Failure Is an Option: Impediments to Short Selling and Options Prices

Richard B. Evans
Darden School at the University of Virginia

Christopher C. Geczy and David K. Musto
The Wharton School at the University of Pennsylvania

Adam V. Reed
Kenan-Flagler Business School at the University of North Carolina

Regulations allow market makers to short sell without borrowing stock, and the transactions of a major options market maker show that in most hard-to-borrow situations, it chooses not to borrow and instead fails to deliver stock to its buyers. A part of the value of failing passes through to options prices: when failing is cheaper than borrowing, the relation between borrowing costs and options prices is significantly weaker. The remaining value is profit to the market maker, and its ability to profit despite competition between market makers appears to result from the cost advantage of larger market makers. (*JEL* G12, G13, G29)

In the United States, the market for short exposure clears differently from the market for long exposure. This difference has recently attracted considerable interest from both the SEC and the market participants who frequently short sell or whose stock is sold short.[1] The interest is both in the economics of clearing and in the pricing of the affected assets, which could be high or inefficient. Our goal is to establish the role of an unfamiliar yet important clearing tactic: failing to deliver.

Short sales are usually accomplished through equity loans. The short seller borrows shares from an equity lender, which he delivers to the buyer. This debt of shares to the lender gives him short exposure going forward. But there is another way to create the same exposure: by failing to deliver the shares. If the

---

We gratefully acknowledge important input from two anonymous referees, Michael Brandt, Greg Brown, Jennifer Conrad, George Constantinides, Patrick Dennis, Darrell Duffie, Bin Gao, Eitan Goldman, Rick Green, Jonathan Karpoff, Richard Rendleman, Erik Stafford, and seminar participants at Notre Dame, USC, UT, Wharton, and the 2002 Western Finance Association Meetings. We thank Wes Gray and Robert Turnquest for excellent research assistance. Send correspondence to Adam Reed, Campus Box 3490, McColl Building, Chapel Hill, NC 27514; telephone: (919) 962-9785. E-mail: adam_reed@unc.edu.

[1] In July of 2004 the SEC passed regulation SHO, which limits the ability of certain market participants to sell stock short without borrowing to cover their position. The discussion period for regulation SHO attracted considerable attention from the business press.

doi:10.1093/rfs/hhm083 Advance Access publication January 5, 2008

short seller delivers nothing to the buyer, thereby incurring a debt of shares to the buyer, this also gives him short exposure going forward. This alternative moves the risk that the short seller does not repay his debt from the equity lender to the buyer, but just as equity lenders have a mechanism for ensuring performance—i.e., collateral—so do the buyers. The clearing corporation intermediating the trade takes margin and marks it to market, thereby defending buyers against their sellers' nonperformance. If equity loans are expensive, unavailable, or unreliable, as research shows they can be (e.g., D'Avolio, 2002; Geczy, Musto, and Reed, 2002; Jones and Lamont, 2002; Lamont, 2004), then this alternative appears desirable, to short sellers if not to buyers. But, considering the market rules that bind short sales to equity loans, how is it feasible?

The answer, we show, lies in the special access to delivery failures that options market makers enjoy. Traders are generally obliged to locate shares to borrow before shorting, but those engaged in bona fide hedging of market making activity are exempt from this requirement. So, unlike traders in general, a market maker can short sell without having located shares to borrow. If he does not locate shares to borrow, then he fails to deliver, someone on the other side fails to receive, and therefore retains the purchase price, and the clearing corporation starts taking margin. While it lasts, this arrangement is effectively an equity loan from the buyer to the seller at a zero rebate. But whether it lasts depends on the reaction of the trader to failing. If a buyer does not get his shares, then he can demand them, in which case a short seller who failed is bought in: he must go buy the shares and hand them over. If that short seller wants to maintain his short exposure, he must short again, so this demand increases his shorting cost by this roundtrip transactions cost. Thus, the cost of failing to deliver is the cost of a zero-rebate equity loan plus the expected incidence of buy-in costs. If this incidence is low enough, then failing is a valuable alternative to borrowing some difficult-to-borrow stocks. We show that the alternative to fail is valuable and important to the pricing and trading of options.

First, we show that shorting costs move options out of parity. That is, synthetic shorts constructed from options trade below spot market prices when shorting is costly—i.e., when interest rebates on equity loans are low—and this disparity grows as the rebate falls. However, this growth slows when the rebate falls below zero, consistent with options market makers choosing failure over negative rebates, and sharing some of the savings. Furthermore, the short interest of a major options market maker grows as a fraction of marketwide short interest as rebates fall, consistent with the market maker having and sharing an advantage in getting short exposure to hard-to-borrow stocks.

We can see the advantage directly in the market maker's shorting experience. Half the time the market maker shorts a hard-to-borrow stock, it fails to deliver at least some of the shares. And it never accepts a negative rebate, always choosing to fail instead. This advantage could in principle be offset by frequent buy-ins, but we find a very low frequency of buy-ins, executed with small price concessions.

How much of this advantage does the market maker share? Estimating the market maker's trading profits and net rebate reductions and buy-ins, we calculate a significant average profit. This profit seems at odds with the competitiveness of options markets, but we show that it corresponds to the way the clearing corporation handles buy-ins. The highest volume options market makers, such as our data supplier, likely benefit from the clearing corporation's practice of assigning buy-ins to the oldest fails. That is, when a number of short sellers' brokers are failing on the same stock and a buyer's broker demands shares, the clearing corporation passes this demand to the broker whose failure started first. This favors the few highest-volume traders because, since their portfolios turn over so much, their fails are rarely the oldest. Thus, we hypothesize that options market competition tends to oligopoly as stocks grow hard to short.

We test our hypothesis, in a number of ways, that market maker competition weakens as specialness grows. First, we test whether the market maker we observe is responsible for an increasing share of short exposure as specialness increases. Second, we test whether a measure of competition, the option bid–ask spread, changes when stocks go on special. Third, we test whether a low turnover increases the likelihood of buy-ins, and we examine the incidence of buy-ins for our data provider. Fourth, we test whether an increase in market maker competition squeezes the profitability of put-call parity arbitrage. The results of these tests support our hypothesis by showing that at least some of the profits result from limits to competition.

The rest of this paper is organized as follows. Section 1 reviews the literature, Section 2 describes the database, Section 3 presents the results, and Section 4 concludes. Appendix A provides background information and relevant details regarding short selling and delivery, and Appendix B describes our methodology for calculating the term structure of interest rates.

1. Review of Literature

This paper is not the first to document that shorting frictions are associated with breakdowns of put-call parity. Lamont and Thaler (2003) find that impediments to short selling prevent traders from exploiting seemingly profitable arbitrage strategies resulting from the misalignment of stock prices in equity carve-outs. Similarly, Ofek, Richardson, and Whitelaw (2004) measure the relationship between increased borrowing costs and put-call disparity and find that cumulative abnormal returns for arbitrage strategies involving put-call disparity exceed 65%. But, as in Jarrow and O'Hara (1989), market imperfections prevent most arbitrageurs from turning the misalignment into a profit. The put-call parity trades studied here can be performed only by market participants who can always borrow stock or short sell without borrowing stock. In other words, rebate rates are valid only if stocks are found and borrowed. Our study has the unique advantage of a coherent approach that combines actual borrowing costs and feasibility for one market participant: a large options market maker.

Furthermore, this paper is not the first to discuss settlement fails; there is a strand of literature that studies settlement and settlement failures in markets other than U.S. equities. In the context of monetary policy, Johnson (1998) finds that technological improvements in the banking settlement system have affected monetary policy. In the context of foreign exchange, Kahn and Roberds (2001) show that settlement through a private intermediary bank can mitigate some of the unique risks associated with foreign exchange settlement. Fleming and Garbade (2002) find that treasury market settlement fails jumped following the September 11 attacks as a result of the destruction of communication facilities.[2]

In this paper, we identify the possibility of profiting from the misalignment, due to short-sale costs, of stock and options markets for market participants who have the option to fail to deliver shares, and we show how limited access to this option is a barrier to entry that prevents competition from realigning market prices. This work relates primarily to three topics in the finance literature: equity lending, the relation of observed prices to Black-Scholes, and deviations from put-call parity. We briefly review each.

1.1 The equity lending market

A number of recent papers have examined variation in the cost of borrowing stock in the equity lending market. Reed (2007) uses one year's daily equity-loan data to measure the reduction in informational efficiency resulting from short-sale costs. Geczy, Musto, and Reed (2002) measure the impact of equity-loan prices on a variety of trading strategies involving short selling. The paper finds that prices in the equity lending market do not preclude short sellers from getting negative exposure to effects on average, but in the case of stock-specific merger arbitrage trades, short-selling impediments reduce profits substantially. Christoffersen, Geczy, Musto, and Reed (2005, 2007) use the same database to study stock loans that are not necessarily related to short selling. The first paper finds an increase in both quantity and price of loans on dividend record dates when the transfer of legal ownership leads to tax benefits, and the second paper finds an increase in quantity of loans on voting record dates when important proposals are under consideration. Using another database of rebate rates, Ofek and Richardson (2003) demonstrate that short selling was generally more difficult for Internet stocks in early 2000, and D'Avolio (2002) uses 18 months' daily data to relate specialness to a variety of stock-specific characteristics. Jones and Lamont (2002) study borrowing around the crash of 1929; the paper finds that hard-to-borrow stocks had low future returns. Finally, Duffie, Gârleanu, and Pedersen (2002) formulate a search model of the equity lending market.

2 More recently, Boni (2005) explored market-wide failing data from the clearing corporation.

1.2 Predicted and observed options prices

By relating short selling to options prices, this paper also contributes to the large literature on the difference between Black-Scholes (1973) options prices and observed options prices. MacBeth and Merville (1979) and Rubinstein (1985) show that, empirically, implied volatilities are not equal across option classes and that deviations are systematic. As in Derman and Kani (1994), these systematic deviations are commonly referred to as the volatility smile. Longstaff (1995) shows that the difference between Black-Scholes and actual options prices increases with option bid-ask spreads and decreases with market liquidity. Although Longstaff's results are contested in a later work (Strong and Xu, 1999), he provides a novel approach to testing the impact of market frictions on options prices. Dumas, Fleming, and Whaley (1998) test a range of time- and state-dependent models of volatility meant to account for observed deviations from Black-Scholes prices. The paper concludes that these models still leave a large mean-square error when explaining market prices. Using Spanish index options, Peña, Rubio, and Serna (1999) find evidence consistent with U.S. markets; they find a positive and significant contribution of the bid-ask spread to the slope of the volatility smile. Dennis and Mayhew (2002) examine the contribution of various measures of market risk and sentiment on individual index options and find that both are correlated with the smile.

1.3 Tests of put-call parity

Some of the evidence on the impact of short-sale impediments on options prices is presented here in terms of put-call parity. Tests of put-call parity date back to Klemkosky and Resnick (1979), who find options market prices to be largely consistent with put-call parity. In a related paper that focuses on the speed of adjustment of option and stock markets, Manaster and Rendleman (1982) conclude that closing options prices contain information about equilibrium stock prices that is not contained in closing stock prices. Although the implied-stock-price measure employed in our work differs substantially from that of Manaster and Rendleman (1982), the approach of comparing actual and implied stock prices is similar.

2. Data

We combine several databases in this study. A prominent options market maker provided a database of rebate rates, failing positions, and buy-ins. For equity options prices, implied volatilities, options volume, and open interest, we use the OptionMetrics database. Finally, the interest rate term structure is estimated using commercial paper rates from the Federal Reserve (see Appendix B for details).

2.1 Rebate rates, fails, and buy-ins

A large options market making firm has generously provided a database of their rebate rates, fails, and buy-ins for 1998 and 1999. The rebate rates are the interest rates on cash collateral for stock loans quoted by the market maker's prime broker. Different borrowers are likely to face different rates; this database is a description of one large market maker's experience. As discussed in Geczy, Musto, and Reed (2002), rebate rates allow us to measure the difficulty of borrowing shares, or specialness.

We construct a measure of specialness for each stock on each date. Specifically, specialness on any stock is the difference between the general collateral rate and the rebate rate on that stock. Following Geczy, Musto, and Reed (2002), we estimate the general collateral rate as the Federal Funds Rate minus the equity lender's fixed commission. Specialness is zero for most stocks, and it is positive for specials, or hard-to-borrow stocks. We augment our market maker's specialness database with specialness from a large custodian lender as described in Geczy, Musto, and Reed (2002).

The rebate rates cover all stocks in the Russell 3000 index, and we have limited our other databases to the Russell 3000 using constitution lists from the Frank Russell Company. The Russell 3000 includes the 3000 largest stocks in the United States based on May 31 market capitalization. In 1997, stocks larger than \$171.7 were included. The cutoff was \$221.9 in 1998 and \$171.2 in 1999.

The database also indicates when this market maker is failing to deliver shares on any of its short positions. Even though we do not have data on any of this market maker's specific trades, we do have information about this market maker's buy-ins. The buy-in database has purchase dates, settlement dates, and execution prices for every buy-in occurring in 1998 and 1999.

2.2 Options data

We use the Ivy DB OptionMetrics database for U.S. equity options prices, spreads, and volume. We use the average of the lowest closing ask and the highest closing bid, or the midquote, as the options price. We apply three filters that are common elsewhere in the options literature (e.g., Dumas, Fleming, and Whaley, 1998; Bakshi, Cao, and Chen, 1997). First, we remove options with fewer than six calendar days to maturity to mitigate liquidity bias. Second, we remove options with prices less than \$0.375 to minimize price discreteness. Third, as presented in Table 1, no arbitrage restrictions are applied to the option quotes.

Table 1 indicates that the intersection of the rebate and options databases contains 19,723,466 observations. After filtering, the database contains 11,437,401 observations. This is daily price data for options with various strike prices and maturity dates on 449,721 unique stock/days. On average, there are 890 stocks per day on the 504 trading days in the sample.

Table 1
Option database filters

Filters	Filters in isolation		Filters in sequence			
	Observations excluded	Fraction of original excluded	Obsservations excluded	Fraction of original excluded	Observations remaining	Fraction of original remaining
					19,723,466	100%
$C, P < 0.375$	3,564,681	18.07%	3,564,681	18.07%	16,158,785	81.93%
tau > 180	3,866,290	19.60%	3,744,692	18.99%	12,414,093	62.94%
tau < 6	1,074,310	5.45%	533,918	2.71%	11,880,175	60.23%
$C > S$	0	0.00%	0	0.00%	11,880,175	60.23%
$C < S\text{-}PV(K)$	578,906	2.94%	442,774	2.24%	11,437,401	57.99%
$P < K\text{-}S$	0	0.00%	0	0.00%	11,437,401	57.99%
$P > K$	0	0.00%	0	0.00%	11,437,401	57.99%

This table presents the number of observations excluded by each filter applied to the options database. As in Bakshi, Cao, and Chen (1997), we delete observations in which call prices are higher than the underlying stock prices ($C > S$). We delete observations in which call prices are less than the present value of payoffs if exercised ($C < S - PV(K) - PV(\text{Div})$). We delete observations in which put prices are less than the current value of exercise ($P < K - S$). We delete observations in which put prices are above their strike prices ($P > K$). We also delete options with less than 6 calendar days to maturity or greater than 180 calendar days to maturity and options with a price less than \$0.375.

3. Results

The empirical results are ordered as follows. First, we address the significance to a market maker of its option to fail, both in the incidence of failing and in the relation of failing to high equity loan costs. Next, we relate equity loan costs to the price of a synthetic short position as determined by put-call parity, and we ask if this relation is sensitive to whether the option to fail is in the money—i.e., whether the rebate rate is less than zero. Then we gauge whether the mispricing of the synthetic short position is a result of expensive puts or cheap calls using implied volatility as a measure of price. We then calculate the expected cost of buy-ins, which is the product of the incidence of buy-ins and their execution quality. Using this actual incidence and price of buy-ins, we compute the market maker's net profits from providing synthetic shorts in hard-to-borrow situations. Finally, in response to the positive profits we document, we address the possibility that options market competition is limited when the underlying stock is hard-to-borrow.

3.1 Specialness and delivery failure

Our database shows the data supplier's short position, for each stock in the Russell 3000 and each day in 1998–99, and in particular it shows whether the position was achieved through borrowing, failing, or both. It also tells us the rebate received on borrowed shares, whether failed shares were bought in, and if so, at what price. Thus, we can sort short positions into five major categories: *General Collateral, Reduced Rebate, Reduced Rebate and Fail, Fail Only*, and *Buy-In*. *General Collateral* indicates that a stock has been loaned at the normal rebate rate, i.e., the stock is easy to borrow. *Reduced Rebate* indicates that the rebate rate is below the general collateral rate—i.e., the stock is

Table 2
Rebate rates, failure, and buy-in frequency

Panel A. Incidence of loan states in the database

Loan state	Frequency	Percentage of database	Average rebate rate
General collateral	1,379,594	91.24	4.98%
Reduced rebate	63,343	4.19	1.72%
Reduced rebate or fail	59,322	3.92	1.50%
Fail	9,655	0.64	0.34%
Buy-in	86	0.01	0.00%

Panel B. Rebate rates for failing and non-failing positions

	Rebate > 0	Rebate = 0	Rebate < 0	Total
No failing	98.61%	1.39%	0.00%	95.43%
Partial failing	59.36%	40.64%	0.00%	3.92%
Failing	10.35%	89.65%	0.00%	0.64%
Total	96.50%	3.50%	0.00%	100%

This table presents statistics on the 1998–99 rebate rate, fail, and buy-in database from a large options market maker. Panel A shows the overall incidence of five equity-loan states in the database: General Collateral, Reduced Rebate, Reduced Rebate/Fail, Fail Only, and Buy In, and the average rebate rate associated with each state. Panel B shows the percentage of daily stock positions in each one of the three categories: (i) No Failing, in which there are no shares failing delivery, (ii) Partial Failing, in which there is at least one share failing delivery, and (iii) Failing, in which every share is failing delivery. Percentages are based on the total number of observations, 1,512,000.

special. *Reduced Rebate and Fail* indicates that some shares have been borrowed at a reduced rebate, and that the market maker failed to deliver some shares that were sold short. *Fail Only* indicates that the market maker failed to deliver any of the shares in this short position. *Buy-In* indicates that the counterparty of the short-sale transaction is forcing delivery on some or all of the shares in the short position. Table 2, panel A, reports the incidence of each.

Consistent with earlier work, a large majority, 91.24%, of stocks are available for borrowing at general collateral rates. The remaining 8.76% are the specials. Breaking out this 8.76%, we find 4.19% in which borrowing simply continues at lower rebates, but in the remaining 4.57% the market maker fails to deliver, partially or completely. Failing is thus an important part of the story; more than half of the time the option to fail is used when stocks are on special. Any analysis of the relationship between short-sale impediments and options prices is at least incomplete, and perhaps severely biased, without consideration of the option to fail.

We would expect options market makers to fail more often as rebate rates fall to zero. Our sample bears this out. Panel B of Table 2 shows 89.65% of the failing positions occurring when rebates in our sample are at the lowest rate, zero, and only 1.39% of the *non*-failing positions have rebates at zero. So, failing is prevalent when rebates hit zero, and delivery is prevalent when rebates are positive. We also find, in unreported results, that the probability of at least some failure grows 15.66% for each 1% decrease in the rebate.[3] Thus we conclude that failure is tightly linked to low rebates.

3.2 Specialness and options prices

We expect options prices to reflect the costs of hedging, including the costs of short selling. We use our measure of short-sale costs, specialness, and two measures of options prices to characterize this relationship. First, we use put-call parity to measure misalignments of stock and options markets. Second, we refer to the options' implied volatilities, as calculated by the Cox-Ross-Rubinstein (1979) methodology, to gauge whether puts or calls are more responsible for what we find.

3.2.1 Put-call parity. The effect of short-sale costs on options prices can be seen via the European put-call parity relation. Put-call parity states that the value of a European call option plus the discounted value of the option's strike price is equal to the value of the underlying asset plus the value of a European put with the same strike price and maturity, $C + e^{-r\tau}K = P + S$, where C is the price of a European call option on stock S with strike price K, $e^{-r\tau}K$ is the present value of K, and P is a put option with strike price K. C and P are assumed to have the same time to maturity, τ.

This relationship allows a trader to replicate the payoffs of any single instrument in the equation with a combination of the other three instruments. For example, the stock price implied by this put-call parity relationship, or the *implied* stock price, is $S^i = C - P + e^{-r\tau}K$. For stocks with dividends paid during the life of the option, the present value of dividends is added to the right-hand side of the equation.

After computing the stock price implied by put-call parity, we compute the percentage deviation of the implied stock price from the actual stock price. This is computed by subtracting the implied stock price from the actual stock price and normalizing by the actual stock price:

$$\Delta_{j,t} = \frac{S_{j,t} - S^i_{j,t}}{S_{j,t}}, \tag{1}$$

where $S_{j,t}$ is the price of stock j on day t from the spot market and $S^i_{j,t}$ is the price of stock j on day t implied by put-call parity. We refer to $\Delta_{j,t}$ as put-call *disparity*. Table 3 shows the distribution of this measure, which shows some dispersion; the 5th percentile is −0.98% and the 95th percentile is 1.95%.

Some of this dispersion does not relate to arbitrage opportunities. Dividend-related early exercise differentiates the European options of the parity relation

[3] Taking those observations for which specialness is positive and the rebate rate is positive, we run a logistic regression of failing on specialness with cross-sectional and time-series fixed effects. The dependent variable is equal to 1 if there is any failing in a particular stock on a given day and is equal to zero otherwise. The coefficient estimate on specialness is 0.1455 (p-value < 0.0001). The odds-ratio point estimate is 1.1566.

Table 3
The distribution of put call disparity and specialness

	Put call disparity	Specialness	Rebate rate
Average	0.0036	0.48%	4.47%
Median	0.0028	0.00%	4.85%
Standard deviation	0.0179	1.30%	1.34%
Minimum	−0.9988	−0.07%	0.00%
Maximum	0.5617	5.80%	5.80%
5th Percentile	−0.0098	0.00%	0.00%
10th Percentile	−0.0053	0.00%	3.00%
90th Percentile	0.0140	1.92%	5.33%
95th Percentile	0.0195	4.50%	5.40%

This table describes the distribution of put-call disparity, specialness, and rebate rates in the sample of 4,560,217 strike-price- and maturity-matched put-call pairs. *Put-call Disparity* is the difference between the stock price and the options-implied stock price normalized by the stock price, i.e., $(S - S^i)/S$. *Specialness* is the difference between the general rebate rate and the specific rebate rate for a stock. *Rebate Rate* is the interest rate on cash collateral in a stock loan.

from the American options of the database; we address this in the following by excluding stocks that paid dividends in the last year. Early exercise also arises with deep in the money puts, which we address by using only the option pair with moneyness (S/K) closest to one, and shortest time to expiration. Near the money and close to maturity, options tend to be more actively traded; this also mitigates the effects of stale prices, which we further address by removing options for which the volume or open interest is equal to zero.

Another source of dispersion is microstructure effects. As Battalio and Schultz (2005) document, end-of-day prices deliver noisy estimates of actual arbitrage opportunities, as they may not be prices that were simultaneously available for trade. And even if prices are simultaneous and midpoint prices are exactly in line with put-call parity, if the stock price were at either an ask or a bid while the options were both at their midpoints, put-call disparity for the average stock in our sample would be 0.53%.[4] But as long as these measurement errors do not correlate with rebates, they do not interfere with our tests. That is, they are as likely to subtract from our measured profits as add to them.

We test the null hypothesis that short selling is not associated with put-call disparity by regressing put-call disparity on specialness and control variables. We find (p-values in parentheses):

$$\begin{aligned} \Delta_{j,t} = {} & \underset{(<0.001)}{-0.007} + \underset{(<0.001)}{0.060}{}^{*}\mathit{Specialness} + \underset{(<0.001)}{0.006}{}^{*}\mathit{Moneyness} \\ & + \underset{(<0.001)}{0.000}{}^{*}\mathit{Time\text{-}to\text{-}Maturity}, \end{aligned} \tag{2}$$

[4] The median bid-ask spread in stocks in our sample is \$0.23 (mean \$0.25). The median stock price in the sample is \$21.69 (mean \$27.04). If put and call prices are midpoints and the stock is at either the bid or the ask, then put-call parity will be different from zero by 0.23/(2*21.69) = 0.0053.

where *Moneyness* is defined as the stock price divided by the strike price and *Time-to-Maturity* is the number of calendar days to expiration of the option. We also include both time-series and cross-sectional fixed effects.

The significantly positive coefficient on *Specialness* confirms that as specialness increases, so does put-call disparity. Thus, specialness passes through to options prices, consistent with findings elsewhere. The new question is whether the *option to fail* passes through to options prices. Because the option to fail can reduce a market maker's shorting costs when rebates are negative, but not when they are positive, its effect would be a *weakening* of the relation between specialness and options prices as rebates go negative. That is, reducing the rebate from 2% to 1% increases the market maker's shorting costs by exactly that much, as failing is not a cheaper alternative to borrowing in either case, but reducing it from –1% to –2% increases it *less*, to the extent that the market maker is failing rather than borrowing. To test this hypothesis we need to use specialness data, which show it when the market rebate is negative, so we use the "custodial" data instead.

The test design is the same as before, but now focuses on stocks that are currently on special, and we add a regressor, Specialness*$ID_{\text{Rebate}<0}$. This regressor is zero when the rebate is positive, and is equal to the specialness, i.e., has the same value as *Specialness*, when the rebate is negative. The null hypothesis is that the coefficient on Specialness*$ID_{\text{Rebate}<0}$ is not significantly negative. We find:

$$\begin{aligned}\Delta_{j,t} = &-0.003 + 0.088^{*}\text{Specialness} - 0.048^{*}(\text{Specialness}^{*}ID_{\text{Rebate}<0}) \\ &\;(0.351) \qquad (<0.001) \qquad\qquad\qquad\quad (0.005) \\ &+0.002^{*}\text{Moneyness} + 0.000^{*}\text{Time-to-Maturity}. \\ &\;(0.261) \qquad\qquad\qquad\quad (<0.001)\end{aligned} \tag{3}$$

The regression rejects the null; the relation between rebates and options prices is indeed significantly weaker (at the point estimate, about 50% less) for negative than for positive rebates. This indicates that the option to fail plays a significant role in options pricing.

3.2.2 Distinguishing the effect on puts and calls. The relation between specialness and synthetic shorts indicates some combination of puts growing expensive and calls growing cheap. To gauge whether one is more important than the other, we need to separate puts from calls and relate their prices to the model. The testable questions become, are the implied volatilities of the puts significantly high, and are the implied volatilities of the calls significantly low?

For each option, we use implied volatility from OptionMetrics, which uses the industry-standard Cox, Ross, and Rubinstein (1979) binomial tree method for calculating implied volatilities. Using a fixed effects regression, we examine the relationship of implied volatilities with the moneyness, time to maturity, and specialness as well as time-series and cross-sectional fixed effects. The

Table 4
Implied volatilities and short-sale constraints

	Calls			Puts		
Intercept	1.053***	1.053***	1.052***	0.928***	0.918***	0.911*
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Moneyness	−0.038**	−0.038**	−0.038**	0.093***	0.092***	0.092***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Time to maturity	−0.000***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Specialness		−0.026			0.606***	
		(0.049)			(0.049)	
100 bp indicator			0.000			0.0168***
			(0.002)			(0.002)
Observations	84,915	84,915	84,915	84,915	84,915	84,915
R^2	0.664	0.664	0.664	0.678	0.678	0.678

This table presents estimation results from a panel regression of the following form:

$\sigma_{j,t}^{\text{implied}} = \gamma_0 + \gamma_1 \text{Moneyness}_{j,t} + \gamma_2 \text{Time - to - Maturity}_{j,t} + \gamma_3 \text{Specialness}_{j,t} + ID(j) + ID(t) + e_{j,t}$. $\sigma_{j,t}^{\text{implied}}$ is the implied volatility of a put or a call option for stock j on day t. Implied volatilities are calculated using the Cox-Ross-Rubinstein binomial tree method. One put and one call are selected on each stock every day; of the options with the shortest time to maturity, those with moneyness closest to one are chosen. Implied volatility is then regressed on *Moneyness*, *Time-to-Maturity*, and two specifications of *Specialness*: the actual value of specialness and an indicator that takes the value of one if the stock has specialness greater than 100 bp, and zero otherwise. The ID functions represent a fixed-effects treatment of both cross-sectional (by stock) and daily (by day) effects. ***Indicates statistical significance at the 0.1% level. **Indicates statistical significance at the 1% level. *Indicates statistical significance at the 5% level. Standard errors are included in parentheses.

estimation results from the following regression are presented in Table 4:

$$\sigma_{j,t}^{\text{implied}} = \gamma_0 + \gamma_1 \text{Moneyness}_{j,t} + \gamma_2 \text{Time-to-Maturity}_{j,t} + \gamma_3 \text{ Specialness}_{j,t} + e_{j,t}, \tag{4}$$

where moneyness is defined as S/K, and time to maturity is measured in calendar days. We include fixed effects for both time-series and cross-sectional effects. Consistent with the results for index options from Derman and Kani (1994) and Longstaff (1995), we find that the implied volatility of put options increases with moneyness. The coefficient on moneyness for call options is statistically negative but the slope is almost flat. Consistent with Bakshi, Cao, and Chen (1997), our regression results show that implied volatility decreases with time to maturity.

For puts, both the presence and the magnitude of specialness are statistically significant and positive. Thus we conclude that specialness increases the prices of puts. Call prices do not show this sensitivity; neither the presence nor the magnitude of specialness has a statistically significant effect on call prices. Thus, when we separate the synthetic short into its components, we detect a significant positive effect of specialness on the cost of buying puts, and no effect on the revenue from writing calls.

3.3 Abnormal profits

3.3.1 Buy-in costs. The other cost of failing, besides the foregone interest from the with held purchase price, is the expected cost of being bought in. When

Table 5
Buy-in execution

	Execution cost	Buy-in quantity	Trading days to settlement
Mean	0.0053	9,512	3.01
Median	0.0028	3,915	3
SD	0.0178	14,450	0.24
t-stat	2.75		
p-value	0.01		

This table presents statistics on the execution of buy-ins in 1998 and 1999 for a major market making firm. After merging the database with TAQ, there are 85 buy-in observations on 24 unique stocks. The execution cost of the buy-ins is examined by comparing the buy-in prices to the volume-weighted average price (VWAP) over the trading day, $(S_{BUYIN} - S_{VWAP})/S_{VWAP}$. For *Execution Cost*, we report the mean, median, and standard deviation of the execution costs. Additionally, we report a t-test of the null hypothesis that the difference is zero. If multiple buy-in events are recorded on a single day, the buy-in price used in the calculations is the quantity-weighted execution price. We report the quantity of shares bought in, *Buy-In Quantity,* and the number of trading days from buy-in to settlement, *Trading Days to Settlement*.

the market maker is bought in, the clearing corporation executes the purchase, and the market maker must execute a new short sale to restore its position. Thus, the market maker's expected buy-in cost is the probability of a buy-in times this round trip cost. Table 2 shows that 86 of the 69,063 failing positions, or 0.12%, were bought in over the 2-year period. Taking this realization as the expected incidence of buy-ins, the expected incidence of buy-in costs is this figure times the expected transactions cost, conditional on a buy-in.

Because the clearing corporation executes the buy-in, execution quality may not be optimized, and may therefore be costly to the market maker. To gauge this other leg of buy-in cost we relate the transaction prices of the 86 buy-ins to the prevailing market prices. Table 5 shows that the buy-in trades are executed at prices 0.53% worse than the volume-weighted average price (VWAP) for the given stock on the buy-in day. The departure from VWAP is statistically significant, but even so, if we assume that the market maker pays the same 0.53% to put the short back on, the overall expected buy-in cost is (0.12%)(2)(0.53%), or 0.1 bp. So the expected buy-in cost is, for our market maker at least, vanishingly small.

3.3.2 Abnormal profit strategies. In this section we combine the data on rebate rates and buy-in costs to calculate the profitability—to the market maker who provided our data—of providing synthetic shorts on hard-to-borrow stocks. Here we consider why the profits we document could be available in equilibrium.

Our profitability measure follows a simple trading strategy, designed to avoid the attribution issues documented by Battalio and Schultz (2005). In particular, we decide whether to put on the trade based on whether the stock is on special, *not* on whether our database shows disparity, and then we hold the trade to expiration. If instead we went in and out of the trade depending on the apparent disparity in the data, we would mistake some measurement error for profitability.

We focus on liquid options and reduce the influence of known biases by selecting the option pair with maturity as short as possible and moneyness closest to one. We also reduce the incidence of early exercise bias, without risking any look-ahead bias, by using only stocks that didn't pay dividends in the past year.

Our database allows us to calculate the profits net of the exact shorting costs. If our market maker received a rebate, we use that rebate, but if they failed, we use a rebate of zero. Assuming the market maker's opportunity cost of capital at time t is the risk-free rate $r(t)$, and denoting his concurrent rebate for a given stock $q(t)$, the short-sale cost paid by the borrower on day t can be written as $r(t) - q(t)$. If our data show that a stock was bought in, we subtract the cost of the buy-in by subtracting the buy-in price from the stock's VWAP that day. The assumption is that the market maker keeps the position going by shorting anew at VWAP on the same day. The profits from this strategy can be written as follows:

$$\underbrace{[S(0) - S(T)]}_{\text{Short Stock Position}} + \underbrace{[S^i(T) - S^i(0)]}_{\text{Synthetic Long Position}} - \underbrace{\left[\sum_{t=0}^{T} S(t)(r(t) - q(t))\right]}_{\text{Reduced Rebate Costs}}$$

$$\underbrace{- 1_{\text{Buy-In Indicator}} \left[\sum_{t=0}^{T} (S_{\text{Buyin}}(t) \quad S_{\text{VWAP}}(t))e^{r(T-t)}\right]}_{\text{Buy-In Round Trip Transaction Cost}}, \tag{5}$$

where the position is opened at $t = 0$ and closed at $t = T$.

We look at the profits to short selling actual stock and buying synthetic stock whenever that stock goes on special. Such a strategy would involve 6,086 option pairs and yield an average profit of \$0.1346 per trade. The profit is statistically significant with a p-value of less than 0.001, and corresponds to \$13.46 per options contract. Thus, the market maker profits, in equilibrium, by providing these synthetic shorts. In our final section we propose and test a hypothesis on the reason that this happens.

3.4 Negative stubs

Mitchell, Pulvino, and Stafford (2002; hereafter MPS), and Lamont and Thaler (2003) document the existence of a phenomenon where specialness and delivery failure can play an important role. This phenomenon, often referred to as a "negative stub," is an apparent mispricing between a parent firm and its subsidiary in which stock prices imply that a firm's market value is less than the value of its ownership stake in a subsidiary firm. In the absence of trading costs and constraints, this mispricing may indicate an arbitrage opportunity where the trading strategy would involve buying the parent's shares and selling the subsidiary's shares short. In our context, each incidence of a negative stub presents an opportunity to measure changes in our key statistics when the source of specialness is easily identifiable as the increase in borrowing demand arising from the negative stub arbitrage.

We compare specialness, delivery failure, and put-call disparity for the subsidiaries across two sets of times: the times when stub values are negative, and the times when they are not. Starting with the sample of negative stubs from MPS (2002), we find nine subsidiaries with rebate data, from our prime broker database, on both negative stub and non-negative stub dates. We see that specialness is significantly higher when there is a negative stub situation. Average specialness is 0.76% for stock/days with negative stubs and 0.61% for stock/days without negative stubs. The difference is statistically significant with a p-value of 0.013. Similarly, delivery failure is significantly more likely when there is a negative stub situation. Specifically, delivery failure happens in 20.06% of the stock/days with negative stubs and 15.45% of the stock/days without negative stubs. The difference is statistically significant with a p-value of 0.0007. Within the negative stub sample, we find three subsidiaries with options data in our cleaned database. In this subsample, the put-call disparity is significantly higher when subsidiaries have a negative stub situation. Disparity is 0.0027% for stock/days with negative stubs and 0.0002% without negative stubs. The difference is statistically significant with a p-value of 0.0048.

These negative stub results reinforce the results above. First, the increase in specialness indicates that the market is working as we would expect: the appearance of the negative stub arbitrage boosts demand by short sellers for loans of the subsidiaries' shares, thus driving the prices up. Second, the increase in delivery failure indicates that our data provider exercises its option to fail to deliver when high borrowing demand makes it more valuable to do so. Finally, consistent with Lamont and Thaler (2003), the increase in put-call disparity in the negative stub situations reflects the spread of short-selling costs across the different channels of short exposure. The results in the negative stub sample are consistent with what we find in the whole sample, and these situations allow us to study a sample in which the source of specialness is an easily identifiable arbitrage opportunity.

3.5 Why aren't abnormal profits competed away?

What accounts for the equilibrium profitability we document? Why aren't more market makers buying these cheap synthetic long positions, bidding them up to zero profitability? We hypothesize that the clearing corporation handles fails in a way that favors higher volume market makers, resulting in weaker competition when stocks grow special.

The hypothesis follows from how the clearing corporation assigns buy-ins. If a fail must be bought in, this buy-in is assigned to the oldest fail (see Appendix A). Assuming that a higher volume market maker is more likely to move from a short position to flat (or positive) and back again, it is *less* likely to have the oldest fail and therefore less likely to be bought in. Thus, we hypothesize that higher volume market makers, such as our data provider, enjoy a cost advantage with hard-to-borrow stocks, and that this advantage limits competition to make markets in the affected options.

In the analysis that follows, we test the hypothesis in several ways. First, we test whether our data provider delivers an increasing share of short exposure as specialness increases. If our explanation for the persistence in profitability is right, we would expect to see this market maker buying more cheap synthetic long positions, and hedging those positions, when specialness is high. Second, we test whether competition changes when stocks go on special. If specialness weakens competition by preventing low volume market makers from posting competitive quotes, then spreads should widen when stocks go on special. Third, we test whether low turnover increases the likelihood of buy-ins. If high volume protects our market maker from buy-ins, then the incidence of buy-ins should increase when turnover decreases. Similarly, because our market maker is one of the largest, its buy-in protection should be relatively strong, so its overall incidence of buy-ins should be relatively light, a hypothesis we can test in back-of-the-envelope fashion. Fourth, we test whether an increase in market maker competition squeezes the profitability of put-call parity arbitrage. We do this by testing for a relationship between the profitability of the put-call arbitrage and the increased options market competition of multiple listed options.

3.5.1 Market maker's share of short interest. In our first test, we look at the relationship between short interest and specialness. If our market maker faces lower competition when specialness is high, then its market share of short exposure, i.e., its total short position as a fraction of economy-wide short interest, should grow as specialness grows. This would suggest that shorting via this options market maker, rather than some other way, becomes more attractive as shorting constraints tighten. Regressing the market share *(Market Maker's SI)/(Market SI)* on *Specialness,* we find (p-values in parentheses):

$$\text{(Market Maker's SI) / (Market SI)} = \underset{(0.0280)}{-0.069} + \underset{(0.020)}{0.040} * \text{Specialness.} \tag{6}$$

Market share increases significantly with specialness, indicating that shorting frictions encourage shorting via this options market maker, and therefore that shorting frictions impose less cost on this options market maker than on traders in general.

So this market maker gains short interest market share as specialness grows, but in principle *all* options market makers could be gaining short interest market share. To address competition *among* options market makers we need a measure of the current competition to make a market in a stock's options, and the natural candidate is the price charged, the current bid-ask spread. In our second test, we examine the connection between bid-ask spreads and specialness.

3.5.2 Option market spreads. Option spreads are subject to limits that often bind.[5] The SEC, which sets these limits, finds that quoted spreads are at their maxima between 21% and 57% of the time (see Securities and Exchange Commission, 2000; hereafter SEC for details). This is consistent with our sample; we find 35% of put options and 32% of call options at their maximum spreads. Thus the relevant measure of spread width is whether it is at the maximum.

Accordingly, to relate spreads to specialness, we fit a probit model where the dependent variable $ID_{Max.Spread}$ indicates that both put and call options are quoted at their maximum spread and specialness is included as an explanatory variable. The estimation results from the probit model with firm-fixed effects are as follows (p-values in parentheses):

$$\begin{aligned} \text{ID}_{\text{Max. Spread}} = {} & 1.122 * \text{Specialness} && - 0.151 * \text{ID}_{\#\text{EXCH} > 1} \\ & (<0.001) && (<0.001) \\ & - 0.138 * \text{Ln}\,(\text{Volume}) && + 0.004 * \text{Price} \\ & (<0.001) && (<0.001) \\ & + 0.888 * \text{Implied Vol.} && + 0.086 * \text{IO} \\ & (<0.001) && (<0.001) \\ & + 4.041 * \text{Turnover} && - 0.113 * Ln\,(\text{Mkt. Cap.})\,. \\ & (<0.001) && (<0.001) \end{aligned} \tag{7}$$

The probit rejects the null; incidence of maximum spreads increases significantly with specialness. Thus the evidence supports the hypothesis that specialness weakens options market competition.

To have a clean test of this hypothesis, we have to control for other possible determinants of the bid-ask spread. We focus on the determinants of option bid-ask spreads shared by at least two of the following four papers: Battalio, Hatch, and Jennings (2004); De Fontnouvelle, Fishe, and Harris (2003; hereafter DFH); Mayhew (2002); and Neal (1987). We include a multiple exchange listing indicator, $ID_{\#EXCH>1}$, which is one if the option trades on more than one exchange and zero otherwise.[6] In a result that is consistent with the

[5] The Securities and Exchange Rule 1014(c)(i)(A) and Advice F-6 prescribe maximum quote spreads for equity options. The rule establishes maximum widths as follows: \$0.25 for options priced between \$0.50 and \$2, \$0.375 for options priced between \$2 and \$5, \$0.5 for options priced between \$5 and \$10, \$0.75 for options priced between \$10 and \$20, and \$1 for options priced above \$20.

[6] The multiple exchange indicator is constructed using releases of the Option Clearing Corporation's Option Symbol Directory, which lists exchanges (AMEX, CBOE, PHLX, and PCX) for each option. Dates before October 1998 are assigned the exchange status from the October 1998 directory. After that, dates are assigned the exchange status of the closest report. Specifically, days before January 15, 1999, are assigned the October 1999 status; days between January 15, 1999, and September 1, 1999, are assigned the April 1999 status; and days after September 1, 1999, are assigned the January 2000 status.

previous literature, we find a statistically significant coefficient estimate of −0.15, indicating that options listed on multiple exchanges have a lower probability of having quotes at their maximum spread. This result is consistent with the idea of market competition decreasing the probability of maximum spreads. Also consistent with the previous literature, our measure of option-series volume, *Ln*(Volume), which is the log of the sum of call and put volume over the previous 22 days, is significantly negative. The coefficient of –0.14 indicates that option volume decreases the probability that an option trades at its maximum bid-ask spread. Previous literature finds option price to have a significant positive relationship to bid-ask spreads. We also find a positive relationship; our measure of options prices, *Price*, which is the average of bid and ask prices, has a statistically significant coefficient estimate of 0.004. Options-implied stock volatility has been a mixed predictor in the options bid-ask spread previous literature. Table 5 of DFH (2003), for example, generally finds a significantly positive relationship between implied volatility and bid-ask spreads even though Mayhew (2002) finds a negative relationship. Our results are similar to the findings of DFH (2003). Using our measure of implied volatility, *Implied Vol.*, which is the average of put and call implied volatility, we find a statistically significant coefficient of 0.89, indicating that an increase in implied volatility increases the probability of quotes at their maximum spread. Overall, we find the controls largely confirm previous findings on the determinants of the bid-ask spread, and we find that specialness is an important new predictor of options being quoted at their maximum spreads.

In addition to controlling for previously known determinants of the bid-ask spread, we must also ensure that specialness, and not the covariates of specialness, is driving the result. To this end, we include the three variables most highly correlated with underlying stocks' specialness from D'Avolio (2002): institutional ownership, turnover, and market capitalization. We measure institutional ownership, or *IO*, by dividing the number of shares held by institutions in the most recent 13f SEC filings by the number of shares outstanding. We find that institutional ownership is an important predictor of maximum spread incidence; we find a statistically positive coefficient estimate of 0.09. Turnover is a statistically significant determinant of maximum option spread incidence. Using our measure of turnover, which is defined as the previous month's average daily stock market volume divided by the previous month's shares outstanding, we find a statistically positive coefficient estimate of 4.04. We find that the natural log of market capitalization has a negative effect on the incidence of maximum spreads. In particular, the statistically significant coefficient of –0.11 indicates that larger stocks are less likely to have options quoted at their maximum spreads. Even though these measures are statistically related to the maximum spread incidence, they are not driving the result; after controlling for these covariates of specialness, the statistically positive coefficient of 1.12 supports the hypothesis that specialness weakens options market competition.

Table 6
Market competition and implied stock prices

Variable	Estimate		
Intercept	−0.007***	−0.007***	−0.002
	(0.002)	(0.002)	(0.001)
Specialness	0.060***	0.065***	0.116***
	(0.0030)	(0.003)	(0.007)
Specialness*$ID_{DATE>AUG99}$		−0.026***	
		(0.006)	
Specialness*$ID_{\#EXCH>1}$			−0.076***
			(0.008)
Moneyness	0.006***	0.006***	0.004***
	(0.001)	(0.001)	(0.001)
Time-to-maturity	0.000***	0.000***	0.000**
	(0.000)	(0.000)	(0.000)
R^2	0.155	0.155	0.165
Number of observations	84,915	84,915	58,804

This table presents estimates from a panel regression of the following form: $\Delta_{j,t} = \mathrm{a} + \mathrm{bSpecialness}_{j,t} + \mathrm{c\,Moneyness}_{j,t} + \mathrm{d\,Time\text{-}to\text{-}Maturity}_{j,t} + e_{j,t}$. $\Delta_{j,t}$ is the put-call disparity of stock j on day t. *Specialness*$_{j,t}$ is the difference between the general rebate rate on day t and the specific rebate rate for a stock j on day t. *Moneyness*$_{j,t}$ is the price of stock j on day t divided by the strike price of the option pair. *Time-to-Maturity* is the number of calendar days to expiration of the option. The second and third specifications include indicator variables that test the relationship between options market competition and put-call disparity. $ID_{DATE>AUG99}$ is an indicator that takes the value of one for dates after August 1999. $ID_{\#EXCH>1}$ is an indicator that takes the value of one for options trading on multiple exchanges. The sample and fixed effects are described in Table 4. The asterisks denote statistical significance as follows: *** significant at 0.1%, ** significant at 1%, and * significant at 5%. Standard errors are included in parentheses.

3.5.3 Turnover and buy-in protection. The third test of the hypothesis that turnover gives large market makers a cost advantage by protecting them from buy-ins is an examination of whether this advantage dissipates when volume drops. That is, we can test whether our data provider's probability of being bought in increases as option turnover declines. We do this by fitting a probit to the sample of all failing stocks, where the dependent variable $ID_{\text{Buy-in}}$ indicates whether (1) or not (0) the position is bought in, and the explanatory variables are *Option Turnover* (the average daily put and call volume over open interest) along with other variables associated with buy-ins. Estimation results are as follows (p-values in parentheses):

$$\begin{aligned}
\mathrm{ID}_{\text{Buy-in}} = {} & -3.367 * \text{Intercept} && - 0.287 * \text{Option Turnover} \\
& (<0.001) && (0.035) \\
& + 0.330 * \text{Specialness} && + 0.082 * \text{Log (SI)} \\
& (<0.001) && (0.061) \\
& - 0.353 * \text{Log (Shares)} && - 0.041 * \text{SD} + 4.240 * \mathrm{ID}_{\text{Price} < \$5}. \\
& (<0.001) && (0.086)\ (<0.01)
\end{aligned} \tag{8}$$

We find that turnover is significantly negative, as predicted: the less options turn over, the more the position is bought in.[7] To control for factors

[7] The option turnover variable is used to test the hypothesis that lowering the age of delivery failure lowers the position of the market maker on the list of buy-in allocations. The hypothesis predicts a negative coefficient. Consequently, the p-value for option turnover is from a 1-sided test.

other than turnover that may influence the probability of a buy-in, we include several additional variables. We include monthly short interest, log*(SI)*, the number of shares outstanding, log*(Shares)*, the standard deviation (SD) of daily returns over the previous 6 months, and the indicator variable $ID_{\text{Price}<\$5}$, which takes the value of 1 if the share price is below \$5 and 0 otherwise.

Furthermore, we can see directly that our data supplier experiences an abnormally low incidence of buy-ins on its fails. On the average day, across the 502 sample trading days, this large market maker is failing on 4.4 shares. This is about 1.75% of the ~250M fails on all NYSE/AMEX/NASDAQ stocks on the average day, in Boni (2006) (see Figure 1 of that paper). In Table 2 we see that our market maker experienced 86 buy-ins across the 502 days, or about 1/6 buy-in per day. If this is about 1.75% of buy-ins, we should see about 10 buy-ins per day. But the DTCC reports more than 4,300 buy-in notices per day,[8] and the fraction of notices that result in buy-ins is presumably greater than 1/430. Thus our data provider's fails appear relatively unlikely, compared to other traders' fails, to beget buy-ins.

3.5.4 Multiple listing and competition. The final test of our explanation for persistence in put-call parity arbitrage is whether increasing competition generally reduces put-call disparity. We focus on a natural experiment. In August 1999, the Chicago Board Options Exchange, the American Stock Exchange, the Philadelphia Stock Exchange, and the Pacific Exchange all began listing options that had been exclusively listed on competing exchanges, and as DFH documents, this succeeded in increasing competition to make markets. Thus, if a lack of market making competition is indeed behind the market misalignment found above, the incidence of disparity should decrease across August 1999. We construct two measures of competition: an indicator of dates after August 1999, and a stock-specific measure of multiple exchange listing. We present estimates from a regression of put-call disparity on these competition measures and other control variables. We find a statistically negative coefficient of −0.026 on the marketwide measure and therefore reject the null that an increase in market making competition did not decrease put-call disparity. Similarly, we find a statistically negative coefficient of −0.076 on our stock-specific measure of multiple listing, which is also consistent with multiple listings decreasing put-call disparity. Although this evidence does not set our data supplier apart from other market makers, the multiple listing results are further evidence that reduced competition among options market makers allows put-call disparity to persist.

[8] This figure represents only the notices transmitted via the DTCC's Participant Exchange service; see "DTCC Will Automate and Streamline Buy-In Notification for Securities," a DTCC press release, at http://www.dtcc.com/PressRoom/2005/buyin.html.

4. Conclusion

We show that the option to fail is significant to both the trading and pricing of equity options. We show that it is often in the money, and that when it is, market makers profit and so do their customers. The profit to market makers is puzzling, considering their competitiveness, but we resolve this puzzle by documenting limits to competition in options on hard-to-borrow stocks, and tracing these limits to the clearing corporation's rules for assigning buy-ins.

The popularity of failing and the price improvements it provides short sellers encourage us to step back and consider the economic case for delivery. Delivery provides 100% insurance to both sides of a trade; by exchanging cash for securities, traders eliminate 100% of their mutual exposure. One hundred percent insurance is unambiguously optimal when it is free, but not when it is costly, so the search for efficiency should bring traders to a mechanism for buying less insurance at a lower price when delivery is costly. This appears to be what they get by failing and margining through the clearing corporation. So, although failing may sound mischievous or abusive, both our results here and basic economic reasoning indicate that its role is positive.

The SEC has taken a cautious approach by introducing its new regulation SHO, which strengthens delivery requirements for "threshold securities," those with substantial current fails. Its effect on large market makers such as our data provider is likely small, since their hedging trades are still exempt from the locate requirement, and their fails do not age much (see Boni, 2006, for a complete description of the regulation), but it has the potential to alter the cost of short exposure, so its impact is an important new empirical question. Fortunately the lists of threshold securities are public, so proprietary data may not be necessary to answer it.

Appendix A: The Details of Short Selling and Delivery

Short sellers sell stock they do not own. In the United States, exchange procedure generally requires short sellers to deliver shares to buyers on the third day after the transaction ($t + 3$).[9] Short sellers typically borrow stock and use the proceeds from the sale as collateral for the loan. Additionally, regulators and brokerages impose varying margin requirements on short positions. To close, or cover, the position, the short seller buys shares and returns the shares to the lender.

A.1 Borrowing and rebate rates

Typically, a short seller borrows shares from her broker. The proceeds from the short sale are used as collateral for the stock loan. The collateral earns interest, and the broker returns some of the interest to the short seller. The interest rate

[9] See Bris, Goetzmann, and Zhu (2004) for a description of short selling in other countries.

the short seller earns is known as the rebate rate. Rebate rates are generally lower for smaller investors, but for a given investor, lower rebate rates indicate more expensive loans. The majority of loans are cheap, but there are a few expensive loans in stock specials.[10]

Specials tend to be driven by episodic corporate events resulting in arbitrage opportunities for short sellers (see Geczy, Musto, and Reed, 2002; or D'Avolio, 2002; for examples). Well-placed investors, such as hedge funds, will be able to borrow stock specials and will earn the reduced rebate.

A.2 Short selling when borrowing is difficult

Exchange rules require most market participants to demonstrate that they can obtain hard-to-borrow shares before they short sell.[11] Market makers require an affirmative determination of borrowable or otherwise attainable shares. In market parlance, the short seller needs a "locate" before short selling. However, there is an exception to the rule. An example is NASD's rule 3370(b), which exempts the following transactions from the affirmative determination requirement: "... bona fide market making transactions by a member in securities in which it is registered as a NASDAQ market maker, to bona fide market maker transactions in non-NASDAQ securities in which the market maker publishes a two-sided quotation in an independent quotation medium, or to transactions which result in fully hedged or arbitraged positions."

A.3 Fails and buy-ins

If the short sale is made on day t, the short seller's clearing firm generally delivers shares on day $t + 3$. However, the National Securities Clearing Corporation (NSCC) procedures state that "each member has the ability to elect to deliver all or part of any short position."[12] If a clearing firm decides to deliver less than the full amount of shares to its buyers, the firm is failing to deliver shares.

If the clearing firm fails, the best-case scenario for the short seller is for the buyer's broker to allow the fail to continue as long as the short position is open. In this case, the short seller's cost of short exposure is the lost interest on the transaction amount. When borrowing shares, the short seller would also lose the full interest income on his collateral in the case of a zero rebate rate. Economically, a failed delivery is the same as delivery of borrowed stock at a zero rebate rate as long as the buyer's broker allows the fail to continue.

---

[10] Fitch IBCA's publicly available report "Securities Lending and Managed Funds" estimates that the industry average spread from the fed funds rate to the general collateral rate on U.S. equities is 21 bps.

[11] During our sample period, NYSE Rule 440C and NYSE Information Memorandum 91–41 require affirmative determination (a "locate") of borrowable or otherwise attainable shares for members who are not market makers, specialists, or odd-lot brokers in fulfilling their market making responsibilities. Similarly, NASD Rule 3370 and NASD Rules of Fair Practice, Article III, Section 1, Interpretation 04 Paragraph (b)(2)(a) (See SEC Release No. 34–35207), and Securities Exchange Act Release No. 27542 (AMEX) require affirmative determination of borrowable shares during the period treated in the paper (SEC Release No. 34–37773).

[12] NSCC Procedures, VII.D.2.

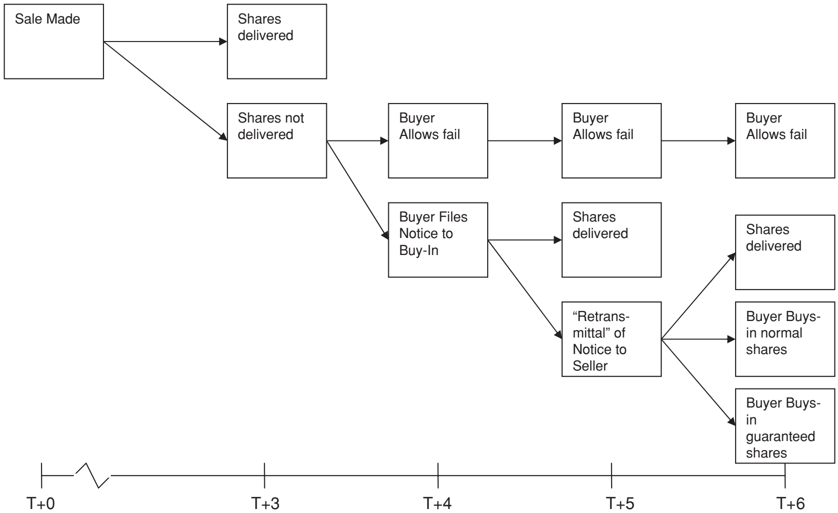

Figure 1
Settlement and buy-in timing.
Delivery status as a function of trading date. The nodes represent possible delivery outcomes. The horizontal axis represents the minimum number of trading days required to reach each outcome.

In the worst-case scenario, the buyer's broker insists on delivery by filing a notice of intention to buy in with the NSCC at $t+4$ in accordance with NSCC's Rule 10.[13] The notice is retransmitted from the NSCC to the seller's broker on $t+5$, and the seller has until the end of day $t+6$ to resolve the buy-in liability. If the seller does not resolve the liability, a "buy-in" occurs: the buyer purchases shares on the seller's account to force delivery.[14] If her position is bought in, the seller may then sell short again to establish the short position. In this case, the short seller will pay the execution costs of the buy-in and the following short sale every 6 days.[15] Figure A1 shows the sequence of events in each scenario.

The NSCC allocates buy-ins across clearing firms, and clearing firms allocate buy-ins across clients. Failing clients can protect themselves against buy-ins at both levels. In the first stage, the NSCC ranks clearing firms according to the date of failed deliveries, and the NSCC allocates buy-ins to the clearing firms

[13] The Securities and Exchange Commission's Customer Protection Rule requires clearing firms to possess shares in fully paid accounts. Clearing firms may attempt to acquire shares to be in compliance with the SEC's rule.

[14] The seller's clearing firm buys shares in a buy-in for NYSE and AMEX stocks; the buyer's clearing firm buys in shares of NASDAQ stocks.

[15] NASD Rule 11810(c)(1)(B) gives buyers the option to buy guaranteed delivery shares, and there have been complaints regarding the purchase price of guaranteed shares. A limited supply of guaranteed delivery shares, combined with the transparency of the underlying purpose for the purchase, may inflate prices. Second, according to NASD Regulation's general counsel Alden Adkins in Weiss (1998), "There are no hard and fast rules dictating the prices at which buy-ins can take place. But [Adkins] says the prices must be 'fair'—and that the person who sets the price must be prepared to defend it."

with the oldest failed delivery first.[16] As a result, clearing firms that frequently change from short to long net positions are less likely to be bought in.

Once the NSCC allocates buy-ins to a clearing firm, that clearing firm must allocate buy-ins among its clients. Clearing firms have discretion over this second stage of the selection decision, and, unlike the first stage, there are no marketwide rules.

Anecdotal evidence suggests that clearing firms use their discretion; they allocate a disproportionately small number of buy-ins to protected clients.

Appendix B. Risk-Free Interest Rates

We construct a database of daily risk-free interest rates using Federal Reserve 1-, 7-, 15-, 30-, 60- and 90-day AA financial commercial paper discount rates, which we convert to bond-equivalent yields.[17] The risk-free rate corresponding to option maturity is calculated by linearly interpolating between the two closest interest rates. For example, the risk-free rate for an option with a maturity of 6 days would be calculated by linearly interpolating between the 1-day and the 7-day discount rates.

The method of linear interpolation is an approximation to the true term structure, and the error inherent in the approximation is greatest for near-term maturities. By using the rates on commercial paper, the error is minimized relative to rates on treasury bills or other fixed-income instruments that are only reported for greater maturities. As a check on our procedure, we also calculate the risk-free rate with daily GOVPX data on treasury bills using a procedure similar to that of Bakshi, Cao, and Chen (1997). The correlation coefficient between the 3-month AA financial commercial paper rate and the 3-month treasury bill rate reported by the Federal Reserve is 0.98. As a further check, we regress our 3-month commercial paper rate on the Federal Reserve's 3-month T-bill rate from September 1997 to August 2001. The intercept is not significantly different from zero, the slope is statistically significant (the coefficient is 0.90), and the R^2 is 0.95.

[16] This description provided here is a slight simplification of the actual procedure. For a more specific example of what really happens, assume that $N + 0$ represents the date the Buy-In Notice is filed. Filing such a notice will give the firm higher priority in settlement on the first business day after filing, $N + 1$, and on the second business day after filing, $N + 2$, if the long position remains unfilled. On date $N + 1$, if the position remains unfilled, NSCC submits "retransmittal notices" to the firm(s) with the oldest short position in the buy-in stock. These notices specify the buy-in liability for the short firm and the name of the long firm instigating the buy-in. "If several firms have short positions with the same age, all such members are issued retransmittal motices, even if the total of their short positions exceeds the buy-in position." Once they receive the retransmittal notice, other settling trades may move them to a flat or even a long position in the stock but do not exempt them from their buy-in liability. The short firm has until the end of day $N + 2$ to resolve their buy-in liability. Before the retransmittal notice is received, a buy-in liability is removed once a net long position of sufficient size is established.

[17] Bond Equivalent Yield = (Discount/100)(365/360)/(1-(Discount/100)(Time to Maturity/360))
This is equivalent to the yield formula reported in the *Wall Street Journal* and is commonly used in option markets and for debt instruments with maturities of less than 1 year.

References

Bakshi, G., C. Cao, and Z. Chen. 1997. Empirical Performance of Alternative Option Pricing Models. *Journal of Finance* 52:2003–49.

Bakshi, G., N. Kapadia, and D. Madan. 2003. Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *Review of Financial Studies* 16:101–43.

Battalio, R., B. Hatch, and R. Jennings. 2004. Toward a National Market System for U.S. Exchange-Listed Equity Options. *Journal of Finance* 59:933–62.

Battalio, R. H., and P. H. Schultz. 2005. Options and the Bubble. Working Paper, University of Notre Dame.

Black, F., and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81:637–54.

Boni, L. 2006. Strategic Delivery Failures in U.S. Equity Markets. *Journal of Financial Markets* 9:1–26.

Bris, A., W. N. Goetzmann, and N. Zhu. 2004. Efficiency and the Bear: Short Sales and Markets Around the World. Working Paper, Yale University.

Christoffersen, S. C., C. G. Geczy, D. K. Musto, and A.V. Reed. 2005. Cross-Border Taxation and the Preferences of Taxable and Non-Taxable Investors: Evidence from Canada. *Journal of Financial Economics* 78:121–44.

Christoffersen, S. C., C. G. Geczy, D. K. Musto, and A. V. Reed. 2007. Vote Trading and Information Aggregation. *Journal of Finance*, forthcoming.

Cox, J. C., S. A. Ross, and M. Rubinstein. 1979. Options Pricing: A Simplified Approach. *Journal of Financial Economics* 7:229–63.

D'Avolio, G. 2002. The Market for Borrowing Stock. *Journal of Financial Economics* 66:271–306.

De Fontnouvelle, P., R. P. H. Fishe, and J. H. Harris. 2003. The Behavior of Bid-Ask Spreads and Volume in Options Markets During the Competition for Listings in 1999. *Journal of Finance* 58:2437–64.

Dennis, P., and S. Mayhew. 2002. Risk-Neutral Skewness: Evidence from Stock Options. *Journal of Financial and Quantitative Analysis* 37:471–93.

Derman, E., and I. Kani. 1994. Riding on a Smile. *Risk* 7:32–39.

Duffie, D., N. Gârleanu, and L. H. Pedersen. 2002. Securities Lending, Shorting, and Pricing. *Journal of Financial Economics* 66:307–39.

Dumas, B., J. Fleming, and R. E. Whaley. 1998. Implied Volatility Functions: Empirical Tests. *Journal of Finance* 53:2059–106.

Fleming, M. J., and K. D Garbade. 2002. When the Back Office Moved to the Front Burner: Settlement Fails in the Treasury Market After 9/11. *Federal Reserve Bank of New York Economic Policy Review* 8:35–57.

Geczy, C. C., D. K. Musto, and A. V. Reed. 2002. Stocks Are Special Too: An Analysis of the Equity Lending Market. *Journal of Financial Economics* 66:241–69.

Hull, J. C. 2000. *Options, Futures, and Other Derivatives*. Upper Saddle River, NJ: Prentice Hall.

Jarrow, R., and M. O'Hara. 1989. Primes and Scores: An Essay on Market Imperfections. *Journal of Finance* 44:1263–87.

Johnson, O. E. G. 1998. The Payment System and Monetary Policy. *International Monetary Fund Papers on Policy Analysis and Assessment* 98:4.

Jones, C. M., and O. A. Lamont. 2002. Short Sale Constraints and Stock Returns. *Journal of Financial Economics* 66:207–39.

Kahn, C. M., and W. Roberds. 2001. The CLS Bank: A Solution to the Risks of International Payments Settlement? *Carnegie-Rochester Conference Series on Public Policy* 54:191–226.

Klemkosky, R. C., and B. G. Resnick 1979. Put Call Parity and Market Efficiency. *Journal of Finance* 42:1141–55.

Lamont, O. 2004. Short Sale Constraints and Overpricing: *The Theory and Practice of Short Selling*. New Hope, PA: Frank J. Fabozzi Associates.

Lamont, O. A., and R. H. Thaler. 2003. Can the Market Add and Subtract? Mispricing in Tech Stock Carve-Outs. *Journal of Political Economy* 111:227–68.

Longstaff, F. A. 1995. Option Pricing and the Martingale Restriction. *Review of Financial Studies* 8:1091–124.

Macbeth, J. D., and L. J. Merville. 1979. An Empirical Examination of the Black-Scholes Call Option Pricing Model. *Journal of Finance* 34:1173–86.

Manaster, S., and R. J. Rendleman, Jr. 1982. Option Prices as Predictors of Equilibrium Stock Prices. *Journal of Finance* 37:1043–57.

Mayhew, S. 2002. Competition, Market Structure, and Bid-Ask Spreads in Stock Option Markets. *Journal of Finance* 57:931–58.

Mitchell, M., T. Pulvino, and E. Stafford. 2002. Limited Arbitrage in Equity Markets. *Journal of Finance* 57:551–84.

Neal, R. 1987. Potential Competition and Actual Competition in Equity Options. *The Journal of Finance* 42:511–31.

New York Stock Exchange. 1991. Rule 440 C Deliveries Against Short Sales. Information Memo No. 91–41.

Ofek, E., and M. Richardson. 2003. DotCom Mania: The Rise and Fall of Internet Stock Prices. *Journal of Finance* 58:1113–38.

Ofek, E., M. Richardson, and R. Whitelaw. 2004. Limited Arbitrage and Short Sales Restrictions: Evidence from the Options Markets. *Journal of Financial Economics* 74:305–42.

Peña, I., G. Rubio, and G. Serna. 1999. Why Do We Smile? On the Determinants of the Implied Volatility Function. *Journal of Banking & Finance* 23:1151–79.

Reed, A. V. 2007. Costly Short Selling and Stock Price Adjustments to Earnings Announcements. Working Paper, University of North Carolina.

Rubinstein, M. 1985. Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Options Classes from August 23, 1976, through August 31, 1978. *Journal of Finance* 40:455–80.

Securities and Exchange Commission. 1995. Release No. 34–35207; File No. SR-NASD-95–01. *Federal Register* V. 60, no. 10.

Securities and Exchange Commission. 1996. Release No. 34–37773; File No. SR-Amex-96–05. *Federal Register* V. 61, no. 197.

Securities and Exchange Commission. 2000. Payment for Order Flow and Internalization in the Options Markets. Special Study, SEC.

Strong, N., and X. Xu. 1999. Do S&P 500 Index Options Violate the Martingale Restriction? *Journal of Futures Markets* 19:499–521.

Weiss, G. 1998. Were the Short Sellers Ripped Off? *Business Week* 3572:118–19.