determinants and trading patterns around expiration cycles. Section 4 relates the proportional bid–ask spread to maturity- and moneyness-substitutions. In Section 5, we present some alternative tests and briefly discuss the robustness checks. Section 6 concludes the paper.

2. Data and option expiration cycles

2.1. Data

The data source is Ivy DB's OptionMetrics, which covers all exchange-traded options on US stocks. The dataset starts on January 1, 1996 and the cut-off date for this study is June 30, 2007, covering a period of eleven and half years. Among other things, the database provides the day-end best bid and ask quotes, open interest, volume, delta and implied volatility (the last two items are calculated via a binomial tree with a constant interest rate). For the underlying stocks, the dataset provides the daily high/low/close prices and trading volume. To ensure that we have enough option observations for the purpose of studying the expiration cycle effects, we restrict our sample to the top 100 stocks ranked by option volume. Specifically, for each calendar year, we rank the stocks by their total option volume and select the top 100 stocks. This procedure results in 287 distinct tickers. The option data are further screened in the following manner:

- To minimize the impact of tick size on bid–ask spreads, we delete observations where the bid is lower than $0.125 or $1/8. This screening criterion is conservative in that after 2001 the tick size was reduced from $1/16 and $1/8 (for option prices below and above $3) to $0.05 and $0.10, respectively. Therefore, $0.125 represents the largest possible tick size in the entire dataset.
- To avoid potential recording errors, we delete observations whose trading volume is larger than five times the sum of today's and yesterday's open interests; observations with an option volume larger than the stock volume is also deleted.[3]

[3] OptionMetrics reports bid and ask quotes associated with both non-zero and zero trading volumes. In this study we retain the zero-volume quotes since our focus is on the behavior of bid–ask spreads. For certain analyses (e.g., when trading volume is a dependent variable in the regression), we delete the zero-volume observations. The zero-volume observations account for about 30% of the screened dataset.

- Since we study the effect of expiration cycles on the liquidity of short-term options, we only keep options with a maturity shorter than or equal to eight months or 240 days.
- To avoid potential biases caused by deep in-the-money and deep out-of-the-money options, we delete options whose moneyness (denned as the exercise price divided by the stock price) is outside the range of [0.9, 1.1].
- To eliminate outliers, we delete observations with too large a percentage bid–ask spread. We arbitrarily set the cut-off percentage bid–ask spread at 200% for out-of-the-money options and 150% for in-the-money options.

2.2. Expiration cycles

When studying the impact of trading activities on option bid–ask spreads, the expiration cycle is a natural aspect to focus on since it is unique to option trading. In the US stock options expire on the third Saturday of the month. In addition, it is stipulated that at any given point in time, there must be four expiration months available for trading. Each optioned stock must also have options expiring in the current month and in the following month (our discussions here do not apply to LEAPs which usually expire in January only). Among the four available expiration months, the last two are not arbitrary. They are determined according to which expiration cycle the stock belongs to. When a stock is initially chosen for option listing, it is randomly assigned to one of the three expiration cycles: January cycle, February cycle or March cycle. Within each cycle the available expiration months are separated by a quarter. Therefore, the January cycle is also called the JAJO cycle (January, April, July and October). By the same token, the February cycle is called the FMAN cycle (February, May, August and November) and the March cycle the MJSD (March, June, September and December). Within our screened sample, the number of stocks belonging to the JAJO, FMAN and MJSD cycles are, respectively, 131, 68 and 88.

The following example explains how the expiration cycles work. Microsoft happens to belong to the January cycle. In January, the four available expiration months are January, February, April and July. January, April and July are cycle months and February is added so that options are available for the next month. In February, the February options become current month options while March options are introduced. Together with the April and July options, there are still four expiration months available. When March rolls in, the previous April options automatically satisfy the next month requirement but there are only three expiration months available now: March, April and July. As a result, October is added as an expiration month.

As the example shows, the first two expiration months are always the current month and the next month, but the third expiration month can be anywhere between 2 and 4 months from now. Similarly, the nearest fourth expiration month is 5 months from now and the farthest is 7 months from now. For ease of exposition, we label the options expiring in the current month "first options," those expiring in the next month "second options," and so on. Regardless of which cycle a stock belongs to, there are only three possible cases concerning the actual maturity of the third and fourth options. Let TTM stand for time to maturity. Then we have:

	Third options	Fourth options
Case 1	2 < TTM < 3	5 < TTM < 6
Case 2	3 < TTM < 4	6 < TTM < 7
Case 3	4 < TTM < 5	7 < TTM < 8

It is apparent that in Case 1, the second and third options expire in two consecutive months; in Case 2, there is a 1-month gap between the second and third options; in Case 3, the gap is 2 months.

2.3. Summary statistics

The general properties of the bid–ask spreads for equity options are summarized in Table 1. Panel A contains the dollar and proportional bid–ask spreads sorted by option price and moneyness (denned as the ratio of exercise price over the stock price). Controlling for price level, out-of-the-money options have a larger dollar bid–ask spread; controlling for moneyness, the dollar bid–ask spread is clearly higher for options with higher prices. The proportional bid–ask spread also increases as we go from in-the-money to out-of-the-money; controlling for moneyness, the proportional bid–ask spread decreases in the option price. The above observations apply to both calls and puts. That the size of the proportional bid–ask spread is affected by the option's moneyness and price level is surprising, for it implies that the bid–ask spread on a per-dollar basis is not the same across options. An obvious and plausible explanation is the higher market-making cost associated with more leveraged options. In other words, an option with a higher return volatility or leverage will command a bigger spread to compensate the market maker for a higher inventory risk. The option return volatility is indeed inversely related to the option price, hence the observed pattern.[4] Incidentally, the increase of the dollar bid–ask spread in the option price is clearly due to the price level itself. This is why we choose to use the proportional spread to measure liquidity, in contrast to Neal (1987) and George and Longstaff (1993) who examine the dollar spread.

Panel B contains bid–ask spreads sorted by time to maturity and moneyness. Controlling for time to maturity, in-the-money options have a larger dollar spread, seemingly contradicting the patterns in Panel A. In fact, the contradicting pattern is precisely due to the fact that we are not controlling for the price level within each maturity range. The same explanation applies to the increasing pattern of dollar spreads as the maturity increases, controlling for moneyness. In other words, the patterns in the dollar spread are simply due to price levels. The proportional bid–ask spread follows a similar and more evident pattern as in Panel A across moneyness after controlling for time to maturity – the relationship between moneyness and proportional spreads is consistent under different controls while that between moneyess and dollar spreads is not. This is the other reason why we do not focus on the dollar bid–ask spread in this study. As for time to maturity, short-term options have a higher proportional bid–ask spread controlling for moneyness. Again this is due to inventory risk. As pointed out by George and Longstaff (1993) and Chong et al. (2003), short-term options are more risky to make market for, since the chance of early exercise is higher and the gamma risk is higher.

Panel C shows how the spread is related to the option return volatility and the stock return volatility which is the 30-day historical volatility in this study. Unlike Panels A and B, the two-way sorting is done by forming equal-sized quartiles or quintiles. As expected, the proportional bid–ask spread goes up as the option return volatility increases. In contrast, the dollar spread goes down as the return volatility goes up, which again reflects the price level effect – a higher option return volatility is usually related to a lower option price. The result reinforces the need to control for the price level when studying the bid–ask spread. As for the stock return volatility, it is negatively associated with the proportional bid–ask spread, controlling for the option return volatility. As a

[4] In this study, the option return volatility σ_o is defined as the absolute value of the price elasticity times the stock return volatility: $\frac{S}{P}\left|\frac{\partial P}{\partial S}\right|\sigma_s$ where S and P are respectively the stock and option prices, $\left|\frac{\partial P}{\partial S}\right|$ is the absolute value of the option's delta, and σ_s is the stock return volatility.

Table 1

Summary statistics for option bid–ask spreads. This table reports the dollar and proportional bid–ask spreads for options. The sample covers the period of January 1, 1996 to June 30, 2007 and contains 287 stocks that make up the top 100 stocks each year in option trading volume. Proportional bid–ask spread is the dollar spread divided by the average of the bid and ask quotes. Panel A reports the spreads sorted by option price (the average of the bid and ask quotes) and moneyness (exercise price divided by the stock price). Panel B reports the spreads sorted by time to maturity and moneyness. Panel C reports the spreads sorted by the stock return and option return volatilities. The stock return volatility is the annualized standard deviation of the past 30 day returns; the option return volatility is the option price elasticity (in absolute value) times the stock return volatility. The quartiles and quintiles in Panel C are respectively of equal size. "Option r.v." stands for option return volatility. Panel D has the same structure as Panel C except that the sorting is done by option volume and option return volatilities. OptionMetrics reports best bid and ask quotes even if there is no volume, hence the zero volume group.

Panel A: Bid–ask spreads sorted by price and moneyness

		Price range					
		$0.125–2	$2–4	$4–6	$6–8	$8–10	>$10
Call options: dollar bid–ask spread							
Moneyness	0.90–0.95	0.126	0.198	0.256	0.313	0.333	0.560
	0.95–1.00	0.126	0.202	0.266	0.324	0.340	0.598
	1.00–1.05	0.125	0.208	0.273	0.334	0.344	0.620
	1.05–1.10	0.125	0.214	0.277	0.337	0.341	0.641
Call options: proportional bid–ask spread							
Moneyness	0.90–0.95	0.090	0.065	0.052	0.045	0.037	0.034
	0.95–1.00	0.106	0.069	0.054	0.047	0.038	0.035
	1.00–1.05	0.146	0.073	0.056	0.049	0.039	0.035
	1.05–1.10	0.175	0.076	0.057	0.049	0.038	0.036
	No. of obs.	1,726,050	1,512,046	925,186	512,218	285,169	568,548
Put options: dollar bid–ask spread							
Moneyness	1.05–1.10	0.129	0.199	0.261	0.317	0.332	0.565
	1.00–1.05	0.132	0.208	0.272	0.330	0.343	0.609
	0.95–1.00	0.131	0.213	0.273	0.334	0.338	0.639
	0.90–0.95	0.133	0.218	0.275	0.332	0.345	0.659
Put options: proportional bid–ask spread							
Moneyness	1.05–1.10	0.094	0.065	0.053	0.046	0.037	0.033
	1.00–1.05	0.112	0.072	0.055	0.048	0.039	0.035
	0.95–1.00	0.152	0.075	0.056	0.049	0.038	0.036
	0.90–0.95	0.185	0.078	0.057	0.048	0.039	0.038
	No. of obs.	1,943,949	1,565,369	858,747	434,732	223,811	433,643

Panel B: Bid–ask spreads sorted by time to maturity and moneyness

		Time to maturity (days)				
		0–30	31–60	61–90	91–120	121–240
Call options: dollar bid–ask spread						
Moneyness	0.90–0.95	0.269	0.283	0.297	0.311	0.332
	0.95–1.00	0.198	0.236	0.258	0.276	0.305
	1.00–1.05	0.154	0.192	0.219	0.241	0.276
	1.05–1.10	0.147	0.169	0.190	0.210	0.245
Call options: proportional bid–ask spread						
Moneyness	0.90–0.95	0.061	0.056	0.052	0.049	0.046
	0.95–1.00	0.093	0.071	0.062	0.058	0.052
	1.00–1.05	0.172	0.103	0.081	0.071	0.061
	1.05–1.10	0.226	0.159	0.118	0.096	0.075
	No. of obs.	1,069,621	1,370,153	543,520	513,050	2,032,873
Put options: dollar bid–ask spread						
Moneyness	1.05–1.10	0.270	0.276	0.286	0.296	0.308
	1.00–1.05	0.202	0.234	0.251	0.265	0.285
	0.95–1.00	0.159	0.192	0.214	0.233	0.256
	0.90–0.95	0.151	0.170	0.188	0.204	0.229
Put options: proportional bid–ask spread						
Moneyness	1.05–1.10	0.064	0.059	0.055	0.052	0.048
	1.00–1.05	0.101	0.076	0.067	0.062	0.056
	0.95–1.00	0.183	0.111	0.088	0.078	0.067
	0.90–0.95	0.245	0.172	0.129	0.105	0.084
	No. of obs.	1,055,695	1,350,296	537,618	508,511	2,008,131

Panel C: Bid–ask spreads sorted by stock return and option return volatilities

		Stock return volatility				
		0–0.219	0.219–0.302	0.302–0.403	0.403–0.580	0.580+
Call options: dollar bid–ask spread						
Option r.v.	0.000–4.723	0.224	0.288	0.309	0.341	0.424
	4.723–7.987	0.178	0.240	0.264	0.292	0.405
	7.987–16.284	0.148	0.199	0.220	0.239	0.358
	>16.287	0.111	0.143	0.156	0.168	0.240

Table 1 (*continued*)

Panel C: Bid–ask spreads sorted by stock return and option return volatilities

		Stock return volatility				
		0–0.219	0.219–0.302	0.302–0.403	0.403–0.580	0.580+
Call options: proportional bid–ask spread						
Option r.v.	0.000–4.723	0.051	0.052	0.052	0.051	0.044
	4.723–7.987	0.069	0.066	0.062	0.058	0.049
	7.987–16.284	0.088	0.083	0.078	0.073	0.060
	>16.287	0.190	0.173	0.156	0.138	0.105
Put options: dollar bid–ask spread						
Option r.v.	0.000–6.703	0.209	0.262	0.279	0.296	0.337
	6.703–12.121	0.168	0.226	0.247	0.274	0.359
	12.121–24.394	0.145	0.195	0.213	0.233	0.359
	>24.394	0.116	0.152	0.164	0.175	0.252
Put options: proportional bid–ask spread						
Option r.v.	0.000–6.703	0.057	0.057	0.056	0.054	0.047
	6.703–12.121	0.079	0.074	0.069	0.063	0.053
	12.121–24.394	0.102	0.094	0.088	0.080	0.063
	>24.394	0.205	0.193	0.176	0.156	0.115

Panel D: Bid–ask spreads sorted by volume and option return volatility

		Volume (number of contracts)					
		0	1–11	12–37	38–105	106–352	352+
Call options: dollar bid–ask spread							
Option r.v.	0.000–4.723	0.326	0.299	0.261	0.230	0.201	0.169
	4.723–7.987	0.348	0.332	0.286	0.247	0.212	0.172
	7.987–16.284	0.327	0.317	0.276	0.239	0.204	0.158
	>16.287	0.219	0.219	0.200	0.179	0.156	0.119
Call options: proportional bid–ask spread							
Option r.v.	0.000–4.723	0.057	0.050	0.048	0.046	0.045	0.044
	4.723–7.987	0.068	0.058	0.058	0.056	0.054	0.052
	7.987–16.284	0.091	0.075	0.073	0.070	0.067	0.062
	>16.287	0.243	0.172	0.159	0.149	0.136	0.119
		Volume (number of contracts)					
		0	1–10	11–29	30–79	80–266	266+
Put options: dollar bid–ask spread							
Option r.v.	0.000–6.703	0.294	0.251	0.220	0.199	0.178	0.154
	6.703–12.121	0.321	0.280	0.242	0.212	0.187	0.156
	12.121–24.394	0.314	0.284	0.243	0.212	0.185	0.148
	>24.394	0.218	0.212	0.192	0.172	0.151	0.119
Put options: proportional bid–ask spread							
Option r.v.	0.000–6.703	0.061	0.053	0.051	0.050	0.049	0.049
	6.703–12.121	0.071	0.064	0.062	0.061	0.060	0.058
	12.121–24.394	0.091	0.078	0.077	0.077	0.074	0.070
	>24.394	0.218	0.167	0.161	0.157	0.148	0.132

matter of fact, this negative relation is also due to inventory risk, albeit through the volume channel. We postpone the detailed discussions to the regression analysis.

Lastly, in Panel D we sort the spreads by option trading volume and option return volatility. As in Panel C, we also use equal-sized grouping, except that, for volumes, the first group contains all observations with zero-volume quotes. Once again, the proportional spread goes up as the option return volatility increases. As for trading volume, the proportional spread does decrease as the trading volume increases, but the magnitude of changes in the spread is rather modest.

3. Liquidity determinants and trading activities around expiration cycles

3.1. Liquidity determinants

The summary statistics in Table 1 identify the following potential determinants for liquidity or proportional bid–ask spread: time to maturity (*TTM*), moneyness (*M*), option return volatility (σ_o), stock return volatility (σ_s), option trading volume (*V*), and option price (*P*). To help further situate our paper in the literature and develop priors on these liquidity determinants, we briefly survey the literature on bid–ask spreads.

The literature on stock bid–ask spreads is rich. Stoll (1978a) and Ho and Stoll (1981) develop theoretical models for the properties of bid–ask spreads. The models predict that trading volume should reduce spreads (due to economies of scale) while price volatility should increase spreads (due to risk bearing). Roll (1984), French and Roll (1986), and Glosten (1987) derive a positive relationship between price volatility and spreads from a statistical point of view. Empirical evidence generally supports these predictions. For instance, Tinic and West (1972) and Branch and Freed (1977) find that spreads are negatively related to trading volume; Benston and Hagerman (1974) and Branch and Freed (1977) find that spreads are positively related to price volatility. These findings are further confirmed by McInish and Wood (1992) in an intraday setting.

As alluded to earlier, the literature on the property of option bid–ask spreads is extremely thin, with George and Longstaff (1993), Chong et al. (2003) and Cao and Wei (2010) being the only three exceptions. Chong et al. (2003) focus on OTC currency op-

Table 2
Correlations between option liquidity determinants. This table reports the pair-wise correlations for the option liquidity determinants. For each stock, we use the time-series data in the entire sample to calculate pair-wise correlations. We then average the pair-wise correlations across stocks. This is the upper number in each entry in the table. The lower number is the *t*-value for the cross-sectional average. The liquidity variables are defined/calculated as follows: moneyness (*M*) is exercise price over stock price, stock return volatility (σ_s) is the annualized return standard deviation over the past 30 days, option return volatility (σ_o) is the option price elasticity (in absolute value) times the stock return volatility, option volume (*V*) is the number of contracts divided by 100,000, and the option price (*P*) is in dollars. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	M	σ_s	σ_o	V	P
Panel A: Call options					
TTM	0.034	−0.034	−0.245	−0.196	0.437
	18.698	**−19.791**	**−65.198**	**−44.249**	**68.414**
M		0.018	0.300	0.076	−0.440
		11.040	**55.487**	**46.024**	**−38.393**
σ_s			0.064	−0.003	0.173
			16.836	−0.627	**15.695**
σ_o				0.079	−0.274
				30.817	**−40.296**
V					−0.168
					−38.033
Panel B: Put options					
TTM	−0.028	−0.037	−0.238	−0.158	0.399
	−23.312	**−20.662**	**−75.146**	**−38.515**	**69.947**
M		−0.004	−0.331	−0.056	0.478
		−2.854	**−56.044**	**−30.013**	**40.368**
σ_s			0.070	−0.002	0.193
			16.043	−0.452	**17.997**
σ_o				0.064	−0.279
				23.593	**−38.705**
V					−0.138
					−33.195

tions and find a negative relation between maturity and spreads, consistent with our findings in Table 1. Cao and Wei (2010) mostly focus on liquidity commonality and do not study the property of option spreads per se. George and Longstaff (1993) is the only known study on this topic and they examine S&P 100 index options. Although they also find a negative relation between maturity and spreads, they fail to detect any link between volatility and spreads. Neal (1987), who studies competition in equity options, also fails to detect a relation between option spreads and volatility. The main reason is, as we will show later, that both studies examine the dollar bid–ask spreads as opposed to proportional spreads.

In this study, we focus on the proper measure of liquidity: the proportional bid–ask spread. Specifically, to gauge the joint determination of bid–ask spreads, for each stock, we run the following regression using all call or put option observations:

$$PBA = \beta_0 + \beta_1 TTM + \beta_2 M + \beta_3 \sigma_o + \beta_4 \sigma_s + \beta_5 V + \beta_6 P + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon, \tag{3.1}$$

where *PBA* stands for proportional bid–ask spread, $YearDummy_i$ is year dummy for 1996, 1997, …, 2006 and all other variables are defined as before.[5] As discussed earlier, time-to-maturity captures the inventory risk faced by market makers. Short-term options have a higher gamma risk and are more likely to be early exercised. Market makers would set a larger spread per dollar of trading, i.e., a larger proportional spread to compensate for the higher inventory risk. Therefore, consistent with George and Longstaff (1993) and Chong et al. (2003), we expect $\beta_1 < 0$. Moneyness reflects the leverage effect or the effect of option return volatility. For calls, the higher the exercise price, the lower the option price and the higher the return volatility. It is the opposite for puts. We therefore expect $\beta_2 < 0$ for calls and $\beta_3 < 0$ for puts. We have already established that a higher option return volatility leads to a higher inventory risk. Hence we should have $\beta_3 > 0$ for both calls and puts. The stock return volatility and option trading volume affect the proportional bid–ask spread in a related fashion. From the view point of inventory risk, a higher trading volume makes order balancing easier, resulting in a lower proportional spread (Stoll, 1978a; Ho and Stoll, 1981). In other words, we should expect $\beta_5 < 0$. Meantime, a higher volume is usually associated with or driven by a higher stock return volatility, we therefore should expect the sign for the stock return volatility to be the same as that for trading volume: $\beta_4 < 0$. Finally, we would expect $\beta_6 < 0$ for both calls and puts since price and stock return volatility are inversely related. The statistical test is done in a Fama–MacBeth fashion: we run the regression in (3.1) for each stock and then compute the average coefficient and its *t*-value across stocks. The adjusted R^2 is averaged over all stocks.

Before we proceed to the regressions, we need to ascertain to what extent the liquidity determinants are correlated with each other. To this end, we calculate the average pair-wise correlations as follows. We use the time-series data in the entire sample to calculate pair-wise correlations for each stock and then average them across stocks. Table 2 contains the results. First and foremost, option price is highly correlated with time to maturity and moneyness and, to a less extent, with option return volatility. The correlation between option return volatility and time to maturity/moneyness is modest, ranging in absolute value from 0.238 to 0.331. The correlation between option return volatility and stock return volatility is not high (around 0.07), mainly due to the fact that time-series variations in option return volatility are mostly driven by option's price elasticity. Finally, all the other correlations, though significant, are negligible in magnitude.[6]

The above findings indicate that the price variable should be dropped in multivariate regressions since it is correlated with most of the other variables. We should also exercise caution when option return volatility, time to maturity and moneyness appear in the same regression since the first is modestly correlated with the latter two. In Section 5.3, we demonstrate that regression (3.1) is indeed properly specified. Table 3 reports the results for various versions of the regression.

To begin, in the multivariate regression (VIII), all explanatory variables have highly significant coefficients with the expected sign, for both calls and puts. The adjusted R^2 is 55.09% for calls and 57.05% for puts when options are lumped together. When the regression is run separately for in-the-money and out-of-the-money options, the adjusted R^2 remains more or less the same. It is remarkable that a relatively parsimonious model can explain more than half of the time-series variations in liquidity.

The two-variable versions of the regression (I–V) offer some interesting observations. It is apparent that the option return volatility has by far the most explanatory power (the unreported R^2 from univariate regressions for *TTM* is 25.28% (calls) and 27.36% (puts), and for σ_o it is 46.19% and 47.57%, respectively). This is

[5] Here, time to maturity (*TTM*) is in years and trading volume (*V*) is the number of contracts scaled downward by 100,000. Regressions such as (3.1) are not pure time-series regressions since for each day we have a panel of option observations. In this study we loosely refer them as time-series regressions since the sample contains daily observations of more than 11 years. Moreover, for brevity, we omit the indexing of observations.

[6] The near-zero correlation between trading volume and stock return volatility seems to contradict the previous statement that "a higher volume is usually associated with or driven by a higher stock return volatility." Some careful thinking reveals no contradiction after all. Consistent with the literature, we have used historical data (30 days) to estimate the return volatility. However, the type of relationship between volatility and volume discussed in the current context refers to the contemporaneous volatility which could be proxied by the daily return squared. The average correlation between return squared and volume is 0.064 with a *t*-value of 18.127, consistent with our statement regarding option volume and stock return volatility.

Table 3
Option liquidity determinants. This table reports the results from various versions of regression (3.1):

$$PBA = \beta_0 + \beta_1 TTM + \beta_2 M + \beta_3 \sigma_0 + \beta_4 \sigma_s + \beta_5 V + \beta_6 P + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon,$$

where for each stock, the proportional bid–ask spread (*PBA*) is regressed on time to maturity in years (*TTM*), moneyness (*M*) defined as exercise price over stock price, stock return volatility (σ_s) which is the annualized return standard deviation over the past 30 days, option return volatility (σ_o) which is the option price elasticity (in absolute value) times the stock return volatility, option volume (*V*) which is the number of contracts divided by 100,000, the option price (*P*), and the year dummy (*YearDummy*) for 1996–2006. The regression is run separately for calls and puts and the coefficients are averaged across 287 stocks in the sample. For each type of options (calls or puts) we also run separate regressions for in-the-money (ITM) and out-the-money (OTM) options. The *t*-statistic is calculated using the regression coefficients. Each entry consists of the average coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is averaged over the 287 time-series regressions. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	Regression	*TTM*	*M*	σ_o	σ_s	*V*	*P*	Adj R^2 (%)
Panel A: Call options								
All	I	−0.143	0.515					38.35
		−49.415	**32.637**					
	II	−0.084		0.001				50.37
		−43.274		**22.810**				
	III	−0.137			−0.010			25.43
		−50.269			**−4.553**			
	IV	−0.136				0.011		25.40
		−48.379				0.241		
	V	−0.057					−0.015	34.80
		−24.277					**−14.808**	
	VI	−0.084		0.001	−0.039			50.83
		−43.761		**22.343**	**−12.346**			
	VII	−0.084		0.001	−0.039	−0.185		50.93
		−42.130		**22.287**	**−12.330**	**−3.660**		
	VIII	−0.098	0.310	0.001	−0.036	−0.423		55.09
		−44.617	**30.306**	**22.725**	**−12.679**	**−7.463**		
ITM	I	−0.061	0.302					43.48
		−36.130	**32.345**					
	II	−0.033		0.003				48.03
		−27.018		**32.801**				
	III	−0.062			0.000			38.15
		−36.111			0.242			
	IV	−0.062				−0.012		38.14
		−34.659				−0.264		
	V	−0.025					−0.007	44.28
		−15.531					**−13.866**	
	VI	−0.027		0.003	−0.043			49.91
		−27.515		**30.389**	**−18.575**			
	VII	−0.028		0.003	−0.044	−0.437		50.13
		−27.590		**30.888**	**−18.850**	**−6.397**		
	VIII	−0.031	0.077	0.003	−0.040	−0.444		50.53
		−27.145	**12.927**	**29.433**	**−18.109**	**−6.518**		
OTM	I	−0.242	0.734					39.21
		−44.090	**25.653**					
	II	−0.145		0.001				53.37
		−45.677		**23.262**				
	III	−0.232			−0.031			34.79
		−47.927			**−9.216**			
	IV	−0.237				−0.829		34.83
		−48.037				**−12.502**		
	V	−0.118					−0.023	38.45
		−25.131					**−12.018**	
	VI	−0.144		0.001	−0.067			54.03
		−45.940		**23.114**	**−12.567**			
	VII	−0.148		0.001	−0.066	−0.664		54.21
		−45.876		**23.155**	**−12.416**	**−10.524**		
	VIII	−0.158	0.342	0.001	−0.065	−0.617		55.19
		−43.515	**24.487**	**22.999**	**−12.243**	**−9.837**		
Panel B: Put options								
All	I	−0.158	−0.590					41.57
		−49.274	**−36.198**					
	II	−0.096		0.001				51.95
		−46.723		**22.011**				
	III	−0.152			−0.021			27.54
		−51.140			**−8.065**			
	IV	−0.151				−0.024		27.46
		−49.549				−0.454		
	V	−0.071					−0.019	37.71
		−31.708					**−16.700**	

(*continued on next page*)

Table 3 (*continued*)

	Regression	*TTM*	*M*	σ_o	σ_s	*V*	*P*	Adj R^2 (%)
	VI	−0.095		0.001	−0.050			52.54
		−47.611		**21.231**	**−13.936**			
	VII	−0.095		0.001	−0.050	−0.235		52.61
		−45.997		**21.183**	**−13.919**	**−3.206**		
	VIII	−0.110	−0.350	0.001	−0.044	−0.486		57.05
		−45.799	**−34.820**	**21.484**	**−14.103**	**−5.702**		
ITM	I	−0.066	−0.355					45.79
		−36.652	**−24.689**					
	II	−0.043		0.003				50.26
		−29.049		**34.052**				
	III	−0.068			−0.003			39.92
		−36.327			**−1.963**			
	IV	−0.068				0.133		39.83
		−35.336				**4.989**		
	V	−0.030					−0.009	46.52
		−16.737					**−11.119**	
	VI	−0.037		0.003	−0.059			52.87
		−31.106		**28.374**	**−17.426**			
	VII	−0.037		0.003	−0.059	−0.339		52.96
		−31.198		**28.060**	**−17.389**	**−6.935**		
	VIII	−0.038	−0.047	0.003	−0.055	−0.346		53.32
		−28.992	**−3.316**	**28.967**	**−20.298**	**−7.019**		
OTM	I	−0.260	−0.817					42.72
		−46.014	**−32.879**					
	II	−0.161		0.001				55.52
		−45.762		**22.734**				
	III	−0.252			−0.037			38.19
		−48.569			**−8.057**			
	IV	−0.255				−1.010		38.07
		−48.636				**−9.079**		
	V	−0.125					−0.031	42.54
		−26.370					**−13.236**	
	VI	−0.160		0.001	−0.077			56.34
		−46.790		**21.498**	**−14.302**			
	VII	−0.162		0.001	−0.076	−0.848		56.44
		−46.826		**21.512**	**−14.259**	**−7.029**		
	VIII	−0.172	−0.364	0.001	−0.074	−0.787		57.43
		−44.440	**−28.400**	**20.978**	**−14.318**	**−6.542**		

not too surprising since σ_o is an encompassing measure of market-making risk. Despite its overwhelming importance, we can only find two studies in the literature (Neal, 1987; George and Longstaff, 1993) that relate option return volatility to bid–ask spread. Unfortunately neither studies find explanatory power for the option return volatility chiefly because they use the dollar spread to measure option liquidity.

Stock return volatility and option trading volume have the least explanatory power relatively speaking, especially before controlling for the option return volatility. Moneyness and option price have comparable explanatory power, which is somewhere between the explanatory power of option return volatility and that of the stock return volatility or trading volume. In fact, the improvement in the adjusted R^2 is quite minimal by adding moneyness, stock return volatility and trading volume to the option return volatility. The best case is for put options (when in-the-money and out-of-the-money options are combined) where the adjusted R^2 increases from 51.95% to 57.05%.

Finally, as for the separate regressions for in-the-money and out-of-the-money options, time to maturity, moneyness and trading volume all have higher statistical significance for out-of-the-money options while the two return volatilities have higher statistical significance for in-the-money options.

3.2. Trading activities around expiration cycles

As alluded to earlier, the structure of the expiration cycles is a special feature of options, which may have some profound implications for liquidity. Based on the previous descriptions of how the expiration cycles work, the most natural and interesting options to examine would be the second options – options with a maturity beyond the current month but before the next month's expiration date. There are three facts relevant for this set of options: (1) they are always available for all stocks, (2) there are always options available before them, i.e., the current month options, and (3) there may be a 1- or 2-month gap between this set of options and the third options. How far the third options are from the second options may affect the trading pattern and liquidity of the second options. For a visual appreciation of the patterns, we plot in Fig. 1 the average open interest (across all stocks) for the second options against the time around the current month expiration date.[7] We see that the open interest increases steadily as the second options approach the current expiration date; once becoming first options or current month options, the open interest levels off. Fig. 2 plots the trading volume. The increase prior to the expiration date is even more drastic. The last trading day prior to expiry of the first options sees the sharpest increase in volume (from 695 contracts to 887 contracts), reflecting the rebalancing activities of traders. Finally, Fig. 3 presents the plot for the proportional spread. Here we see that the proportional spread increases as the option maturity decreases. It increases more as options are getting near their own maturity. The small dip on Day −1 corresponds to the sharpest increase in volume.

It is the rebalancing activity evident in Fig. 2 that reminds us of a potential substitution effect. When the second options belong to

[7] We only present plots for calls. The patterns are almost identical for puts. Please also note that we index the expiration date (Saturday) as Day 0 and all other days are calendar days relative to Day 0.

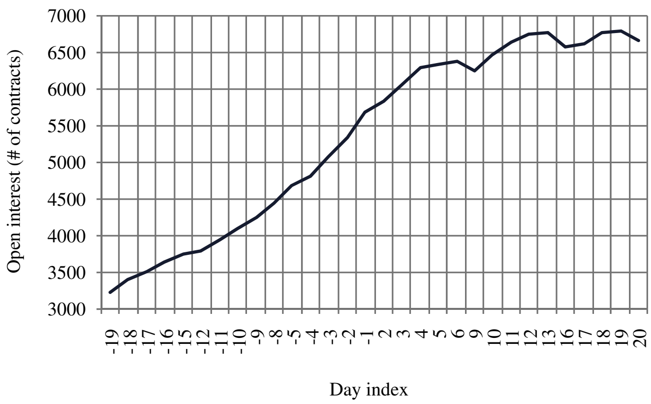

Fig. 1. Call option open interest around expiry date. *Note:* The open expiration date (Saturday) as Day 0 and all other days are calendar days relative to Day 0. In other words, the plot traces the average open interest for the second options as the time to maturity decreases. The sample covers the period of January 1, 1996 to June 30, 2007 and contains 287 stocks that make up the top 100 stocks each year in option trading volume.

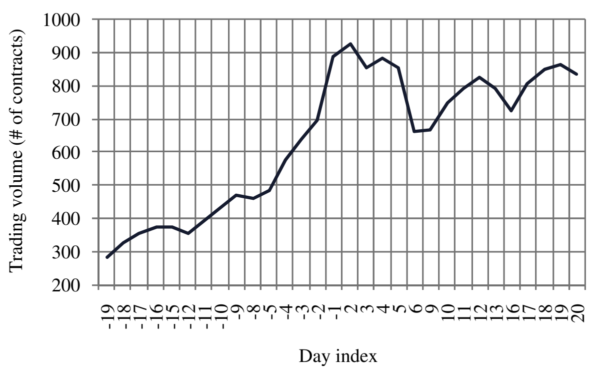

Fig. 2. Call option volume around expiry date. *Note:* The trading volume is averaged across all second options for all stocks. We index the current month's expiration date (Saturday) as Day 0 and all other days are calendar days relative to Day 0. In other words, the plot traces the average trading volume for the second options as the time to maturity decreases. The sample covers the period of January 1, 1996 to June 30, 2007 and contains 287 stocks that make up the top 100 stocks each year in option trading volume.

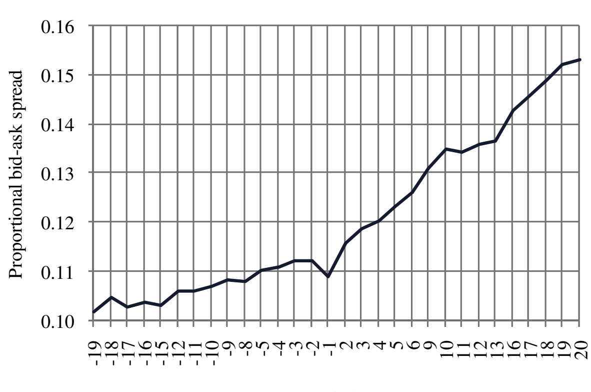

Fig. 3. Call option proportional bid–ask spread around expiry date. *Note:* The proportional bid–ask spread is averaged across all second options for all stocks. We index the current month's expiration date (Saturday) as Day 0 and all other days are calendar days relative to Day 0. In other words, the plot traces the proportional spread for the second options as the time to maturity decreases. The sample covers the period of January 1, 1996 to June 30, 2007 and contains 287 stocks that make up the top 100 stocks each year in option trading volume.

Case I as denned in Section 2, the available maturities cover the current month and the next two immediate months. Traders have more flexibility and can easily substitute between the 2- and 3-month options. As the cycle carries us to Case 2, there is a 1-month gap between the second options and the third. Since the third options now have a maturity of 4 months, it is too much of a stretch to qualify them as short-term options. In this case, conceivably, the second options will have more activities since they now assume some of the roles previously fulfilled by the 3-month options. Carrying this logic to Case 3 leads to the prediction that the second options will see the heaviest activities when they belong to Case 3 – when there is a 2-month gap between the second options and the third.

To verify our conjecture, for the second options of each stock we run the following regression by deleting zero-volume observations:

$$V = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 TTM + \beta_4 M + \beta_5 \sigma_o + \beta_6 \sigma_s + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon, \tag{3.2}$$

where all variables are defined as before except that D_1 and D_2 are dummy variables indicating if the current observation belongs to Case 2 or Case 3 (Case 1 is the base-case). We also run the regression for open interest. If our conjecture turns out to be true, then the coefficients for D_1 and D_2 should be positive and significant, and the coefficient of D_2 should be bigger than that of D_1. Table 4 reports the results.

Our conjecture is supported by the empirical findings. For the volume regressions, the coefficient for D_1 is positive and significant in all cases once we control for all other explanatory variables; the coefficient for D_2 is positive and significant for call options; for put options, the D_2 coefficient is positive in all regressions but not significant in the full regression. We also test $D_2 = D_1$. It is seen that the t-value of this test is significant at the 5% level for call options (indicating that $D_2 > D_1$), but is insignificant for puts. The results from the open interest regressions are much stronger. The coefficient for D_1 is significant in most cases and that for D_2 is significant in all cases. Furthermore, the t-value for the test $D_2 > D_1$ is highly significant and the adjusted R^2 is much higher compared with volume regressions. Taken together, the empirical evidence supports the notion that investors use 2-month options to fulfil the role otherwise fulfilled by 3-month or 4-month options. In other words, there is a substitution among options with different maturities due to expiration cycles.

The signs for *TTM* and *M* are as expected: longer maturity options and away-from-the-money options trade less. The option return volatility has the same sign as moneyness for calls and the opposite sign for puts precisely because of the relation between the two variables. For calls, moneyness and option return volatility are positively related; for puts they are negatively related. More interesting is the sign for the stock return volatility. It is negative for in-the-money options and positive for out-of-the-money options, and this is true for both calls and puts. When the stock return volatility is high, the trading of out-of-the-money options becomes more active while in-the-money options become less active. There seems to be a shift in trading from in-the-money to out-of-the-money options. In other words, a higher stock return volatility induces a substitution among options with differing exercise prices. In this study, we remain agnostic about the cause for the moneyness-substitution. One potential reason is, a higher volatility stimulates more speculative trading in options, and out-of-the-money options provide a higher leverage and hence attract more interest.

Table 4
Maturity substitution in trading volume and open interest. This table reports the results from various versions of regression (3.2):

$$V = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 TTM + \beta_4 M + \beta_5 \sigma_o + \beta_6 \sigma_s + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon,$$

where for each stock, the option volume (number of contracts) for the second options (V) is regressed on dummy variable D_1 which takes the value of one when the option belongs to Case 2, dummy variable D_2 which takes the value of one when the option belongs to Case 3 (please see the text for definitions of the second options and Cases 2 and 3), time to maturity in days (TTM), moneyness (M) defined as exercise price over stock price, stock return volatility (σ_s) which is the annualized return standard deviation over the past 30 days, option return volatility (σ_o) which is the option price elasticity (in absolute value) times the stock return volatility, and the year dummy (*YearDummy*) for 1996–2006. The regression is run separately for in-the-money (ITM) and out-of-the-money (OTM) calls and puts. The coefficients are averaged across 287 stocks in the sample. The *t*-statistic is calculated using the regression coefficients. Each entry consists of the average coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is averaged over the 287 time-series regressions. We also report the *t*-statistic for the test $D_2 = D_1$. Regression (3.2) is repeated by replacing the trading volume by open interest. To conserve space, we only report the dummy variable coefficients, the *t*-value for the inequality test and the average adjusted R^2. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	Regression	Volume								Open interest			
		D_1	D_2	TTM	M	σ_o	σ_s	$D_2 = D_1$	Adj R^2 (%)	D_1	D_2	$D_2 = D_1$	Adj R^2 (%)
Panel A: Call options													
ITM	I	8.046	63.625					**2.934**	5.49	111.1	3916.3	**11.897**	31.51
		0.573	**5.000**							0.938	**13.180**		
	II	15.174	80.184	−12.056	4583.3			**3.326**	9.00	129.6	3943.0	**11.786**	35.21
		1.034	**6.208**	**−12.813**	**7.224**					1.081	**13.119**		
	III	27.388	90.330	−10.115		53.911	−592.991	**3.246**	9.68	170.4	4042.8	**11.475**	35.94
		1.844	**7.246**	**−11.532**		**6.378**	**−3.055**			1.473	**12.753**		
	IV	30.239	93.140	−10.731	1359.5	41.085	−407.841	**3.372**	9.96	157.8	4021.8	**11.540**	36.35
		2.117	**7.759**	**−11.571**	**3.981**	**4.624**	−1.833			1.343	**12.827**		
OTM	I	69.556	145.597					**3.030**	9.90	313.4	4606.6	**11.413**	37.40
		4.869	**7.056**							**2.291**	**13.146**		
	II	78.458	148.093	−15.421	−3226.8			**2.366**	12.06	374.7	4594.4	**11.395**	40.85
		4.704	**6.108**	**−13.424**	**−3.967**					**2.593**	**13.474**		
	III	80.572	134.414	−18.601		−2.886	869.514	**2.050**	12.71	362.8	4595.7	**11.430**	41.43
		4.632	**6.830**	**−11.902**		**−8.830**	**4.568**			**2.483**	**13.505**		
	IV	78.888	136.858	−18.054	−1552.5	−3.483	766.696	**2.062**	12.86	358.4	4601.3	**11.408**	41.69
		4.468	**6.257**	**−12.888**	−1.380	**−4.976**	**4.769**			**2.497**	**13.411**		
Panel B: Put options													
ITM	I	31.640	42.472					0.789	4.40	196.6	2158.5	**10.855**	27.42
		3.935	**3.814**							**2.533**	**13.224**		
	II	35.686	36.090	−8.350	−4339.6			0.016	7.38	236.2	2137.7	**10.591**	31.45
		4.215	1.480	**−6.668**	**−4.650**					**2.905**	**13.355**		
	III	43.750	42.689	−7.378		33.322	−543.482	−0.040	7.67	288.6	2190.3	**9.992**	32.09
		4.271	1.752	**−6.343**		**4.273**	**−2.349**			**3.108**	**13.183**		
	IV	40.064	40.551	−8.066	−2404.1	16.071	−189.973	0.019	8.02	266.7	2174.0	**10.116**	32.77
		4.123	1.667	**−6.911**	**−5.835**	**2.200**	−0.907			**2.921**	**13.178**		
OTM	I	55.068	71.851					0.701	6.33	199.6	3483.5	**9.282**	34.03
		3.249	**4.250**							**2.172**	**10.196**		
	II	62.936	47.557	−14.530	3660.8			−0.374	8.16	248.2	3327.4	**9.615**	37.15
		3.619	1.275	**−5.100**	**2.818**					**2.659**	**10.862**		
	III	56.280	65.022	−17.854		−1.713	843.729	0.246	8.67	232.5	3452.9	**9.418**	37.87
		3.329	**2.078**	**−4.598**		**−6.214**	**3.360**			**2.465**	**10.506**		
	IV	54.893	51.102	−16.380	2011.9	−1.714	563.047	−0.085	8.81	222.1	3419.0	**9.387**	38.18
		3.310	1.237	**−5.236**	1.680	**−4.273**	**2.338**			**2.348**	**10.450**		

4. Option liquidity versus substitution effects

4.1. Time-series behavior of liquidity due to substitutions

The substitution phenomena uncovered above have direct implications for option liquidity, the ultimate focus of this paper. Let us begin with the moneyness-substitution. When the stock return volatility is high, trading shifts from in-the-money options to out-of-the-money options. The higher volume in out-of-the-money options reduces the inventory risk for market makers, making it possible for them to lower the proportional spread; the exact opposite applies to in-the-money options. Hence the following hypothesis:

Hypothesis 1. The difference in the proportional bid–ask spread between out-of-the-money options and in-the-money options is negatively related to the stock return volatility.

We employ the following regression to test Hypothesis 1:

$$PBA_{diff} = \beta_0 + \beta_1 TTM + \beta_2 \sigma_s + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon, \quad (4.1)$$

where PBA_{diff} is the difference between the proportional bid–ask spreads of out-of-the-money and in-the-money options. We run the regression separately for the first, second, and third plus fourth options; we also run it for all options combined. As before, the regression in (4.1) is run for each stock and the coefficients are averaged across stocks. We note that each observation of PBA_{diff} is averaged from multiple observations having the same time to maturity but different exercise prices. We will have an observation for a particular maturity only when we have both in-the-money and out-of-the-money observations for that maturity. Table 5 contains the results.

The hypothesis is strongly supported by the regression results for both calls and puts. The negative relationship between the spread differential and the stock return volatility is significant for all maturity groups. That in-the-money and out-of-the-money options tend to have different liquidity levels due to leverage is well known; but the finding that the difference is also a function of the

Table 5
Impact of moneyness-substitution on option liquidity. This table reports the results from regression (4.1):

$$PBA_{diff} = \beta_0 + \beta_1 TTM + \beta_2 \sigma_s + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon,$$

where for each category of options (please see the text for definition of first options, second options, and so on), the difference between the proportional bid–ask spreads of out-of-the-money and in-the-money options (PBA_{diff}) is regressed on time to maturity in days (*TTM*), stock return volatility (σ_s) which is the annualized return standard deviation over the past 30 days, and the year dummy (*YearDummy*) for 1996–2006. The spread differential PBA_{diff} is calculated by averaging over all the spreads available for a particular maturity. It has a valid observation for a particular maturity only when there are observations for both in-the-money and out-of-the-money options. The regression coefficients are averaged across 287 stocks in the sample. The *t*-statistic is calculated using the regression coefficients. Each entry consists of the average coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is averaged over the 287 time-series regressions. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	Call options			Put options		
	TTM	σ_s	Adj R^2 (%)	*TTM*	σ_s	Adj R^2 (%)
1st options	−0.0050	−0.0720	25.66	−0.0052	−0.0793	29.55
	−42.100	**−12.946**		**−42.105**	**−9.865**	
2nd options	−0.0012	−0.0511	27.72	−0.0013	−0.0524	33.55
	−23.424	**−9.259**		**−24.005**	**−6.021**	
3rd & 4th options	−0.0002	−0.0151	22.29	−0.0002	−0.0176	28.55
	−19.016	**−5.829**		**−20.753**	**−7.179**	
All options	−0.0005	−0.0349	27.30	−0.0006	−0.0386	31.09
	−39.177	**−12.509**		**−40.139**	**−9.552**	

stock return volatility is new. The shifting in trading volume from in-the-money to out-of-the-money options as the volatility goes up narrows the spreads of the latter and widens those of the former through the inventory risk channel.

The moneyness-substitution effect is stronger for short-term options as indicated by the magnitude of the volatility coefficient. The current month options have the strongest moneyness-substitution effect. Take put options as an illustration. The volatility coefficient is −0.079, meaning that when the stock return volatility goes up by, say 5%, the PBA of out-of-the-money options will go down by about 40 basis points relative to the PBA of in-the-money options ($0.079 \times 0.05 \doteq 40$ bp).

The reasoning behind Hypothesis 1 can also be applied to the maturity-substitution. Specifically, the market-making cost due to inventory risk for the second options should rank in descending order for Cases 1, 2 and 3. Hence the following hypothesis:

Hypothesis 2. The proportional bid–ask spread for the second options depends on how far apart they are from the third options. It is the largest when the third options are immediately next to the second options; it is the smallest when the third options are 2 months away.

Hypothesis 2 can be tested based on the regression structures for Tables 2 and 3. Specifically, we add the two dummy variables in (3.2) into regression (3.1) to have the following regression for each stock:

$$PBA = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 TTM + \beta_4 M + \beta_5 \sigma_o + \beta_6 \sigma_s + \beta_7 V + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon. \quad (4.2)$$

As before, we average the coefficients and perform statistical tests accordingly. Table 6 contains the results.

First and foremost, the coefficients for D_1 and D_2 are negative and significant in all cases, supporting the first half of Hypothesis 2 – the spread of the second options is lower when the third options are not available for the month immediately following the expiry month of the second options. Before controlling for liquidity determinants, the two dummies explain up to 40% of the time-series variations in the proportional spreads, especially when in-the-money and out-of-the-money options are examined separately. The statistical significance for the two dummies remains after controlling for all liquidity determinants.

Second, as for the second half of the prediction in the hypothesis, viz., $D_2 < D_1$ (or $|D_2 > D_1|$), the support is not as strong. The *t*-value for the test $D_2 = D_1$ is significant only for some cases of call options. For all other cases, the relation $D_2 < D_1$ is still true albeit statistically insignificant. Here, many of the *t*-values would have been significant had we tested at the 10% or for one-tail only. Regardless, the important conclusion is, the maturity-substitution does affect the second options' liquidity. The fact that traders use the second options to fulfil the role of the third options (when they are far away) increases the trading volume of the second options and thus narrows their spread through the inventory risk channel.

Finally, it is important to note the economic significance of this substitution effect. As apparent in Table 6, the coefficients for D_1 and D_2 in the full regression range from −0.0006 to −0.0043 with an average of about −0.0020 or 20 basis points. This is by no means a large portion of the percentage bid–ask spread for a typical option, which is 8.6% in our sample. Should the liquidity impact be much larger, a trader could time his/her trades to save liquidity costs (e.g., trade the 2-month options only when the third options are not available for the immediate following month). A differential of 20 basis points does not seem to be worth the while to time the trades. Therefore our results have more implications for liquidity studies than trading. The same perhaps can be said about the moneyness-substitution effect.

At this point, it is natural to ask if a similar substitution effect also exists for the first options. Conceivably, when traders substitute for the third options, they may also shift some of the volume to the first options. If this is the case, the first options' spread should narrow when the third options are not in the third month. To verify this conjecture, we repeat the regression in (4.2) for the first options. Table 7 contains the results.

Although the impact of substitution on spreads still exists, it is limited to Case 2 only – when the third options are in the fourth month or when there is only 1 month gap between the second and third options. The statistical significance is also weaker, especially for put options. The intuitive explanation for the overall findings is as follows. When the third options are in the fourth month, some volume is shifted to both the first and the second options since they are reasonably close; but when the third options are in the fifth month, the first options are too short to be reasonable substitutes hence their spreads are not affected. In fact, the results in Table 4 lend indirect support to this reasoning since the coefficient of D_2 is much larger than that of D_1. In other words, the shifting of volume mostly goes to the second options when the third options are the farthest.

Table 6
Impact of maturity-substitution on option liquidity – second options. This table reports the results from various versions of regression (4.2):

$$PBA = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 TTM + \beta_4 M + \beta_5 \sigma_o + \beta_6 \sigma_s + \beta_7 V + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon,$$

where for each stock, the proportional bid–ask spread (*PBA*) of the second options is regressed on dummy variable D_1 which takes the value of one when the option belongs to Case 2, dummy variable D_2 which takes the value of one when the option belongs to Case 3 (please see the text for definitions of the second options and Cases 2 and 3), and all the determinant variables appearing in Table 3 other than option price (*P*). The regression is run separately for calls and puts and the coefficients are averaged across 287 stocks in the sample. For each type of options (calls or puts) we also run separate regressions for in-the-money (ITM) and out-of-the-money (OTM) options. The *t*-statistic is calculated using the regression coefficients. Each entry consists of the average coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is averaged over the 287 time-series regressions. We also report the *t*-statistic for the test $D_1 = D_2$. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	Regression	D_1	D_2	TTM	M	σ_o	σ_s	V	$D_1 = D_2$	Adj R^2 (%)
Panel A: Call options										
All	I	−0.0018	−0.0045						**3.531**	23.66
		−3.791	**−7.672**							
	II	−0.0019	−0.0034	−0.288	0.700				**2.259**	48.00
		−4.543	**−6.630**	**−28.930**	**25.928**					
	III	−0.0010	−0.0025	−0.129	0.328	0.0012	−0.056	−0.597	**2.615**	60.00
		−3.086	**−5.572**	**−16.063**	**19.208**	**23.707**	**−14.696**	**−9.345**		
ITM	I	−0.0009	−0.0024						**3.941**	39.66
		−3.709	**−8.982**							
	II	−0.0009	−0.0020	−0.086	0.298				**3.309**	47.35
		−3.625	**−8.033**	**−19.382**	**30.165**					
	III	−0.0006	−0.0016	−0.072	0.199	0.0013	−0.020	−0.389	**2.841**	48.48
		−2.724	**−6.449**	**−14.648**	**19.535**	**12.163**	**−8.769**	**−7.642**		
OTM	I	−0.0031	−0.0050						1.457	32.39
		−3.983	**−4.918**							
	II	−0.0038	−0.0057	−0.554	1.335				1.546	49.60
		−5.538	**−5.887**	**−23.971**	**18.936**					
	III	−0.0021	−0.0039	−0.267	0.583	0.0010	−0.095	−0.624	1.733	58.47
		−3.554	**−4.517**	**−15.107**	**13.893**	**22.856**	**−10.994**	**−9.189**		
Panel B: Put options										
All	I	−0.0024	−0.0033						0.967	25.89
		−4.051	**−5.095**							
	II	−0.0025	−0.0029	−0.332	−0.796				0.561	51.95
		−4.991	**−5.066**	**−28.135**	**−27.669**					
	III	−0.0015	−0.0026	−0.155	−0.362	0.0010	−0.068	−0.689	1.631	62.83
		−3.623	**−5.114**	**−18.702**	**−19.087**	**24.389**	**−14.977**	**−9.117**		
ITM	I	−0.0009	−0.0018						1.920	42.18
		−2.585	**−4.956**							
	II	−0.0008	−0.0013	−0.101	−0.353				1.262	50.69
		−2.283	**−4.522**	**−19.301**	**−25.533**					
	III	−0.0006	−0.0010	−0.086	−0.219	0.0012	−0.026	−0.351	1.032	51.95
		−2.056	**−3.861**	**−18.496**	**−13.090**	**10.022**	**−10.227**	**−5.393**		
OTM	I	−0.0042	−0.0045						0.208	36.45
		−4.858	**−4.448**							
	II	−0.0044	−0.0051	−0.602	−1.402				0.542	53.65
		−4.922	**−5.174**	**−28.027**	**−25.707**					
	III	−0.0025	−0.0043	−0.293	−0.557	0.0009	−0.108	−0.771	1.791	61.57
		−3.726	**−5.279**	**−19.079**	**−16.454**	**20.498**	**−14.016**	**−9.201**		

4.2. Cross-sectional determinants of the substitution effects

In this section, we attempt to relate the substitution effects to firm characteristics. Presumably, certain type of stocks are more prone to substitution effects than others. In the absence of theoretical guidance, we can think of the following characteristics that may potentially affect the extent of substitution: the ratio of stock volume to option volume (V_{ratio}), stock return volatility (σ_s), option's proportional bid–ask spread (*PBA*), firm size (*SIZE*) – stock price times the number of shares outstanding, share turnover (*TO*) – stock trading volume divided by the number of shares outstanding, and probability of informed trading (*PIN*). Moreover, which expiration cycle the stock belongs to may also have an impact. With the above in mind, we run the following cross-section regression:

$$D\_coef = \beta_0 + \beta_1 CD_1 + \beta_2 CD_2 + \beta_3 V_{ratio} + \beta_4 \sigma_s + \beta_5 PBA + \beta_6 SIZE + \beta_7 TO + \beta_8 PIN + \varepsilon, \quad (4.3)$$

where *D_coef* is either the coefficient for the dummy variables (D_1 or D_2) in (4.2) or the coefficient for the volatility variable in (4.1), and CD_1 and CD_2 are dummy variables that indicate respectively the JAJO and FMAN cycles (MJSD is the base cycle). For options, we average the observations (option volume and *PBA*) within the day to obtain daily aggregate measures. We then compute the daily value for all independent variables (e.g., V_{ratio} would be the daily stock volume divided by the average option volume). The daily observations for each variable are then averaged within the sample to obtain an aggregate measure.[8] The variable *SIZE* is scaled downward by 10^{10}.

In the absence of a known model that explains the substitution effects, we can only speculate on the general relationship of each explanation variable to the substitution effect. The impact of the

[8] We obtain the quarterly PIN measures from Stephen Brown, the University of Maryland.

Table 7
Impact of maturity-substitution on option liquidity – first options. This table reports the results from various versions of regression (4.2):

$$PBA = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 TTM + \beta_4 M + \beta_5 \sigma_o + \beta_6 \sigma_s + \beta_7 V + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon,$$

where for each stock, the proportional bid–ask spread (*PBA*) of the first options is regressed on dummy variable D_1 which takes the value of one when the option belongs to Case 2, dummy variable D_2 which takes the value of one when the option belongs to Case 3 (please see the text for definitions of the second options and Cases 2 and 3), and all the determinant variables appearing in Table 3 other than option price (*P*). The regression is run separately for calls and puts and the coefficients are averaged across 287 stocks in the sample. For each type of options (calls or puts) we also run separate regressions for in-the-money (ITM) and out-of-the-money (OTM) options. The *t*-statistic is calculated using the regression coefficients. Each entry consists of the average coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is averaged over the 287 time-series regressions. We also report the *t*-statistic for the test $D_2 = D_1$. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	Regression	D_1	D_2	TTM	M	σ_o	σ_s	V	$D_2 = D_1$	Adj R^2 (%)
Panel A: Call options										
All	I	−0.0020	0.0004						**2.698**	13.78
		−3.088	0.687							
	II	−0.0026	0.0003	−1.200	1.320				**3.566**	47.39
		−4.392	0.581	**−47.642**	**37.306**					
	III	−0.0024	0.0000	−0.923	0.887	0.0004	−0.058	−0.830	**3.058**	57.41
		−4.223	0.050	**−43.410**	**36.502**	**23.153**	**−15.523**	**−9.679**		
ITM	I	−0.0016	0.0010						**3.401**	28.14
		−2.957	1.838							
	II	−0.0017	0.0004	−0.504	0.657				**3.659**	45.99
		−4.014	1.009	**−32.077**	**31.196**					
	III	−0.0013	0.0005	−0.321	0.242	0.0025	−0.054	−0.481	**3.510**	51.55
		−3.352	1.516	**−27.465**	**16.187**	**26.464**	**−17.194**	**−8.156**		
OTM	I	−0.0023	0.0022						**2.578**	20.24
		−1.867	1.780							
	II	−0.0057	0.0018	−2.692	2.331				**4.351**	45.39
		−4.929	1.425	**−47.339**	**22.743**					
	III	−0.0044	0.0006	−1.937	1.318	0.0003	−0.123	−0.946	**3.227**	53.20
		−4.301	0.531	**−43.919**	**27.067**	**20.737**	**−13.318**	**−10.316**		
Panel B: Put options										
All	I	−0.0004	0.0023						**2.668**	14.63
		−0.540	**3.245**							
	II	−0.0014	0.0022	−1.283	−1.480				**3.588**	49.85
		−2.112	**2.919**	**−45.422**	**−37.059**					
	III	−0.0013	0.0017	−0.988	−1.009	0.0003	−0.068	−1.063	**3.149**	58.71
		−2.074	**2.369**	**−42.443**	**−37.525**	**20.666**	**−15.095**	**−8.880**		
ITM	I	−0.0011	0.0005						**2.020**	29.27
		−2.235	0.814							
	II	−0.0007	0.0003	−0.546	−0.761				1.756	48.44
		−1.752	0.726	**−32.741**	**−30.475**					
	III	−0.0005	0.0005	−0.353	−0.230	0.0025	−0.065	−0.404	1.757	54.67
		−1.246	1.241	**−29.557**	**−9.503**	**20.792**	**−18.062**	**−6.896**		
OTM	I	0.0000	0.0063						**3.259**	22.81
		−0.032	**4.242**							
	II	−0.0034	0.0058	−2.791	−2.393				**4.464**	48.01
		−2.624	**3.618**	**−49.601**	**−32.304**					
	III	−0.0027	0.0040	−2.045	−1.394	0.0003	−0.134	−1.305	**3.359**	55.07
		−2.203	**2.542**	**−46.114**	**−31.602**	**16.617**	**−14.556**	**−9.354**		

proportional spread (*PBA*) is relatively easy to conjecture – a larger spread will most likely be associated with a stronger substitution effect. The volume ratio V_{ratio} measures the relative activity of the stock and options. For stocks with less active trading in options (i.e., higher V_{ratio}), we may not expect a strong substitution effect since the options are not of much interest in the first place. The share turnover variable (*TO*) acts in the same fashion as V_{ratio} in this context. Other things being equal (i.e., controlling for the relative size of the two markets, V_{ratio}), a higher share turnover means less reliance on the option market, and thus a weaker substitution effect. The remaining three variables (σ_s, *SIZE* and *PIN*) are somewhat related, especially *SIZE* and *PIN*. It is well known that smaller firms tend to have a higher *PIN*, both of which indicate a higher degree of information asymmetry. The prediction of how informed trading impacts the substitution effect can go both ways. On the one hand, insofar as the spread is more sensitive to volume for stocks with more information asymmetry, we may expect smaller firms or stocks with a higher *PIN* to exhibit a stronger substitution effect. On the other hand, options on large firm stocks are more likely traded for non-informational reasons (e.g., differences of opinion) and as a result the trading interest may cover all maturity ranges. In this case, a gap in maturities will have a bigger impact on the spread, i.e., a stronger substitution effect (the same logic applies to the moneyness-substitution). The prediction for the stock return volatility is likewise ambiguous since volatility and firm size are closely related.

We first run the regression in (4.3) for the volatility coefficient from (4.1). To enhance testing power, we pool the call and put options by calculating *PBA* separately for the two types of options. After merging our dataset with the *PIN* measures, we are left with 255 stocks. Therefore the pooled regression has 510 observations. As in (4.1), we run four versions of the regression in (4.3): all options combined, the first options only, the second options only, and the third and fourth options only. Table 8 reports the result. For brevity, we report the full results only for the case of all options combined. The reduced versions of regressions (i.e., the two dummies plus each explanatory variable separately) for all other cases are similar.

Table 8
Cross-section determinants of moneyness-substitution effect. This table reports the results from various versions of cross-section regression (4.3):

$$D_{coef} = \beta_0 + \beta_1 CD_1 + \beta_2 CD_2 + \beta_3 V_{ratio} + \beta_4 \sigma_s + \beta_5 PBA + \beta_6 SIZE + \beta_7 TO + \beta_8 PIN + \varepsilon,$$

where D_{coef} is the coefficient of the stock return volatility in regression (4.1), CD_1 a dummy variable indicating the January expiration cycle, CD_2 a dummy variable indicating the February expiration cycle, V_{ratio} the stock volume over the option volume, σ_s the stock return volatility, *PBA* the option's proportional bid–ask spread, *SIZE* the stock price times the number of shares outstanding (scaled downward by 10^{10}), *TO* the stock's turnover which is stock volume over the number of shares outstanding, and *PIN* the probability of informed trading. For options, we first calculate the daily average of proportional spread and trading volume. Then, for all explanatory variables we calculate the sample average to be used in the cross-section regression. To enhance testing power, we pool call and put options. Corresponding to Table 5, the regression is run for all options combined, the first options only (1st), the second options only (2nd), and the third and fourth options (Other). All versions of regressions are reported for the case of all options combined. For other cases, we only report the result for the full regression (version VII). Each entry consists of the regression coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is in the last column. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	Regression	CD_1	CD_2	V_{ratio}	σ_s	*PBA*	*SIZE*	*TO*	*PIN*	Adj R^2 (%)
All options	I	−0.003	0.013	−0.113						1.82
		−0.637	**2.605**	−0.258						
	II	0.002	0.010		0.112					29.52
		0.544	**2.316**		**14.104**					
	III	−0.003	0.013			−0.066				2.21
		−0.706	**2.679**			−1.441				
	IV	−0.002	0.011				−2.093			8.63
		−0.580	**2.364**				**−6.145**			
	V	0.001	0.016					5.503		8.61
		0.321	**3.338**					**6.136**		
	VI	0.000	0.012						0.251	12.47
		0.092	**2.638**						**7.849**	
	VII	0.001	0.009	−0.125	0.107	−0.138	−0.915	−0.106	0.017	30.85
		0.264	**2.196**	−0.331	**9.455**	**−3.202**	**−2.696**	−1.134	0.426	
1st	VII	0.012	0.019	−1.186	0.152	−0.261	−0.729	0.005	0.230	25.88
		1.743	**2.308**	−1.595	**6.800**	**−3.078**	−1.086	0.028	**2.923**	
2nd	VII	0.003	0.008	−0.009	0.154	−0.210	−1.051	−0.057	0.152	36.27
		0.630	1.416	−0.017	**9.462**	**−3.409**	**−2.154**	−0.423	**2.660**	
Other	VII	0.011	0.016	0.286	0.121	−0.232	−0.558	−0.170	−0.066	20.94
		2.362	**3.053**	0.587	**8.256**	**−4.175**	−1.269	−1.398	−1.284	

Table 9
Cross-section determinants of maturity-substitution effect. This table reports the results from various versions of cross-section regression (4.3):

$$D_{coef} = \beta_0 + \beta_1 CD_1 + \beta_2 CD_2 + \beta_3 V_{ratio} + \beta_4 \sigma_s + \beta_5 PBA + \beta_6 SIZE + \beta_7 TO + \beta_8 PIN + \varepsilon,$$

where D_{coef} is the coefficient of the dummy variable D_2 in regression (4.2), CD_1 a dummy variable indicating the January expiration cycle, CD_2 a dummy variable indicating the February expiration cycle, V_{ratio} the stock volume over the option volume, σ_s the stock return volatility, *PBA* the option's proportional bid–ask spread, *SIZE* the stock price times the number of shares outstanding (scaled downward by 10^{10}), *TO* the stock's turnover which is stock volume over the number of shares outstanding, and *PIN* is the probability of informed trading. For options, we first calculate the daily average of proportional spread and trading volume. Then, for all explanatory variables we calculate the sample average to be used in the cross-section regression. To enhance testing power, we pool call and put options, which is done in three different ways: (1) all options combined, (2) in-the-money calls with out-of-the-money puts (ICOP), and (3) out-of-the-money calls with in-the-money puts (OCIP). All versions of regressions are reported for the case of all options combined. For other cases, we only report the result for the full regression (version VII). Each entry consists of the regression coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is in the last column. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	Regression	CD_1	CD_2	V_{ratio}	σ_s	*PBA*	*SIZE*	*TO*	*PIN*	Adj R^2 (%)
All options	I	0.000	−0.002	0.065						1.63
		−0.133	**−2.803**	0.967						
	II	0.000	−0.002		0.005					3.91
		0.258	**−3.024**		**3.599**					
	III	0.000	−0.002			−0.037				6.95
		−0.275	**−2.605**			**−5.470**				
	IV	0.000	−0.002				0.018			1.47
		−0.081	**−2.778**				0.330			
	V	0.000	−0.002					0.570		4.57
		0.580	**−2.396**					**4.069**		
	VI	0.000	−0.002						0.007	1.84
		0.066	**−2.827**						1.422	
	VII	0.000	−0.002	0.050	0.004	−0.037	0.048	0.024	0.007	9.61
		0.324	**−2.530**	0.754	1.829	**−4.903**	0.818	1.476	0.938	
ICOP	VII	0.000	−0.002	0.028	0.006	−0.055	−0.013	0.018	−0.008	14.15
		0.247	**−2.143**	0.331	**2.338**	**−7.357**	−0.167	0.885	−0.890	
OCIP	VII	−0.002	−0.003	0.060	−0.053	−0.053	0.027	0.045	0.024	11.86
		−2.263	**−3.115**	0.746	−1.937	**−6.317**	0.368	**2.285**	**2.947**	

To begin, the volume ratio (V_{ratio}) has no explanatory power throughout and the turnover variable (*TO*) also loses its explanatory power once we control for other variables. In the reduced regression (version V), the share turnover does have a significant positive coefficient, confirming our previous conjecture (keep in mind that the dependent variable is negative). As for σ_s, *SIZE* and *PIN*, it turns out that larger firms with a lower volatility and a lower *PIN* have the strongest substitution effect, indicating that options on these stocks are mostly used for non-informational trading. It should be noted that the statistical significance for the three variables largely remains in the full regressions. As for the spread variable *PBA*, it is not significant in the reduced regression but becomes negative and significant once other variables are controlled for. The negative coefficient confirms our conjecture – the substitution effect is stronger when the average spread itself is larger. Finally, it appears that the February cycle options exhibit the least moneyness-substitution effect compared with the options of the other two cycles.

Next, we run the regression in (4.3) for the dummy variable coefficients from (4.2). Here the regression is also run by pooling calls and puts and is run for three permutations in terms of the specification of the dependent variable: coefficient of D_1, coefficient of D_2 , and the sum of coefficients of D_1 and D_2. The pooling is also done in three ways: all calls and puts, in-the-money calls and out-of-the-money puts, and out-of-the-money calls and in-the-money puts. Since the results across alternative dependent-variable specifications are similar, we only report in Table 9 the results for the coefficient of D_2.

The results are generally similar to those in Table 8. For instance, the coefficient for *PBA* is negative and significant; V_{ratio} has no explanatory power; the coefficients for σ_s and *PIN* are positive when significant. However, firm size no longer has explanation power and σ_s and *PIN* have explanatory power in only a few cases. The adjusted R^2 for each regression is also lower compared with their counterparts in Table 8. It is interesting to observe that the substitution effect for out-of-the-money calls and in-the-money puts exhibits the most sensitivity to the explanatory variables (in terms of *t*-values). Finally, in contrast to Table 8, the dummy variable CD_2 has a significant, negative coefficient indicating that the February cycle options exhibit a stronger maturity-substitution effect than options of the other two cycles. That the February cycle options exhibit a weaker moneyness-substitution effect but a stronger maturity-substitution effect remains a puzzle to us. We do know the following: (1) in our sample of 287 stocks, February cycle stocks form the smallest group (68 stocks compared with 131 for the January cycle and 88 for the March cycle); (2) the order of the average firm size coincides with that of the number of stocks: \$20.7 billion for February cycle stocks, \$26.4 billion for March cycle stocks and \$30.9 billion for January cycle stocks; (3) February cycle options also happen to have the largest proportional spread (0.0948 compared with 0.0899 for the January cycle and 0.0939 for the March cycle); and (4) the average spreads noted in (3) are not statistically different from one another when pair-wise equality tests are performed.

We draw the following conclusions from the cross-section regressions. First, both types of substitution effects are stronger when the options already have a larger proportional bid–ask spread. Second, the moneyness-substitution effect is stronger for firms with a large size, a lower return volatility and a lower *PIN*. In other words, moneyness-substitution effect tends to be stronger for options mostly traded for non-informational reasons. Third, proportional spread and the February cycle dummy are the only two robust explanatory variables for the maturity-substitution effect.

5. Auxiliary tests and robustness checks

In this section, we repeat some of the tests using alternative liquidity measures and perform various robustness checks.

5.1. Miscellaneous robustness checks

To begin, we perform several straightforward robustness checks. Firstly, the 30-day historical return volatility is replaced by the daily average implied volatility or daily return squared in all regressions. These changes have virtually no consequences. In all regressions, different volatility measures yield very similar levels of statistical significance. Secondly, instead of a simple dichotomy of moneyness, for sub-sample tests, we split the sample into four moneyness buckets as [0.90, 0.95), [0.95, 1.00), [1.00, 1.05) and [1.05, 1.10].[9] No discernible differences exist between the two buckets on each side of the moneyness. In addition, all tests are repeated by expanding the moneyness range to [0.8, 1.2] and the results remain unchanged qualitatively. Thirdly, the regressions in Tables 3, 5, 6 and 7 are repeated by deleting zero-volume observations. The adjusted R^2 and *t*-values decrease slightly in certain cases, but the significance remains in all cases. Fourthly, all tests are repeated with various sizes of larger samples. As stated in the data section, we select the top 100 stocks each year according to option volume, yielding 287 distinct stocks.

This data selection procedure is repeated for the top 200, 500 and 1000 stocks. The overall testing power does not improve in a noticeable fashion. The main reason is that, as we expand the sample to include stocks with less liquid options, the benefit of a larger sample in calculating *t*-values is offset by the cost of a bigger standard deviation across coefficient estimates. In other words, as we include more stocks with less liquid options, the first-pass, time-series regressions yield coefficients that are less and less reliable. The good news is, the testing power is more or less the same across different sample sizes. In sum, the finding of substitutions and their impact on spreads are robust to volatility estimations, moneyness grouping and sample selection.

5.2. Stock's bid–ask spread as an additional liquidity determinant

Cho and Engle (1999) argue that the option's bid–ask spread should be positively related to that of its underlying stock due to hedging additives. Clearly, the stock's bid–ask spread should affect the option's. Two questions arise: (1) to what extent does the stock's spread affect the option's? and (2) after controlling for the stock's bid–ask spread, do the option's return volatility and other liquidity determinants still have significant explanatory power? To address these questions, we need the bid and ask quotes for the stocks in our sample. Following Cao and Wei (2010), we retrieve the quotes from TAQ and use the average bid/ask prices in the last 5 min of trading to calculate the end-of-the-day proportional bid–ask spread. We then redo the regressions in Table 3 by incorporating this additional explanatory variable. Table 10 reports the results.

First of all, the underlying stock's proportional bid–ask spread significantly explains the variation in option's spread in all cases, confirming the hedging effect predicted by Cho and Engle (1999). Second, comparing Table 10 with Table 3, we see that the significance of the option return volatility declines slightly but the *t*-values are still very large, implying that the explanatory power of option return volatility remains strong after incorporating the stock's *PBA*. In addition, comparisons of the corresponding *R*-squares reveal that the incremental explanatory power of the stock's *PBA* is negligible in the multivariate regressions. The weak link between the stocks and options *PBAs* is consistent with the finding in Cao and Wei (2010).

5.3. Liquidity determinants – two-stage regressions

As shown in Table 2, the option price is highly correlated with other liquidity determinants, and the option return volatility is modestly correlated with time to maturity and moneyness. For the regressions in Table 3, although we perform diagnostic

[9] This is done for all tables other than Tables 5 and 8.

Table 10
Option liquidity determinants including stock's bid–ask spread. This table reports the results from various versions of regression (3.1) by adding the stock's proportional bid–ask spread PBA_s:

$$PBA = \beta_0 + \beta_1 PBA_s + \beta_2 TTM + \beta_3 M + \beta_4 \sigma_o + \beta_5 \sigma_s + \beta_6 V + \beta_7 P + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon,$$

where all the other variables are defined as in Table 3. The regression is run separately for calls and puts and the coefficients are averaged across the stocks in the sample. For each type of options (calls or puts) we also run separate regressions for in-the-money (ITM) and out-of-the-money (OTM) options. The *t*-statistic is calculated using the regression coefficients. Each entry consists of the average coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is averaged over the time-series regressions. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	Regression	PBA_s	TTM	M	σ_0	σ_s	V	P	Adj R^2 (%)
Panel A: Call options									
All	I	1.525	−0.141	0.497					39.10
		3.229	**−43.036**	**33.971**					
	II	1.635	−0.081		0.001				50.28
		3.195	**−38.914**		**19.405**				
	III	1.942	−0.135			−0.014			26.16
		2.957	**−42.810**			**−7.919**			
	IV	1.663	−0.134				−0.056		26.09
	V	**2.815**	**−41.576**				−0.885		
		0.874	−0.054					−0.017	35.65
		2.987	**−21.758**					**−13.301**	
	VI	2.159	−0.081		0.001	−0.039			50.82
		3.345	**−39.551**		**18.802**	**−14.297**			
	VII	2.140	−0.081		0.001	−0.039	−0.263		50.91
		3.285	**−38.715**		**18.672**	**−14.453**	**−3.689**		
	VIII	1.992	−0.095	0.295	0.001	−0.035	−0.498		54.94
		3.410	**−40.931**	**28.837**	**19.033**	**−15.407**	**−6.170**		
ITM	I	1.371	−0.062	0.304					43.64
		3.885	**−33.489**	**30.095**					
	II	1.002	−0.033		0.003				47.98
		5.218	**−26.412**		**30.949**				
	III	1.453	−0.063			−0.001			38.45
		3.476	**−33.529**			−1.022			
	IV	1.407	−0.063				−0.041		38.36
		3.361	**−32.543**				−0.933		
	V	0.860	−0.024					−0.008	44.42
		4.324	**−14.058**					**−11.834**	
	VI	1.445	−0.027		0.003	−0.046			50.10
		4.659	**−26.811**		**27.965**	**−18.770**			
	VII	1.445	−0.028		0.003	−0.046	−0.500		50.32
		4.523	**−26.942**		**28.154**	**−18.949**	**−6.430**		
	VIII	1.413	−0.031	0.074	0.003	−0.042	−0.505		50.65
		4.468	**−25.719**	**10.667**	**26.658**	**−18.549**	**−6.648**		
OTM	I	1.523	−0.234	0.680					39.83
		2.181	**−42.721**	**34.323**					
	II	1.559	−0.140		0.001				53.23
		2.661	**−40.952**		**19.888**				
	III	1.686	−0.227			−0.032			35.69
		4.202	**−43.962**			**−11.363**			
	IV	1.233	−0.231				−0.905		35.67
		2.525	**−44.346**				**−8.739**		
	V	1.295	−0.117					−0.023	39.17
		2.637	**−24.698**					**−11.906**	
	VI	2.335	−0.139		0.001	−0.062			53.97
		3.511	**−41.849**		**19.391**	**−16.492**			
	VII	2.280	−0.143		0.001	−0.060	−0.756		54.15
		3.420	**−42.432**		**18.698**	**−16.575**	**−7.524**		
	VIII	2.302	−0.152	0.310	0.001	−0.059	−0.708		55.05
		3.175	**−41.168**	**30.901**	**19.041**	**−16.533**	**−7.212**		
Panel B: Put options									
All	I	1.285	−0.155	−0.572					42.40
		3.361	**−46.346**	**−37.867**					
	II	1.540	−0.093		0.001				51.90
		3.276	**−43.843**		**18.840**				
	III	1.434	−0.149			−0.021			28.16
		4.533	**−48.067**			**−10.647**			
	IV	1.213	−0.148				−0.021		28.04
		3.992	**−46.950**				−0.360		
	V	0.500	−0.066					−0.021	38.69
		2.764	**−30.721**					**−14.719**	
	VI	2.077	−0.093		0.001	−0.049			52.56
		3.409	**−44.699**		**18.194**	**−16.376**			
	VII	2.063	−0.093		0.001	−0.049	−0.270		52.62
		3.388	**−43.532**		**18.190**	**−16.386**	**−3.677**		
	VIII	1.928	−0.108	−0.338	0.001	−0.043	−0.537		57.00
		3.357	**−42.803**	**−33.661**	**18.178**	**−16.317**	**−6.198**		

Table 10 (*continued*)

	Regression	PBA_s	TTM	M	σ_0	σ_s	V	P	Adj R^2 (%)
ITM	I	1.259	−0.067	−0.360					46.47
		2.885	**−33.164**	**−22.410**					
	II	1.047	−0.044		0.003				50.63
		4.639	**−26.735**		**31.772**				
	III	1.781	−0.069			−0.005			40.60
		2.959	**−33.082**			**−3.257**			
	IV	1.599	−0.068				0.133		40.47
		2.809	**−32.343**				**3.961**		
	V	0.822	−0.029					−0.010	47.08
		3.541	**−15.307**					**−10.229**	
	VI	1.599	−0.037		0.003	−0.061			53.45
		4.897	**−29.197**		**26.028**	**−17.437**			
	VII	1.586	−0.037		0.003	−0.061	−0.398		53.54
		4.953	**−29.392**		**25.714**	**−17.393**	**−7.188**		
	VIII	1.393	−0.039	−0.050	0.003	−0.057	−0.401		53.86
		4.340	**−27.103**	**−3.140**	**26.070**	**−21.858**	**−7.291**		
OTM	I	1.404	−0.255	−0.777					43.45
		3.044	**−44.037**	**−37.634**					
	II	1.737	−0.158		0.001				55.41
		3.123	**−42.390**		**19.523**				
	III	1.808	−0.247			−0.039			39.03
		3.341	**−45.284**			**−13.531**			
	IV	1.351	−0.250				−1.143		38.89
		3.317	**−45.496**				**−7.664**		
	V	1.198	−0.121					−0.032	43.41
		2.469	**−26.037**					**−13.349**	
	VI	2.567	−0.156		0.001	−0.074			56.32
		3.116	**−43.601**		**18.319**	**−16.994**			
	VII	2.556	−0.159		0.001	−0.073	−0.951		56.43
		3.087	**−43.823**		**18.354**	**−16.990**	**−7.057**		
	VIII	2.515	−0.168	−0.344	0.001	−0.071	−0.889		57.37
		3.090	**−42.034**	**−29.126**	**17.902**	**−16.948**	**−6.628**		

Table 11
Option liquidity determinants – two-stage regressions. This table reports the results for the two-pass regressions in (5.1) and (5.2):

$$\sigma_0 = \beta_0 + \beta_1 TTM + \beta_2 M + \beta_3 \sigma_s + \beta_4 V + \beta_5 P + \sum_{i=1}^{11} \gamma_i YearDummy_i + \varepsilon,$$

$$PBA = \gamma_0 + \gamma_1 PBA_s + \gamma_2 \varepsilon + \eta,$$

where in the first-pass, for each stock, option return volatility (σ_o) is regressed on time to maturity in years (*TTM*), moneyness (*M*) defined as exercise price over stock price, stock return volatility (σ_s) which is the annualized return standard deviation over the past 30 days, option volume (*V*) which is the number of contracts divided by 100,000, the option price (*P*), and the year dummy (*YearDummy*) for 1996–2006; in the second pass, the option's proportional bid–ask spread (*PBA*) is regressed on the stock's proportional bid–ask spread (PBA_s) and the residual from the first-pass. We run two versions of the second pass: with or without the stock's proportional bid–ask spread (PBA_s). The regressions are run separately for calls and puts and the coefficients are averaged across the stocks in the sample. For each type of options (calls or puts) we also run separate regressions for in-the-money (ITM) and out-of-the-money (OTM) options. The *t*-statistic is calculated using the regression coefficients. Each entry consists of the average coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is averaged over the time-series regressions. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

	First-pass regression						Second-pass regressions				
							Without stock PBA		With stock PBA		
	TTM	M	σ_s	V	P	Adj R^2 (%)	ε	Adj R^2 (%)	PBA_s	ε	Adj R^2 (%)
Panel A: Call options											
All	−62.613	291.953	28.399	45.382	−2.228	20.30	0.0007	15.26	3.4260	0.0007	17.52
	−25.057	**24.202**	**20.184**	1.360	**−11.480**		**19.213**		**4.681**	**19.220**	
ITM	−7.687	60.004	12.643	84.351	−0.535	61.32	0.0029	5.62	3.1204	0.0030	10.37
	−32.946	**90.568**	**57.484**	**13.906**	**−15.516**		**25.918**		**5.271**	**25.837**	
OTM	−90.487	536.663	49.439	−17.750	−10.039	24.68	0.0006	14.00	3.2087	0.0006	16.18
	−23.344	**23.284**	**16.646**	−0.394	**−22.759**		**19.070**		**3.533**	**19.336**	
Panel B: Put options											
All	−76.209	−380.790	35.932	264.892	−3.013	22.04	0.0006	14.07	3.6054	0.0006	16.34
	−25.653	**−28.529**	**23.267**	1.869	**−12.618**		**18.515**		**5.657**	**18.432**	
ITM	−6.549	−82.514	16.314	87.450	−0.650	65.60	0.0029	5.97	3.7354	0.0030	11.03
	−29.825	**−90.771**	**65.671**	**8.791**	**−15.833**		**20.718**		**3.772**	**20.560**	
OTM	−95.721	−670.473	66.347	251.211	−16.547	26.50	0.0005	12.48	4.0305	0.0006	14.76
	−24.915	**−28.073**	**20.130**	1.408	**−25.089**		**18.842**		**6.273**	**18.808**	

Table 12
Option liquidity determinants – alternative liquidity measures. This table reports the results from various modified versions of regression (3.1) by replacing option's proportional bid–ask spread (*PBA*) with two illiquidity measures:

$$ILLIQ = \beta_0 + \beta_1\sigma_o + \beta_2\sigma_s + \beta_3 V + \beta_4 P + \sum_{i=1}^{11}\gamma_i YearDummy_i + \varepsilon,$$

where all variables are defined as in Table 3 and *ILLIQ* is an illiquidity measure (Amihud, 2002). We use two versions of *ILLIQ* as in Cao and Wei (2010): the absolute *ILLIQ* (*AILLIQ*) and the percentage *ILLIQ* (*PILLIQ*). See Section 5.3 for calculations details. Unlike in Table 3, the observations here are in daily frequency. We therefore also run the regressions for the option's *PBA* as a benchmark case. The regression is run separately for calls and puts and the coefficients are averaged across 260 stocks in the sample. When the independent variable is *AILLIQ* or *PILLIQ*, the coefficients for σ_o, σ_s and *P* are multiplied by 10,000 for ease of presentation. The *t*-statistic is calculated using the regression coefficients. Each entry consists of the average coefficient (the upper number) and its *t*-value (the lower number). The adjusted R^2 is averaged over the 260 time-series regressions. The *t*-values in bold type are significant at the 5% level or higher for two-tail tests.

Regression	Dependent variable: *PBA*					Dependent variable: *AILLIQ*					Dependent variable: *PILLIQ*				
	σ_o	σ_s	*V*	*P*	Adj R^2 (%)	σ_o	σ_s	*V*	*P*	Adj R^2 (%)	σ_o	σ_s	*V*	*P*	Adj R^2 (%)
Panel A: Call options															
I	0.001				41.80	0.082				9.45	0.172				8.41
	16.440					**5.799**					**6.973**				
II		−0.016			20.01		−3.372			8.98		−6.586			6.80
		−7.389					**−3.243**					**−5.806**			
III			−1.167		19.94			−0.127		9.92			−0.123		7.65
			−7.027					**−7.031**					**−7.826**		
IV				−0.037	38.35				−2.523	9.36				−5.056	7.87
				−15.101					**−7.949**					**−7.464**	
V	0.001	−0.064			43.86	0.092	−7.210			9.73	0.190	−14.106			8.99
	15.445	**−16.914**				**5.837**	**−5.616**				**7.048**	**−8.297**			
VI	0.001	−0.062	−0.943		44.06	0.092	−5.808	−0.123		10.80	0.191	−12.707	−0.116		10.02
	15.533	**−16.901**	**−7.389**			**5.833**	**−4.788**	**−6.948**			**7.073**	**−7.780**	**−7.772**		
VII	0.001	−0.045		−0.024	50.10	0.077	−6.240		−1.160	10.04	0.160	−12.507		−2.182	9.43
	18.090	**−19.865**		**−12.641**		**4.317**	**−5.132**		**−4.633**		**6.299**	**−8.004**		**−8.168**	
VIII	0.001	−0.043	−0.877	−0.024	50.26	0.076	−4.883	−0.121	−1.179	11.06	0.159	−11.095	−0.115	−2.237	10.44
	18.402	**−19.506**	**−7.688**	**−12.647**		**4.303**	**−4.230**	**−7.178**	**−4.764**		**6.352**	**−7.393**	**−7.938**	**−8.033**	
Panel B: Put options															
I	0.001				43.30	0.139				9.97	0.261				8.88
	14.591					**7.885**					**7.517**				
II		−0.018			21.16		−4.649			9.19		−6.650			6.81
		−2.252					−1.789					−1.142			
III			−1.603		20.76			−0.359		9.92			−0.366		7.39
			−4.830					**−3.875**					**−3.866**		
IV				−0.044	41.01				−5.377	9.69				−10.422	8.19
				−16.383					**−8.687**					**−8.725**	
V	0.001	−0.071			45.33	0.159	−13.102			10.30	0.291	−23.435			9.42
	14.103	**−14.511**				**6.373**	**−6.872**				**7.293**	**−7.398**			
VI	0.001	−0.070	−0.811		45.44	0.158	−11.329	−0.288		11.21	0.290	−21.614	−0.279		10.22
	14.073	**−14.537**	**−5.138**			**6.435**	**−6.377**	**−6.395**			**7.323**	**−7.404**	**−6.625**		
VII	0.001	−0.046		−0.028	52.05	0.138	−11.588		−2.145	10.57	0.247	−20.673		−4.086	9.89
	13.944	**−15.353**		**−15.499**		**5.678**	**−6.713**		**−5.314**		**6.469**	**−7.337**		**−6.365**	
VIII	0.001	−0.045	−0.786	−0.028	52.15	0.136	−10.024	−0.286	−2.248	11.45	0.244	−18.973	−0.280	−4.251	10.69
	13.935	**−15.586**	**−6.726**	**−15.433**		**5.680**	**−6.068**	**−6.496**	**−5.312**		**6.474**	**−7.690**	**−6.541**	**−6.369**	

analyses by running various simple regressions and exclude the price variable in the multivariate regressions, the correlation between the option return volatility and the option's time to maturity and moneyness may still render the regression misspecified. Since regression (3.1) drives the main findings of this paper, it is imperative that we ensure the robustness of our results. To this end, we run the following two-pass regressions[10]:

$$\sigma_0 = \beta_0 + \beta_1 TTM + \beta_2 M + \beta_3\sigma_s + \beta_4 V + \beta_5 P + \sum_{i=1}^{11}\gamma_i YearDummy_i + \varepsilon, \quad (5.1)$$

$$PBA = \gamma_0 + \gamma_1 PBA_s + \gamma_2\varepsilon + \eta, \quad (5.2)$$

where all variables are denned as in (3.1) and PBA_s is the stock's proportional bid–ask spread. The basic idea is as follows: we first remove, via regression (5.1), the components in the option return volatility relating to all other liquidity determinants such as stock return volatility; we then determine, via regression (5.2), if the residual component in the option return volatility can still explain the variation in the option's bid–ask spread. Here we run two versions of the second-pass regression: with or without the stock's proportional bid–ask spread. Table 11 reports the results. Several observations are in order.

To begin, the results in the first-pass regressions are as expected: the option return volatility is higher when (1) time to maturity is shorter, (2) the option is out-of-the-money, (3) the stock volatility is higher, and (4) the option price is lower. A higher option return volatility also tends to be associated with a higher volume, albeit this relationship is significant only for in-the-money options.

The positive and significant coefficient for the stock *PBA* in the second-pass regression is consistent with the findings in Table 10.

More important, the residual option return volatility still has a significant explanatory power for the option's *PBA* whether we control for the stock's *PBA* or not. This is especially true for out-of-the-money options. Our results strongly validate the findings in Table 3 and confirm the proper specification of (3.1).

As an alternative verification, we have also performed the following two-pass regression: we first run (3.1) by omitting the

[10] We thank an anonymous referee for suggesting this approach.

option return volatility on the right-hand-side, and then regress the residual from the first-pass regression on the option return volatility. Here, we attempt to establish the incremental explanatory power of option return volatility while avoiding the potential multicollinearity problem. It turns out that the option return volatility can explain about 20% of the variations in the option *PBA* residuals. Details are available up request.

5.4. Liquidity determinants – an alternative liquidity measure

As argued earlier, our findings should be judged on contributing to the overall understanding of option liquidity. So far we have used the option's bid–ask spread as a measure of option liquidity as many other authors (e.g., George and Longstaff, 1993) have done. In this sub-section, we examine if the main finding from Table 3 (i.e., option return volatility can significantly explain option liquidity) can be generalized to other liquidity measures. To this end, we again take some guidance from Cao and Wei (2010). Besides the option *PBA*, they consider four additional liquidity measures: (1) option volume in contracts, (2) option volume in dollars, (3) absolute price impact in option prices (Amihud, 2002), and (4) percentage price impact in option prices. Since we have already used trading volume as an explanatory variable in various regressions, the only alternative measure we can consider is the price impact variable. Following Cao and Wei (2010), for each day, we calculate the absolute and percentage price impacts (*AILLIQ* and *PILLIQ*) as the volume weighted price changes with delta-adjustments. To ensure that the two liquidity measures can be estimated with reasonable accuracy, we follow Cao and Wei (2010) and screen out stocks that have fewer than 500 option observations within a calendar year. This additional screening reduces our sample from 287 stocks to 260 stocks.

The aggregation from panel data (across time to maturity and moneyness) to daily frequency has two implications: (1) the time to maturity (*TTM*) and moneyness (*M*) variables, no matter how aggregated, are no longer meaningful in the time-series regression (3.1),[11] and (2) the panel data of option *PBA*, option return volatility, and option price must also be aggregated somehow. To be consistent with the calculation of *AILLIQ* and *PILLIQ*, we also use option volume as the weight when calculating the average within each day. Finally, the daily volume is simply the average volume across contracts.

Since we now have a reduced sample – 260 stocks versus the previous 287 – and the data frequency is daily, we first run a base-case regression of (3.1) by removing *TTM* and *M*. We then repeat the base-case regression by replacing the option *PBA* with *AILLIQ* or *PILLIQ*. Table 12 contains the results.

To begin, the base-case regression results confirm the general findings in Table 3. The *R*-squares are slightly lower since we have dropped two explanatory variables, *TTM* and *M*. When we replace the option *PBA* with *AILLIQ* or *PILLIQ*, the explanatory power of all variables remains, although the *R*-squares are much lower. Moreover, the comparative advantage of option return volatility is no longer obvious: the R^2 of the univariate regression using option return volatility as the explanatory variable is only marginally higher than those of the other explanatory variables. This however should not be taken as evidence against the potent power of option return volatility in explaining option's liquidity. The price impact measure is a very noisy measure of liquidity chiefly due to the fact that not all options are traded each day. Please note that, in order to calculate *AILLIQ* or *PILLIQ*, we need prices of *the same option* in two consecutive days. This condition is not always met, especially for less liquid options.

[11] In a time-series setting, the average time to maturity will simply be a recording of elapsing time and the average moneyness is just a quantity close to unity.

Table 12 does indicate that the general conclusions on liquidity behavior derived from option *PBA* seem to be valid for alternative liquidity metrics such as the price impact measure.

6. Summary and conclusion

The literature on option liquidity is extremely thin. Only a handful of papers exist on the subject. It is fair to say that the quest for understanding option market liquidity has just begun. This paper attempts to make a contribution by offering insights into some specific aspects of equity option liquidity. We ask and attempt to answer two simple questions: (1) what are the determinants of option liquidity? (2) over and above the identified liquidity determinants, how do trading activities affect option liquidity?

Using the proportional bid–ask spread (*PBA*) as a measure of option liquidity, we examine the explanatory power of several liquidity determinants: time to maturity, moneyness, stock return volatility, option return volatility, option trading volume and option price. These determinants collectively explain more than half of the time-series variations in *PBA*. Contrary to common beliefs, stock return volatility and option volume have the least explanatory power compared with other determinants. By far the most powerful explanatory variable is the option return volatility. This variable alone explains over 45% of the variations in the option's *PBA*. Insofar as the option return volatility measures the overall riskiness of the option and thus directly affects the inventory risk faced by market makers, its impressive explanatory power should not be too surprising. However, to our knowledge, the role of option return volatility in determining option liquidity has not been acknowledged in the literature. The only two known studies that examine this determinant are Neal (1987) and George and Longstaff (1993). Somewhat unfortunately, both studies establish an inconsequential role of option return volatility chiefly due to the fact that they use the dollar spread to measure liquidity.

Assigning maturity months according to expiration cycles means that stocks do not always have options available for all consecutive months. The current rule dictates that all option maturities follow a quarterly cycle and that there must be options available for the month following the current month. As a result, there are always options available for the first 2 months at any given point in time; but the next month with options trading does not have to be the third month. We show that the *PBA* of the first 2 months' options is negatively affected by how far the third expiry month is, after controlling for the liquidity determinants. Simply put, when options of the third expiry month are too far away, the previous demand for mediate-term options (with maturities of 90 days and longer) is shifted to short-term options, causing the latter's volume to go up and spread to go down. We term this phenomenon the maturity-substitution effect.

We also uncover a moneyness-substitution effect: a higher stock return volatility induces a demand shift from in-the-money options to out-of-the-money options. This shift in trading volume leads to a relatively narrower spread for out-of-the-money options compared with in-the-money options.

Finally, we investigate how the extent of substitution effect is related to such firm characteristics as size, share turnover, stock volume relative to option volume, option's average *PBA*, stock return volatility and *PIN*. The only attribute that consistently exhibits cross-sectional explanatory power is *PBA*: the higher the *PBA* the stronger both substitution effects.

All told, our study leads to the following three main conclusions. (1) Option liquidity is related to time to maturity, moneyness, trading volume, option price, stock return volatility and option return volatility. Option return volatility is by far the most important determinant of option liquidity, indicating the importance of inventory risk in determining options bid–ask spreads. (2) Trading activities around expiration cycles affect the liquidity of short-term options after controlling for the above listed determinants. (3) The stock return volatility affects the relative liquidity of in-the-money and out-of-the-money options.

Acknowledgements

This research was carried out while Jinguo Zheng was a visiting Ph.D. student at the Joseph L. Rotman School of Management, University of Toronto. We thank Siu Kai Choy for compiling part of the data. We also thank an anonymous referee and the Editor, Ike Mathur for useful comments and suggestions. Jason Wei gratefully acknowledges the financial support from the Social Sciences and Humanities Research Council of Canada.

References

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5, 31–56.
Benston, C., Hagerman, R., 1974. Determinants of bid–ask spreads in the over-the-counter market. Journal of Financial Economics 1, 353–364.
Branch, B., Freed, W., 1977. Bid-asked spreads on Amex and the Big Board. Journal of Finance 32, 159–163.
Brenner, M., Eldor, R., Hauser, S., 2001. The price of options illiquidity. Journal of Finance 56, 789–805.
Cao, M., Wei, J., 2010. Option market liquidity: Commonality and other characteristics. Journal of Financial Markets 13, 20–48.
Chan, J., Hong, D., Subrahmanyam, M., 2008. A tale of two prices: Liquidity and asset prices in multiple markets. Journal of Banking and Finance 32, 947–960.
Cho, Y., Engle, R., 1999. Modeling the impacts of market activity on bid–ask spreads in the option market. NBER Working Paper.
Chong, B., Ding, D., Tan, K., 2003. Maturity effect on bid–ask spreads of OTC currency options. Review of Quantitative Finance and Accounting 21, 5–15.
Deuskar, P., Gupta, A., Subrahmanyam, M., 2009. Liquidity effect in OTC options markets: Premium or discount? Working Paper, New York University.
French, K., Roll, R., 1986. Stock return variances: The arrival of information and the reaction of traders. Journal of Financial Economics 17, 5–26.
Frino, A., Gerace, D., Lepone, A., 2008. Liquidity in auction and specialist market structure: Evidence from the Italian bourse. Journal of Banking and Finance 32, 2581–2588.
George, T., Longstaff, F., 1993. Bid–ask spreads and trading activity in the S&P 100 index options market. Journal of Financial and Quantitative Analysis 28, 381–397.
Glosten, L., 1987. Components of the bid–ask spread and the statistical properties of transaction prices. Journal of Finance 42, 1293–1307.
Glosten, L., Milgrom, P., 1985. Bid, ask, and transaction prices in a special market with heterogeneously informed traders. Journal of Financial Economics 14, 71–100.
Glosten, L., Harris, L., 1988. Estimating the components of the bid–ask spread. Journal of Financial Economics 21, 123–142.
Ho, T., Stoll, H., 1981. Optimal deal pricing under transactions and return uncertainty. Journal of Financial Economics 9, 47–73.
Kalodera, I., Schlag, C., 2004. An empirical analysis of the relation between option market liquidity and stock market activity. Working Paper, University of Frankfurt.
Mantecon, T., Poon, P., 2009. An analysis of the liquidity benefits provided by secondary markets. Journal of Banking and Finance 33, 335–346.
Mayhew, S., Sarin, A., Shastri, K., 1999. What drives option liquidity? Working Paper, Purdue University.
Mclnish, T., Wood, R., 1992. An analysis on intraday patterns in bid/ask spreads for NYSE stocks. Journal of Finance 57, 753–764.
Neal, R., 1987. Potential competition and actual competition in equity options. Journal of Finance 42, 511–531.
Rakowski, D., Beardsley, X., 2008. Decomposing liquidity along the limit order book. Journal of Banking and Finance 32, 1687–1698.
Roll, R., 1984. A simple implicit measure of the effective bid–ask spread in an efficient market. Journal of Finance 39, 1127–1139.
Stoll, H., 1978a. The supply of dealer services in securities markets. Journal of Finance 33, 1133–1151.
Stoll, H., 1978b. The pricing of security dealer services: An empirical study of NASDAQ stocks. Journal of Finance 33, 1153–1172.
Tang, D., Yan, H., 2008. Liquidity and credit default swap spreads. Working Paper, The University of Hong Kong.
Tinic, S., West, R., 1972. Competition and the pricing of dealer services in the over-the-counter stock market. Journal of Financial and Quantitative Analysis 8, 1707–1728.
Vijh, A., 1990. Liquidity of the CBOE equity options. Journal of Finance 55, 1157–1179.
Wang, F., 2010. Informed arbitrage with speculative noise trading. Journal of Banking and Finance 34, 304–313.