Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Financial Economics 74 (2004) 305–342
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/econbase

Limited arbitrage and short sales restrictions: evidence from the options markets ☆

Eli Ofek[a], Matthew Richardson[a,b], Robert F. Whitelaw[a,b,*]

[a] *Stern School of Business, New York University, New York, NY 10012, USA*
[b] *National Bureau of Economic Research, Cambridge, MA 02138, USA*

Received 25 November 2002; accepted 29 May 2003

Available online 13 May 2004

Abstract

We investigate empirically the well-known put–call parity no-arbitrage relation in the presence of short sales restrictions. Violations of put–call parity are asymmetric in the direction of short sales constraints, and their magnitudes are strongly related to the cost and difficulty of short selling. These violations are also related to both the maturity of the option and the level of valuations in the stock market, consistent with a behavioral finance theory of over-optimistic stock investors and market segmentation. Moreover, both the size of put–call parity violations and the cost of short selling are significant predictors of future returns for individual stocks.

JEL classification: G13; G14

Keywords: Limits to arbitrage; Short sales restrictions; Put–call parity; Rebate rate

1. Introduction

The concept of no arbitrage is at the core of our beliefs about finance theory. In particular, two assets with the same payoffs should have the same price. If this

☆ We would like to thank Michael Brandt, Steve Figlewski, Francis Longstaff, Lasse Pedersen, and Eduardo Schwartz for helpful suggestions and David Hait (OptionMetrics) for providing the options data. We are especially grateful to comments from Owen Lamont, Jeff Wurgler, the anonymous referee, and seminar participants at UBC, NYU, USC, Dartmouth, and the NBER.

*Corresponding author. Stern School of Business, New York University, New York, NY 10012, USA.
E-mail address: rwhitela@stern.nyu.edu (R.F. Whitelaw).

0304-405X/$ - see front matter
doi:10.1016/j.jfineco.2003.05.008

restriction is violated, then at least two conditions must be met. First, there must be some limits to arbitrage that prevent the convergence of these two prices (see, e.g., Shleifer, 2000; Barberis and Thaler, 2003). Second, there must be a reason why these assets have diverging prices in the first place. The goal of our paper is to analyze the impact of these two conditions in an obvious no-arbitrage framework.

There is perhaps no better example in finance than the case of redundant assets, for example, stocks and options on these stocks. One of the most commonly cited no-arbitrage relations using stocks and options is that of put–call parity. The put–call parity condition assumes that investors can short the underlying securities. If short sales are not allowed, then this no-arbitrage relation may no longer hold. Of course, even without short sales, the condition does not necessarily fail. Suppose that the stock is priced too high on a relative basis. Then one could form a portfolio by buying a call, writing an equivalent put, and owning a bond; the return on this portfolio would exceed the return on the stock in all possible circumstances. This is a difficult phenomenon to explain in rational equilibrium asset pricing models.

There is a considerable and growing literature that looks at the impact of short sales restrictions on the equity market (see, e.g., Lintner, 1969; Miller, 1977; Harrison and Kreps, 1978; Figlewski, 1981; Jarrow, 1981; Chen et al., 2002; D'Avolio, 2002; Duffie et al., 2002; Geczy et al., 2002; Jones and Lamont, 2002; Mitchell et al., 2002; Ofek and Richardson, 2003; among others). However, there has been much less attention paid to understanding the direct links between short sales and the options market (Figlewski and Webb (1993), Danielson and Sorescu (2001), and Lamont and Thaler (2003) are notable exceptions). Of particular interest to this paper, Lamont and Thaler (2003) document severe violations of put–call parity for a small sample of three stocks that have gone through an equity carve-out, and the parent sells for less than its ownership stake in the carve-out. Lamont and Thaler (2003) view this evidence as consistent with there being high costs to short these stocks.

This paper provides a comprehensive analysis of put–call parity in the context of short sales restrictions. We employ two novel databases from which we construct matched pairs of call and put options across the universe of equities, as well as a direct measure of the shorting costs of each of the underlying stocks, namely their rebate rate. This rebate rate is the interest rate that investors earn on the required cash deposit equal to the proceeds of the short sale. We report several interesting results. First, consistent with the theory of limited arbitrage, we find that the violations of the put–call parity no-arbitrage restriction are asymmetric in the direction of short sales restrictions.[1] These violations persist even after incorporating shorting costs and/or extreme assumptions about transactions costs (i.e., all options transactions take place at ask and bid prices). For example, after shorting costs,

[1] Related phenomena exist in other markets. For example, short-sellers of gold must pay a fee called the lease rate in order to borrow gold. This short-selling cost enters the no-arbitrage relation between forward and spot prices of gold as a convenience yield (see McDonald and Shimko, 1998). Longstaff (1995) examines transaction costs in the market for index options and shows that these costs can increase the implied cost of the index in the options market relative to the spot market.

13.63% of stock prices still exceed the upper bound implied by the options market while only 4.36% are below the lower bound. Moreover, the mean difference between the option-implied stock price and the actual stock price for these violations is 2.71%.

Second, under the assumption that the rebate rate maps one-to-one with the difficulty of shorting, we find a strong general relation between violations of no arbitrage and short sales restrictions. In particular, both the probability and magnitude of the violations can be linked directly to the magnitude of the rebate rate, or, in other words, the degree of specialness of the stock. In a regression context, a one standard deviation decrease in the rebate rate spread implies a 0.67% increase in the deviation between the prices in the stock and options markets. This result is robust to the inclusion of additional variables to control for effects such as liquidity, in either the equity or options markets, stock and option characteristics, and transactions costs.

The above results suggest that the relative prices of similar assets (i.e., ones with identical payoffs) can deviate from each other when arbitrage is not possible. If we take the view that these deviations rule out our most standard asset pricing models, then what possible explanations exist? If markets are sufficiently incomplete, and there is diversity across agents, then it may be the case that these securities offer benefits beyond their risk-return profiles (see, e.g., Detemple and Jorion, 1990; Detemple and Selden, 1991; Detemple and Murthy, 1997; Basak and Croitoru, 2000). Alternatively, if markets are segmented such that the marginal investors across these markets are different, it is possible that prices can differ. Of course, in the absence of some friction that prevents trading in both markets, this segmentation will not be rational.

Third, we provide evidence on this latter explanation by examining a simple framework in which the stock and options markets are segmented and the equity markets are "less rational" than the options markets. This framework allows us to interpret the difference between a stock's market value and its value implied by the options market as mispricing in the equity market. It also generates predictions about the relation between put–call parity violations, short sales constraints, maturity, valuation levels, and future stock returns. Consistent with the theory, we find that put–call parity violations are increasing in both the maturity of the options and the potential level of mispricing of the stocks. We also evaluate the model's ability to forecast future movements in stock returns. Filtering on rebate rate spreads and put–call parity violations yields average returns on the stock over the life of the option that are as low as −12.6%. In addition, cumulative abnormal returns, net of borrowing costs, on portfolios that are long the industry and short stocks chosen using similar filters are as high as 65% over our sample period.

This paper is organized as follows. In Section 2, we review the basics of put–call parity and the lending market, and then describe the characteristics of the data used in the study. Section 3 presents the main empirical results on the violations of put–call parity and their link to short sales restrictions. In Section 4, we apply our analysis to imputing the overvaluation of stocks using evidence from the options market. Section 5 makes some concluding remarks.

2. Preliminaries

2.1. Put–call parity

Under the condition of no arbitrage, it is well known that for European options on nondividend paying stocks, put–call parity holds, i.e.,

$$S = \mathrm{PV}(K) + C - P, \tag{1}$$

where S is the stock price, $\mathrm{PV}(K)$ is the present value of the strike price, and C and P are the call and put prices, respectively, on options with strike price K and the same maturity. For American options, Merton (1973) shows that the puts will be more valuable because at every point in time there is a positive probability of early exercise. That is,

$$S \geqslant \mathrm{PV}(K) + C - P. \tag{2}$$

There are essentially two strands of the literature that investigate Eq. (2) above. The first group of papers contains a series of empirical investigations (e.g., Gould and Galai, 1974; Klemkosky and Resnick, 1979; Bodurtha and Courtadon, 1986; Nisbet, 1992; Kamara and Miller, 1995; Lamont and Thaler, 2003). The evidence from this literature is mixed, but for the most part finds that put–call parity holds as described by Eq. (2).

For example, Klemkosky and Resnick (1979) use a sample of 15 stocks during the first year of put trading on the CBOE and show that the evidence is generally consistent with put–call parity and market efficiency. They present some evidence of asymmetry in violations consistent with that reported in this paper, with approximately 55% of the violations consistent with short sales restrictions and violations of larger magnitudes in this direction. However, they do not estimate or directly control for the early exercise premium, and they acknowledge that this omission may be responsible for the estimated violations in this direction. Consequently, Klemkosky and Resnick do not try to explain these findings and tend to focus on the violations in the opposite direction. Nisbet (1992) examines a somewhat larger sample of options on 55 companies traded on the London Traded Options Market for a six-month period in 1988. She also adds direct estimates of transactions costs to the analysis and generally finds that apparent violations cannot be exploited. Again, ignoring the early exercise premium, the results suggest larger and more numerous violations in the direction documented in this paper. However, this asymmetry is reduced or disappears when she eliminates observations for which early exercise is likely. Interestingly, Nisbet also speculates that short sales restrictions might account for the existence of put–call parity violations, but she does not pursue this topic. Later studies, such as Kamara and Miller (1995), tend to focus on index options as these options are liquid and of the European type. They find fewer instances of violations than previous studies, though the studies are not directly comparable given the underlying are indexes rather than individual equities. The paper closest in spirit to ours is that of Lamont and Thaler (2003), which documents large violations of put–call parity for a sample of three stocks that:

(i) have gone through an equity carve-out, and (ii) the parent sells for less than its ownership stake in the carve-out. The analysis in this paper looks at a much wider universe of stocks and their underlying options.

The second strand of the literature is concerned with analytical valuation formulas for American put options in which explicit values are given for the early exercise premium (e.g., Johnson, 1983; Geske and Johnson, 1984; Ho et al., 1994; Unni and Yadav, 1999). Specifically, Eq. (2) can be rewritten as

$$S = \mathrm{PV}(K) + C - P + \mathrm{EEP}, \tag{3}$$

where EEP is the early exercise premium on the American put option.

At least two conditions must be met for Eq. (3) to fail. First, although it is no longer strictly an arbitrage relation as the value of the early exercise premium is incorporated directly, there must be some limits on arbitrage to permit significant violations of this relation. The most commonly cited limit is short sales restrictions. Without short sales, if the stock price drifts above its implied price in the options market, then there does not exist an arbitrage that will automatically lead to convergence of the two values. There is a large and growing literature in finance that documents both the theoretical and empirical importance of short sales restrictions.[2]

Second, it must be possible that the values given by Eq. (3) can drift apart. That is, why would an investor purchase shares for $S when she could duplicate the payoff of the stock using the bond market and call–put option pairs? Perhaps, it is too difficult or costly to replicate shares in the options market (e.g., transactions costs), or there is some hidden value in owning shares (e.g., Duffie et al., 2002). Alternatively, perhaps options provide some additional value in terms of risk management due to markets being incomplete (e.g., Detemple and Selden, 1991).

The most popular explanation lies at the roots of behavioral finance. Behavioral finance argues that prices can deviate from fundamental values because a significant part of the investor class is irrational. These irrational investors look to other information, e.g., market sentiment, or are driven by psychological (rather than financial) motivations. This class of investors has the potential to move asset prices, and, in the presence of limited arbitrage, there is no immediate mechanism for correcting these resulting mispricings (see, e.g., Shleifer, 2000). In the context of Eq. (3), if the equity and options markets are segmented, i.e., have different investors, then mispricings in the equity market do not necessarily carry through to the options market (see Lamont and Thaler, 2003). In other words, irrational investors do not use the options market.

In particular, as long as the investors in options are different than those in the equity market, and as long as these options investors believe there is a positive probability that asset prices will revert back to their fundamental price by the time the options expire, there can be a substantial difference between the market asset price and the implied asset price from the options market. Of course, these

[2] For example, Lintner (1969), Miller (1977), Jarrow (1981), Figlewski (1981), Chen et al. (2002), Hong and Stein (2002), D'Avolio (2002), Geczy et al. (2002), Ofek and Richardson (2003), Jones and Lamont (2002), and Duffie et al. (2001) to name a few.

differences can only persist in the presence of limited arbitrage, whether that is due to transactions costs or, more directly, short sales restrictions. An interesting feature generated by the fixed expiration of the option is that in a world of mean reversion to fundamental values, the maturity of the option can be an important determinant of the level of violations of Eq. (3).

In this paper, we investigate violations of Eq. (3) and relate them to the conditions described above, namely: (i) limited arbitrage via either short sales restrictions or transactions costs, and (ii) potential periods of mispricing between equities and their corresponding options. We evaluate this latter condition by looking at expected maturity effects, potential structural shifts in mispricing, and the forecastability of future returns.

2.2. The lending market

There has been recent interest in the lending market for stocks. For example, D'Avolio (2002) and Geczy et al. (2002) provide a detailed description and analysis of this market. Beyond the papers described in footnote 2 that show the potential theoretical effects of short sales restrictions and that document empirical facts strongly relating short sales restrictions to stock prices, D'Avolio (2002) and Geczy et al. (2002) present evidence that short sales restrictions exist and are not uncommon.

There are essentially two reasons why short sales restrictions exist. Investors are either unwilling to sell stock short or find it too difficult to do so. In the former case, Chen et al. (2002) provide a detailed account of why investors may be unwilling to short stock. In particular, they focus on an important group of investors, i.e., mutual funds, and argue that though restrictions under the Investment Company Act of 1940 are no longer binding, mutual funds still abide by that act. In fact, Almazan et al. (2002) show that only a small fraction of mutual funds short stocks, and they provide evidence of greater mispricings when mutual funds are absent from the market.

In the latter case, there are both theoretical reasons and supporting empirical evidence that suggests it is difficult to short stocks on a large scale. First, in order to short a stock, the investor must be able to borrow it. In general, there are only a limited number of shares available for trading (i.e., a stock's float is finite), and someone (i.e., an institution or individual) would have to be willing to lend the shares. For example, insiders may be reluctant to sell or be prevented from selling, and, in the extreme case, for six months after an IPO, most of the shares have lockup restrictions. For whatever reason, individuals tend to lend shares less than institutions do. Second, there is no guarantee that the short position will not get called through either the lender demanding that the stock be returned or a margin call. In this case, there is no guarantee that the investor will be able to re-short the stock.

When an investor shorts a stock, she places a cash deposit equal to the proceeds of the shorted stock. That deposit carries an interest rate referred to as the rebate rate. If shorting is easy, the rebate rate closely reflects the prevailing market rate.

However, when supply is tight, the rebate rate tends to be lower. This lower rate reflects compensation to the lender of the stock at the expense of the borrower, and thus can provide a mechanism for evening out demand and supply in the market. One way to measure the difficulty in short selling is to compare the rebate rate on a stock against the corresponding "cold" rate, i.e., the standard rebate rate on stocks that day. Since there is limited demand for short selling the majority of stocks, empirically this cold rate corresponds to the median rebate rate.

There are two ways in which we view the rebate rate spread in this paper. First, it can be used as the actual cost of borrowing a stock, and thus the rate can be employed in Eq. (3) in that context (e.g., D'Avolio, 2002; Mitchell et al., 2002). Second, as pointed out by Geczy et al. (2002) and Ofek and Richardson (2003), the lending market is not a typical well-functioning, competitive market. Thus, it may not be appropriate to treat rebate rates as competitive lending rates, and, instead, we use the rebate rate as a signal of the difficulty of shorting, i.e., the degree to which short sales restrictions are binding.

Alternatively, if investors can short only a limited number of shares, there are other ways to bet against the stock. For example, one could imagine setting up a synthetic short position using the options market. Figlewski and Webb (1993) and Lamont and Thaler (2003) look at this case empirically. In the context of our discussion in Section 2.1, we might expect to see violations of put–call parity as the standard no-arbitrage condition can be violated due to short sales restrictions and overvaluation of stocks. In this case, there would be excess demand for put options relative to call options, leaving a significant spread between the prices. As an extreme example, Lamont and Thaler (2003) show that in the Palm/3Com case, the synthetic short for Palm (i.e., its implied value from options) was substantially lower than the traded price of Palm (approximately 30% lower during the first few weeks). This is consistent with the equity prices reflecting one set of beliefs and the options market reflecting another.

2.3. Data

This paper looks at put–call parity in the options market in conjunction with short sales restrictions as measured by the rebate rate. We employ two unique data sets over the sample period July 1999 to November 2001. Specifically, we look at daily data for 118 separate dates during this period that are approximately five business days apart.

The first dataset comes from OptionMetrics, who provide end-of-day bid and ask quotes, open interest, and volume on every call and put option on an individual stock traded on a U.S. exchange (often more than three million option observations per month). Along with the options data are the corresponding stock prices, dividends, and splits, as well as option-specific data such as implied volatilities, interest rates, maturities, and exercise prices (see the Appendix for details).

The second dataset includes the rebate rate for almost every stock in our options sample. In particular, a financial institution, and one of the largest dealer-brokers, provided us with its proprietary rebate rates for the universe of stocks on the aforementioned dates. The rebate rate quoted represents an overnight rate and thus includes no term contracts, which are also possible in the lending market. The existence of a rebate rate quote is not an implicit guarantee that the financial institution will be able to locate shares of the stock for borrowing. It is simply the rate that will apply if the stock can be located. Moreover, the rebate rate quote may not be the same as that quoted by another institution, although these rates are likely to be highly correlated. For each day, we calculate the short selling cost as the deviation of the rebate rate on a particular stock from the cold rate for the day, i.e., the standard rebate rate on the majority of stocks. We denote this cost as the rebate rate spread throughout the paper. Obviously, this spread will be zero for the majority of firms.

There is one potentially important measurement issue with respect to the rebate rates. It appears that not all the quotes are synchronous. Therefore, if interest rates and the cold rate move during the day, stale rebate rates may appear to deviate from the cold rate even though they did not do so at the time of the original quote. This phenomenon is most obvious in small positive rebate rate spreads, which we set to zero. When the rebate rate spread is small and negative, there is no obvious way to determine if it is truly negative or if it is the result of nonsynchronicity. As a result, we do not adjust these spreads, and there is likely to be some measurement error in rebate rate spreads, especially at low absolute magnitudes.

Table 1A describes our entire sample of option pairs, i.e., puts and calls with the same exercise price and maturity, after we apply a set of preliminary filters. These filters are described in detail in the appendix, but the primary requirements are that the stock be nondividend paying and that both the put and call have positive open interest. Over the sample period, this sample includes a total of 1,359,461 option pairs. These pairs span 118 dates, with approximately 1100 firms per date and ten option pairs per firm (an average of 2.5 different maturities and 4.3 different strike prices per maturity). The median and mean maturity of the options pairs are 115 and 162 days, respectively. The open interest on the call options tends to be larger than on the put options, with the mean and medians being 711 and 133 contracts versus 481 and 63, respectively. Note, however, that the daily volume can be quite low, especially for the put options. In particular, the mean and median volume for the call and puts are 32 and zero versus 16 and zero, respectively. Of course, even though over half the sample of options on any day does not trade, this does not mean that the bid and ask quotes do not represent accurate prices at which the options can be bought and sold. As a robustness check, we duplicate the analysis that follows using only options that had positive trading volume. While the sample sizes are much smaller, the results are qualitatively the same.

For the analysis, we further wish to restrict our sample to homogenous sets of option pairs. Therefore, we break the sample up into three maturity groups: (i) short (i.e., 30–90 days), (ii) intermediate (i.e., 91–182 days), and (iii) long (i.e., 183–365 days). Furthermore, we focus on options that are close to at-the-money (i.e., $-0.1 < \ln(S/K) < 0.1$) and apply a second set of filters to eliminate bad data (see the appendix). The majority of the analysis looks at the at-the-money, intermediate maturity option pairs. If there are multiple option pairs per stock on a given day that

Table 1
Sample description
Panel A reports descriptive statistics for the full sample of paired options. The data span 118 dates between July 1999 and November 2001. The total number of option pairs is 1,359,461. Panel B reports descriptive statistics for the subsample of paired options with $\ln(S/X)$ of less than 10% in absolute value, and maturity between 91 and 182 days. If multiple options pairs fit the criteria for a single firm on a given date, then only one pair is selected. The total number of pairs in Panel B is 80,614, of which 24,542 have negative rebate rate spreads (Reb$<$0).

Variable	Mean	Median	5th pctl	95th pctl
Panel A: full sample of paired options				
Days to expiration	161.918	115.000	37.000	569.000
Ln(S/K)(%)	−2.361	−1.859	−55.513	50.456
Open interest—call	711.4	133	5	2655
Open interest—put	480.6	63	3	1661
Daily volume—call	31.9	0	0	103
Daily volume—put	15.5	0	0	40
Number of firms per date	1083.7	1104	963	1160
Number of option expirations per firm	2.5	2	1	5
Number of strikes per expiration	4.3	3	1	12
Panel B: at-the-money, intermediate maturity sample of paired options				
Stock price	32.195	23.813	7.520	83.375
Expiration (days)	134.554	135.000	95.000	177.000
Ln(S/K)(%)	0.047	0.000	−7.796	7.855
Open interest—call	416.510	101	5	1525
Open interest—put	289.110	50	3	1056
Daily volume—call	20.163	0	0	70
Daily volume—put	12.129	0	0	30
Spread—call (% of mid)	8.580	7.407	2.128	18.182
Spread—call (% of stock price)	1.526	1.311	0.360	3.428
Spread—put (% of mid)	9.176	8.000	2.247	20.000
Spread—put (% of stock price)	1.474	1.254	0.344	3.338
EEP (% of put mid)	0.829	0.709	0.181	1.815
EEP (% of stock price)	0.132	0.117	0.026	0.282
Implied volatility call (%)	74.751	72.813	39.219	118.125
Implied volatility put (%)	73.508	74	40	120
Rebate rate spread (Reb$<$0)	−1.573	0	−6	0
Number of firms per date	683.169	693	561	781
Number of obs. per firm	46.490	38	3	113
Number of options per firm	17.385	13	1	49

match the relevant maturity and moneyness criteria, then we restrict ourselves to the option pairs that are closest to the middle of the range. This provides us with a maximum of one option pair per stock per date.

Table 1B provides a summary of the data for the at-the-money, intermediate maturity option pairs. The sample contains 80,614 pairs of options with median and mean expirations of slightly over 130 days. These observations span 1,734 different stocks, with an average of 683 stocks per date. Compared to the larger sample, the open interest and volume for the calls and puts are of a similar magnitude. Of some interest to the analysis of put–call parity with transactions costs, the mean and median values of the bid–ask spread on calls and puts range from 7.4% to 9.2% as a percentage of the midpoint of the corresponding option quotes. Thus, in the extreme case in which transactions only take place at ask and bid prices, these costs may be especially relevant. Of course, these spreads are much smaller as a percentage of the stock price, with means and medians ranging from 1.3% to 1.5%, but transaction costs are still likely to be substantially higher in the options market than in the stock market.

Table 1B illustrates three other important features of the data. First, the implied volatilities of the stocks are quite high by historical standards, that is, almost 75% on average. Note that these implied volatilities are calculated using the Black-Scholes pricing model for call options assuming no dividends. Second, the early exercise premium for puts is relatively low, representing less than 1% of the value of the option on average and only slightly more than 0.1% of the stock price. We use the method of Ho et al. (1994) to estimate this premium for each put option on each date. All the put–call parity conditions are then adjusted for this estimate as in Eq. (3). Finally, the mean and median annualized rebate rate spreads, conditional on being "special" (i.e., the rebate rate spread being negative), are −1.57% and −0.46%, respectively. The interpretation of these values in terms of both the actual costs of shorting and, more generally, as an indicator of the difficulty of shorting, are discussed in detail in the next section. Note that 24,542 (approximately 30%) of the observations correspond to negative rebate rate spreads, although interpreting all these stocks as special is almost certainly incorrect given the issue of nonsynchronous rebate rate quotes discussed above.

3. Put–call parity: empirical tests

In this section, we perform an initial empirical analysis of Eq. (3). Ceteris paribus, without any underlying theory we might expect 50% of the violations of Eq. (3) to be on either side. However, the limited arbitrage via short sales restrictions provides an asymmetry to Eq. (3). In particular, as stocks' market values rise above those implied by the options markets (if, in fact, that were to occur), there is no arbitrage mechanism that forces convergence. On the other hand, if stock prices fall below their implied value, one can arbitrage by buying shares and taking the appropriate option positions. Thus, to the extent short sales constraints are binding, if prices deviate from fundamental value, Eq. (3) will be violated in one particular direction.

We provide three formal examinations related to Eq. (3). First, using the midpoints of the option quotes and the closing price of the stock, we evaluate violations of Eq. (3). In addition, we directly relate these violations to the spread between the rebate rate and the prevailing market rate. To preview the major results, there are violations of put–call parity primarily in the direction of the asymmetry induced by binding short sales constraints.

Second, to better understand this latter point, we investigate the relation between the rebate rate spread and both the magnitude and direction of these violations. As a test of robustness, we include a number of other control variables, such as ones related to liquidity in both the options and equity market, to the underlying characteristics of the options, and to valuation levels in the equity market. While some of these variables do have explanatory power, it tends to be small relative to that of the rebate rate spread. More important, the rebate rate spread results are robust to the inclusion of all these variables.

Third, the initial analysis assumes transactions take place at the midpoint of the quoted spread. As an alternative, we assume that all purchases and sales in the options market are done at the ask and bid prices, respectively. We also build into the analysis the assumption that the investor can short, but at the cost of the rebate rate spread. This provides us with a more stringent test of the put–call parity condition. We still document important violations though they are significantly reduced in number. We view these violations as evidence that the rebate rate measures more than just the direct cost of shorting. While these transaction costs-based results cannot explain why stock prices and their option-implied values drift apart, it does explain why investors do not exploit these differences.

3.1. Put–call parity violations

We investigate Eq. (3) by taking the midpoint prices of all the option pairs in our filtered sample, the corresponding stock price, and the prevailing market interest rate (see the appendix for details about this interest rate). Table 2 reports both the percentage of violations of put–call parity in both directions, as well as estimates of the cross-sectional distribution of the traded stock price value divided by the option-implied stock price value. That is, in the latter case, we look at the ratio $R = 100\ln(S/S^*)$ where $S^* = \mathrm{PV}(K) + C - P + \mathrm{EEP}$. To the extent that there are asymmetric violations due to short sales constraints, we would expect R to exceed zero.

There are several interesting observations one can make from the results reported in Table 2. First, in the sample period studied here, R exceeds zero for almost two-thirds of the sample. As mentioned previously, ceteris paribus, we would expect this number to be 50%. In fact, it is possible to show that under the null that the true probability is 50%, the 5% tail is approximately 50.70%; thus, the actual percentage of 65.10% is statistically significant at any measurable level.

In calculating the 5% tail above, it is critical to adjust for dependence across the observations. Empirically, there is a negligible cross-sectional correlation between observations for different firms, even contemporaneously; therefore, we only control for serial dependence. It is impossible to estimate the autocorrelations separately for each firm because the data are sparse—on average, each firm only has observations for 46 of the 118 dates (see Table 1B). Consequently, we impose the restriction that the autocorrelation function is the same for every stock. For the full sample, the stock price ratio R has a first-order autocorrelation of 0.60, and autocorrelations decline slowly for longer lags. Not surprisingly, the binomial variable that measures whether R exceeds zero has a much lower first-order autocorrelation of 0.28. Nevertheless, the variance of the estimate of the percentage of positive ratios (i.e., the

Table 2
Distribution of unadjusted stock price ratios
The table reports the distribution of the ratio $R \equiv 100\ln(S/S^*)$ for at-the-money, intermediate maturity options, where S is the stock price and S^* is the stock price derived from the options market using put–call parity and assuming trades of options at the midpoint of the spread. The four test statistics and corresponding P-values test: (1) the equality of the mean ratios across zero (Reb = 0) and negative rebate spread (Reb < 0) stocks, (2) and (3) whether the probability of observing $R>0$ equals 50% for the zero and negative rebate spread stocks, respectively, and (4) whether the probability of observing $R>0$ is equal across zero and negative rebate spread stocks. The test statistics have an asymptotic N(0, 1) distribution under the null hypotheses.

	All	Reb = 0	Reb < 0	Reb < −1
Obs.	80,614	56,072	24,542	8699
Mean	0.30	0.16	0.61	1.21
Percentiles				
1	−2.93	−2.87	−3.04	−3.41
5	−1.22	−1.19	−1.27	−1.37
10	−0.68	−0.67	−0.69	−0.68
25	−0.16	−0.18	−0.12	0.04
50	0.20	0.16	0.35	0.80
75	0.65	0.53	1.02	1.82
90	1.33	1.04	2.04	3.34
95	1.97	1.49	2.97	5.14
99	4.42	2.82	7.68	10.16
$R<0$ (%)	34.90	36.83	30.50	23.80
$R>0$ (%)	65.10	63.17	69.50	76.20

Test	Stat.	P-value
$E[R \mid Reb=0] = E[R \mid Reb<0]$	9.08	0.00
$Pr(R>0 \mid Reb=0) = 50\%$	28.92	0.00
$Pr(R>0 \mid Reb<0) = 50\%$	25.92	0.00
$Pr(R>0 \mid Reb=0) = Pr(R>0 \mid Reb<0)$	7.19	0.00

average of the binomial variable) is more than six times larger than under the assumption of independence. The 5% tail would be 50.28% based on an assumption of independence. The overall effect of the serial dependence is to make similar upward adjustments to the standard errors and downward adjustments to the test statistics that are reported in the tables and discussed later in the paper. Of course, the standard errors themselves are estimates and depend on the estimated autocorrelations. However, for all the major results the p-values are so small that the statistical significance is not in doubt. An alternative way to adjust for the serial dependence in each stock's put–call parity violation is to restrict ourselves to one observation per firm. Specifically, we select the observation for each firm with a rebate rate spread closest to the median value for that firm. The disadvantage of this approach is that it throws away data and we lose the time series structure of our analysis. The advantage is that it requires fewer assumptions on the underlying distribution of the data. (We thank the referee for this suggestion.) We re-run the results of Tables 2, 3, and 5 using this sample of 1,734 observations. The results are similar in spirit to the ones documented in the text, and, if anything, are a little more dramatic. We conjecture that this latter effect might be due to a reduction in the noise in the data by eliminating extreme rebate rate spreads. Finally, we also adjust the standard errors for heteroscedasticity where appropriate, again assuming that the form of heteroscedasticity is the same across all firms.

Second, consistent with this asymmetry, the median and mean of R are 0.30 and 0.20, respectively. While these estimates are significant at conventional levels, the magnitudes do not seem particularly large. Moreover, in studying the cross-sectional distribution of R, the 1% and 99% tails are—2.93 and 4.42, respectively. The tails of R are asymmetric but not markedly so, further suggesting that while violations occur, they tend to be relatively small. Note that these observations look at the sample unconditionally. As discussed in Section 2.1, deviations from fundamental value are not sufficient to generate violations of put–call parity. At a minimum, there must also be some form of limited arbitrage. Therefore, we break the sample into two distinct groups—one with rebate rate spreads equal to zero, and the other with negative rebate rate spreads. If negative spreads map one-to-one with short sales restrictions, then this partition represents one way to condition on stocks that are subject to limited arbitrage.

Table 2 reports the results using the rebate rate partitioning of the data. First, note that of the 80,614 option pairs, 24,542, or approximately 30% of the observations, have negative rebate rate spreads. However, as described in Section 2.3, there is reason to believe that rebate rate spreads are subject to some measurement error, suggesting that observations of small negative rebate rate spreads may not be that informative. It is difficult to determine whether small negative spreads are simply a result of nonsynchronous observations of the rebate rate across stocks or whether they measure a true short selling cost. Thus, we also condition on more significant negative spreads of -1% or greater. This reduces the number of observations to 8,699, or still 10.8% of the sample. It seems that difficulty in shorting stocks is a relatively common phenomenon in our sample.

Second, the option pairs with negative rebate rate spreads also have a greater percentage of put–call parity violations in the expected direction, that is, 69.50% versus 63.17%. These differences are significant at any measurable level with a standard normal test statistic equal to 7.19. All the statistical tests of positive violation probabilities in the paper use the well-known DeMoivre-Laplace normal approximation to the binomial distribution, adjusted for serial dependence in the data as described previously. Given our sample sizes, this asymptotic approximation is essentially perfect. Interestingly, the occurrence of these violations and the underlying ratios are also more persistent for the negative rebate rate stocks. To the extent that rebate rate spreads are persistent—a conjecture that we verify later—this evidence is consistent with short sales constraints being meaningful.

Third, the median and mean of the ratio R are significantly greater for these negative rebate rate stocks, i.e., 0.35 and 0.61 versus 0.16 and 0.16, respectively.

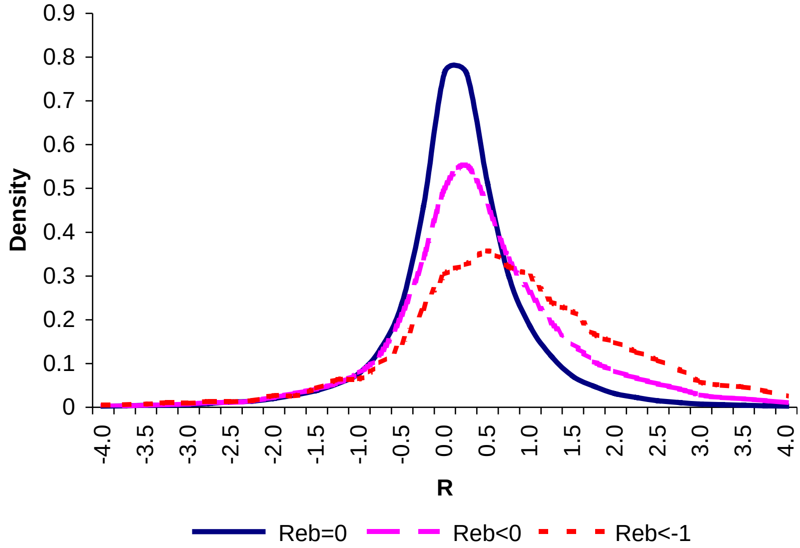

Fig. 1. Distribution of unadjusted stock price ratios. The figure shows the empirical distributions of the ratio $R \equiv 100\ln(S/S^*)$ for at-the-money, intermediate maturity options, where S is the stock price and S^* is the stock price derived from the options market using put–call parity and assuming trades of options at the midpoint of the spread. Three samples are constructed based on the magnitude of the rebate rate spread: (i) a rebate rate spread equal to zero (Reb $=0$); (ii) a negative rebate rate spread (Reb <0); and, (iii) a rebate rate spread less than -1% (Reb <-1).

Fourth, and most important, while the 1% tails of the distribution of R are similar for the two samples (i.e., -3.04 versus -2.87), the 99% tails are dramatically different (i.e., 7.68 versus 2.82). Fig. 1 graphically illustrates this point in a slightly more general manner via a plot of the empirical distributions of put–call parity violations (i.e., R). The left tails of the distributions are almost identical, but, for high stock price ratios, the density for stocks with negative rebate rate spreads is many times greater than that for stocks with zero rebate rate spreads. The theory suggests that the distribution of R should be asymmetric as the limited arbitrage makes itself manifest through the difficulty in shorting stocks, i.e., when $S > S^*$. For negative rebate rate spread stocks this asymmetry is clear in Fig. 1, especially in contrast to the symmetry apparent in the distribution for zero rebate rate spread stocks.

Finally, these results are substantially more dramatic when we condition on spreads less than -1%, with the mean of R doubling to 1.21%, the 99% tail increasing to over 10%, and the proportion of positive violations exceeding 76%. Again, however, the left tail of the distribution is almost identical, with the effect of short sales constraints evident in the further increase in skewness. These results are consistent with the measurement error hypothesis, but they also suggest a relation between the magnitude of the spread and violations of put–call parity. We explore this issue below.

3.2. The rebate rate and put–call parity violations

Table 3A reports regression results of R on the rebate rate spread using the full sample, as well as for the observations with negative rebate rate spreads. There is a strong negative relation between the rebate rate spread and R. While this is expected given the previous results, Table 3A allows us to quantify both the statistical and economic significance of this relation. The t-statistic is over eight, which represents significance at any imaginable level. Conditional on a negative rebate rate spread, a one standard deviation decrease in the rebate rate (i.e., 2.77%) leads to a 0.67% increase in the relative mispricing between the stock price and its implied value from options.

In the context of the above regression, one way to address the issue of whether the rebate rate measures the actual cost of shorting versus the difficulty of shorting would be to regress R on the rebate rate spread for all the observations, but include a dummy variable for whether the rebate rate spread is zero. If the rebate rate proxies for the difficulty of shorting, then we would expect to see a discontinuity at zero. In other words, a very small but negative rebate spread should have different implications than a zero rebate rate spread. As expected, the coefficient on the rebate rate is the same; however, the dummy variable is statistically significant, albeit small at -0.07%. Thus, there is only a small jump in the magnitude of the violation once the rebate rate goes negative. Note that both the magnitude and statistical significance may be reduced by the presence of measurement error in small negative rebate rate spreads, as discussed earlier.

The empirical fact that the rebate rate spread is strongly related to the magnitude of the put–call parity deviation is consistent with the theory of limited arbitrage. However, there are other potential explanations. For example, perhaps the put–call parity deviation reflects the underlying liquidity in the market, and the rebate rate spread simply proxies for this liquidity (or lack thereof). To test this hypothesis, Table 3B reports regressions of R on the rebate rate spread, on proxies for liquidity in both the options and equity market (i.e., open interest, option spreads, option volume, and equity volume), on underlying characteristics of the options (i.e., implied volatility, moneyness, and maturity) and on a proxy for potential mispricing of the underlying stock (i.e., the earnings-price ratio). This latter variable is truncated at a value of -1 to prevent outliers with large negative earnings from distorting inference. All the regressions are estimated using the sample of observations for which we have data on all the variables in order to assure comparability across the regressions. Thus, the sample size is somewhat smaller than in Table 3A, but it is still substantial. The standard deviations of the independent variables over the full sample are reported in the final column to assist in determining the economic significance of the results. Several observations are of interest.

First, the evidence for rebate rates is robust to the addition of controls in the regression. In fact, the coefficient on the rebate rate spread actually increases slightly (from -0.20 to -0.21 for the full sample and from -0.20 to -0.22 for the negative rebate rate sample), and the statistical significance is of similar magnitude. If we drop

Table 3

Regressions for unadjusted stock price ratios

Panel A reports linear regressions of the stock price ratio on rebate rate spreads. The dependent variable is the ratio $R \equiv 100\ln(S/S^*)$ (see Table 2). The independent variables are a zero rebate spread dummy that equals 1 if the firm has a zero rebate spread that day and 0 otherwise, the rebate spread for the firm that day (Reb), and the adjusted rebate spread for the firm that day (Reb^A), which is the average expected rebate rate spread over the life of the option using the 3-state AR(1) model estimated in Table 4. Panel B reports multivariate regressions of the stock price ratio on the rebate spread dummy, the rebate spread and 9 additional variables: (1) the percentage bid–ask spread averaged across the call and put, (2) the daily volume averaged across the call and put (divided by 100), (3) the open interest averaged across the call and put (divided by 1000), (4) the implied volatility of the call option, (5) the natural log of the average daily dollar volume on the stock over the prior 3 months (divided by the mean across all dates and stocks), (6) the ratio of open interest on the put to open interest on the call (divided by 10), (7) the moneyness of the options ($100\ln(S/K)$), (8) the expiration of the option in years, and (9) the earnings price ratio of the stock. The last column reports the standard deviation of these variables and the dependent variable. Standard errors are in parentheses.

Panel A: rebate rate spread

Sample	Const.	Dummy	Reb	Reb^A	R^2	Obs.
Reb<0	0.228[a]		−0.241[a]		0.108	24,542
	(0.044)		(0.030)			
Reb<0	−0.278[a]			−2.746[a]	0.097	24,542
	(0.070)			(0.258)		
All	0.228[a]	−0.068[c]	−0.241[a]		0.074	80,614
	(0.038)	(0.040)	(0.026)			
All	−0.023[a]			−2.250[a]	0.065	80,614
	(0.021)			(0.153)		

Panel B: regressions with control variables

Sample	All	All	All	Reb<0	Reb<0	Reb<0	STD
Dependent variable	Unadjusted stock price ratio						1.400
Constant	0.218[a]	1.730[a]	1.323[a]	0.218[a]	0.862	−0.013	
	(0.036)	(0.248)	(0.218)	(0.040)	(0.557)	(0.477)	
Rebate dummy	−0.071[c]		−0.214[a]				0.452
	(0.038)		(0.036)				
Rebate spread	−0.197[a]		−0.211[a]	−0.197[a]		−0.220[a]	1.604
	(0.026)		(0.026)	(0.030)		(0.030)	
Option spread		−0.032[a]	−0.027[a]		−0.024[a]	−0.027[a]	5.238
		(0.004)	(0.004)		(0.009)	(0.009)	
Option volume		0.005	0.004		0.056	0.049	1.192
		(0.004)	(0.004)		(0.042)	(0.031)	
Open interest		−0.023[a]	−0.027[a]		0.074	0.003	1.494
		(0.005)	(0.004)		(0.087)	(0.052)	
Implied volatility		−0.916[a]	−1.255[a]		−1.358[a]	−1.664[a]	0.244
		(0.072)	(0.062)		(0.151)	(0.136)	
Stock volume		−0.471[b]	0.218		0.831	1.705[a]	0.104
		(0.211)	(0.178)		(0.513)	(0.428)	
Open interest ratio		0.005[b]	0.004[b]		0.001	−0.001	2.343
		(0.002)	(0.002)		(0.006)	(0.007)	
$\mathrm{Ln}(S/K)$ (%)		−0.004[a]	−0.003[a]		−0.006[b]	−0.006[b]	4.669
		(0.001)	(0.001)		(0.003)	(0.003)	

Table 3 (*continued*)

Sample	All	All	All	Reb < 0	Reb < 0	Reb < 0	STD
Expiration (years)		−0.096	−0.119		0.448[b]	0.430[b]	0.071
		(0.091)	(0.089)		(0.230)	(0.217)	
E/P		−0.665[a]	−0.510[a]		−0.516[c]	−0.625[a]	0.125
		(0.152)	(0.130)		(0.279)	(0.238)	
R^2	0.056	0.028	0.099	0.092	0.038	0.150	
Obs.	65,005	65,005	65,005	18,541	18,541	18,541	

[a] Significant at the 1% level.
[b] Significant at the 5% level.
[c] Significant at the 10% level.

the rebate rate spread from the regressions, then the R^2 drops (from 9.9% to 2.8% for the full sample and from 15.0% to 3.8% for the negative rebate rate sample), which suggests the rebate rate spread is by far the most important factor for explaining put–call parity deviations.

Second, to the extent the option liquidity variables are statistically significant, their coefficients actually go in the opposite direction than one might theorize. That is, the greater the liquidity in the options market (as measured by the spread and open interest), the greater the stock price ratio R. We take this as further evidence that the violations are real and not a product of measurement error. The liquidity in the options market is consistent with investors increasing their trading in this market as asset prices drift further from their fundamentals (subject to the difficulty of shorting). Interestingly, R also increases with the volume in the stock market, which is consistent with these asymmetric put–call parity violations generating trade in the stock market as well. An alternative explanation is that it is the stocks that are heavily traded, especially by retail investors, that tend to exhibit mispricing in the first place.

Third, higher (implied) volatility stocks tend to have lower put–call parity deviations in the direction of interest. It is unlikely that this effect is related to our measure of early exercise premiums because although low volatility tends to reduce the value of holding the option, early exercise is important only for options that are in-the-money. Alternatively, volatility might proxy for some characteristic that helps explain put–call parity violations in the context of short sales restrictions.

Finally, the earnings-price ratio has a negative and significant coefficient for both samples. Again this result is consistent with the story, developed in more detail later, that high stock price ratios are a product of overpriced stocks.

In the regression analysis we control for the time to expiration of the option, which enters with a positive and significant coefficient for negative rebate rate stocks. However, theoretically the more appropriate variable is the predicted magnitude of the rebate rate over the option's life. Given an estimate of the rebate rate spread, we can estimate the relation between the magnitude and direction of the put–call parity violations and expected shorting costs over the life of the option. This variable is also useful for controlling directly for expected shorting costs, as in the transactions costs analysis in the next section, and as a measure of the potential revenues that an owner of the stock can receive by lending it out, as discussed below. Finally, the properties of the rebate rate spread itself are of interest since implicit in our analysis is the assumption that "specialness" is persistent, i.e., that if a stock is costly or difficult to short sell today, it will also be expected to have this same characteristic in the future.

To estimate expected rebate rate costs, we need to develop a rebate rate model. For example, one might expect specialness to subside or get worse over time depending on the current rebate rate spread. Alternatively, even if a stock is not special today, there may be some expectation that it will be in the future. In theory, this expectation of future limits on arbitrage could drive a wedge between the equity and options markets.

Our model assumes that rebate rate spreads follow a three-state Markov model, where the states are defined as rebate rate spreads of zero, between zero and -0.5%, and less than -0.5%. The transition probabilities between these states are estimated from the data. Conditional on negative rebate rate spreads and remaining in the current state, we assume an autoregressive time series model (an AR(1)) for the rebate rate over the next period (again, estimated from the data for each state). For transitions between states, we estimate the conditional expected rebate rate spread, conditional on the prior and current state. Thus, each period, we calculate the probability that the stock will go or remain special from week-to-week over the remaining life of the option, and then evaluate the expected cost, i.e., the cost of shorting over the life of the option. The key assumption is that past rebate rate spreads are sufficient to describe the expected movement in these spreads. Table 4 reports the results from the estimation of the model.

The probability transition matrix (Table 4B) shows that conditional on not being special, the probability of going special from week-to-week is very small—approximately 3.93%, only 0.59% of which is for rebate spreads below -0.5%. However, conditional on being special, the probability of remaining special is also high over the next week. For example, conditional on spreads being either between 0 and -0.5% or less than -0.5%, the probabilities of going off special are 15.21% or 2.96%, respectively, while the probabilities of remaining at the same degree of specialness are 77.79% and 88.58%. Mean reversion of negative rebate rate spreads is quite slow, i.e., the AR(1) coefficients equal 0.78 and 0.80, depending on the degree of specialness (Table 4D). Thus, assuming the stock stays special and that its current rebate rate spread is highly negative, the spread is expected to remain this way for quite a long time. This suggests that there are substantial costs to shorting certain stocks over the life of the option.

Table 3A reports regressions using Reb^A, the expected cost of short selling over the life of the option, which is calculated using the parameter estimates in Table 4. Both the explanatory power of the regressions (i.e., approximately 10%) and the economic implications of the coefficient estimates are very similar to those using the current rebate rate. For example, a one standard deviation decrease (i.e., 0.23%) in the adjusted rebate rate leads to a comparable 0.63% increase in the relative

Table 4
Distribution and time series model of rebate rate spreads
The table reports the cross-section and time series properties of the rebate rate spreads for the stocks in the at-the-money, intermediate maturity sample (see Table 1B). The analysis is done on the rebate spread, which is the difference between the actual rebate rate on a stock and the rebate rate on "cold stocks" that day. Panel A provides descriptive statistics on the distribution of the rebate spread for the entire sample. Panel B reports the 1-period transition probabilities between zero and two negative rebate rate spread states. Panel C reports the conditional means in period $t+1$ given the state in period t. Panel D reports estimates of an AR(1) model for rebate spreads conditional on spreads remaining in the same state (standard errors are in parentheses).

Panel A: distribution of rebate rate spreads

Range	Obs.	Mean	Median	5th pctl	95th pctl
$\text{Reb} = 0$	56,072	0	0	0	0
$-0.5 < \text{Reb} < 0$	12,590	-0.13	-0.06	-0.43	-0.01
$\text{Reb} \leqslant -0.5$	11,952	-3.09	-2.07	-7.19	-0.58

Panel B: transition probabilities between rebate spread states

Period		$t+1$		
		$\text{Reb} = 0$	$-0.5 < \text{Reb} < 0$	$\text{Reb} \leqslant -0.5$
t	$\text{Reb} = 0$	96.070	3.336	0.594
	$-0.5 < \text{Reb} < 0$	15.205	77.790	7.005
	$\text{Reb} \leqslant -0.5$	2.964	8.456	88.58

Panel C: means of period $t+1$ rebate spreads per state conditioned on period t state

Period		$t+1$		
	State(t)	$\text{Reb} = 0$	$-0.5 < \text{Reb} < 0$	$\text{Reb} \leqslant -0.5$
t	$\text{Reb} = 0$		-0.060	-2.442
	$-0.5 < \text{Reb} < 0$	0		-0.998
	$\text{Reb} \leqslant -0.5$	0	-0.261	

Panel D: AR(1) model for of negative rebate spreads within states

State	Const	AR(1)	R^2	Obs.
$-0.5 < \text{Reb} < 0$	-0.031[a]	0.783[a]	0.601	8073
	(0.001)	(0.007)		
$\text{Reb} \leqslant -0.5$	-0.666[a]	0.796[a]	0.639	8548
	(0.032)	(0.006)		

[a] Significant at the 1% level.

mispricing between the stock price and its implied value from options. One possible explanation for the similarity in the regression results is that the current rebate rate spread and our model-based short selling costs are highly correlated, with a correlation of 0.90. Interestingly, when the regression is performed over all the observations, including stocks with zero rebate rate spreads, the explanatory power drops. This suggests that our simple rebate rate model is not particularly helpful in explaining violations for zero rebate rate spread stocks.

Finally, if the rebate rate reflects only the extra income that a holder of the stock can make by lending it out (see Duffie et al., 2002), then the coefficient should be less than or equal to one in magnitude. In both Models 2 and 4 in Table 3A, the magnitudes of the coefficients are significantly larger than this bound, suggesting that something more is going on.

One natural question to ask is whether these put–call parity violations are consistent with the magnitude of short sales costs and other transactions costs in the options markets. This is an important question as there is some debate about the competitive nature of the equity lending market. In the next subsection, we bring evidence to bear on this question.

3.3. Transactions costs and put–call parity violations

Over a given horizon, investors can choose to purchase shares directly or replicate the share payoffs by going to the options market. Why would any investor choose the former if the latter market provides a much cheaper way of achieving the same payoffs? One possibility might be that transaction costs in the options market are too high (e.g., Nisbet, 1992). To investigate this hypothesis, we compare separately a long and short position in the stock versus the replication in the options market. In performing these calculations, we assume that the stock purchase is done at the last transaction price (be it a buy or a sell) and that one can borrow or lend at the same rate. In contrast, we assume purchases and sales of options are at the ask and bid prices, respectively. For example, we compare the prices of being long the stock to buying the call at its ask, selling the put at its bid, and lending the strike price. That is,

$$S^{L} \approx \mathrm{PV}(K) + C^{A} - P^{B} + \mathrm{EEP}, \tag{4}$$

where C^{A} and P^{B} are the ask and bid prices of the call and put, and S^{L} represents a long position in the stock. Similarly, a short position in the stock can be written as

$$S^{S} \approx \mathrm{PV}(K) + C^{B} - P^{A} + \mathrm{EEP}, \tag{5}$$

where S^{S} represents a short position in the stock. Combining (4) and (5) together provides a bound on how much the stock price can drift:

$$S^{S} \leqslant S \leqslant S^{L}. \tag{6}$$

The results are reported in Table 5A. Given the evidence in Table 2 that there are relatively few cases in which the stock price is below its implied value from the options market, it is not surprising that there are only a few cases in which the stock price drops below S^{S}. Table 5A shows that only 2.73% of the observations have stock prices that violate this condition. In contrast, violations on the other side are more numerous, with 12.23% of the observations exceeding S^{L}. This means that even in the presence of transactions costs (i.e., the bid–ask spread), it is cheaper to replicate payoffs using options than to purchase the shares directly. Why investors do not do this is a puzzle. At first glance, one reasonable possibility is that long-term investors may not wish to roll over their options positions from period to period (due to transactions costs). However, this argument does not hold for U.S. equity options as the investor can choose to take delivery of the stock upon exercise.

Table 5
Frequency of put–call parity violations after transactions costs
The table reports the distribution of put–call parity of violations (in percent) after accounting for transactions costs in the options market for the at-the-money, intermediate maturity sample (see Table 1B). There are a total of 80,614 observations, of which 24,542 have negative rebate rate spreads. The variable S^{S} is the lower bound on the stock price as derived from put–call parity (the implied short stock price), S^{M} is the stock price as derived from put–call parity when all option trades are traded at the midpoint, and S^{L} is the upper bound on the stock price as derived from put–call parity (the implied long stock price). In Panel A we use the observed stock price S, while in Panel B we use the stock price adjusted for the rebate rate cost over the life of the option (S^{A}) using the two-state AR(1) model estimated in Table 4. The three test statistics test: (1) and (2) whether the probability the stock price exceeds the upper bound is equal to the probability that the stock price is less than the lower bound for zero and negative rebate rate spread stocks, respectively, and (3) whether the probability of exceeding the upper bound is equal across zero and negative rebate rate spread stocks. The test statistics have an asymptotic N(0,1) distribution under the null hypotheses.

Panel A: *unadjusted stock price*

	$S<S^{S}$	$S^{S}\leqslant S<S^{M}$	$S^{M}\leqslant S\leqslant S^{L}$	$S>S^{L}$
All	2.73	32.17	52.87	12.23
Reb = 0	2.77	34.06	54.13	9.04
Reb < 0	2.65	27.85	49.99	19.51

Test	Stat	*P*-value
$\Pr(S<S^{S}\mid \text{Reb}=0)=\Pr(S>S^{L}\mid \text{Reb}=0)$	18.91	0.00
$\Pr(S<S^{S}\mid \text{Reb}<0)=\Pr(S>S^{L}\mid \text{Reb}<0)$	17.90	0.00
$\Pr(S>S^{L}\mid \text{Reb}=0)=\Pr(S>S^{L}\mid \text{Reb}<0)$	10.66	0.00

Panel B: *stock price adjusted for rebate rate cost*

	$S<S^{S}$	$S^{S}\leqslant S<S^{M}$	$S^{M}\leqslant S\leqslant S^{L}$	$S>S^{L}$
All	3.56	39.15	47.70	9.59
Reb = 0	3.21	38.75	50.21	7.82
Reb < 0	4.36	40.05	41.96	13.63

	Stat	*P*-value
$\Pr(S<S^{S}\mid \text{Reb}=0)=\Pr(S>S^{L}\mid \text{Reb}=0)$	15.18	0.00
$\Pr(S<S^{S}\mid \text{Reb}<0)=\Pr(S>S^{L}\mid \text{Reb}<0)$	11.68	0.00
$\Pr(S>S^{L}\mid \text{Reb}=0)=\Pr(S>S^{L}\mid \text{Reb}<0)$	7.08	0.00

These results are even more dramatic when we partition the sample of observations into groups with and without negative rebate rate spreads. Assuming that negative rebate rate spreads proxy for short sales restrictions, Table 5A shows that the violations are much more numerous for stocks that are short sales constrained. The percentages of put–call parity violations in the two samples are 19.51% and 9.04% relative to a long position in the stock, with a corresponding mean violation of 2.71% in the former case. This difference suggests that the equity market prices are further from fundamentals because, without short sales, the prices cannot be either driven back down by equity market short sellers or arbitraged away in the options market. Further evidence to this effect is presented in Section 4 below. On the other side of Eq. (6), and consistent with the asymmetric nature of short sales constraints, the violations are virtually identical, i.e., 2.65% and 2.77% for the two samples.

Given the persistence of short sales constraints as documented in Table 4, one might also expect the persistence of violations in the two tails to differ. In particular, stock prices less than the value of the synthetic short position may be in part due to measurement error, such as nonsynchronous trading in the stock and options markets, and thus should not persist from week to week. In fact, the autocorrelation of these violations is 0.23, less than half the autocorrelation of 0.58 for violations in the other tail of the distribution. Viewed slightly differently, the probabilities of seeing a violation for a particular stock in the following week, conditional on a violation this week, are 25% and 66% for the left and right tail, respectively.

Option spreads, however, are not the only transaction cost faced by investors. If an investor is able to short, then the rebate rate spread represents the cost of shorting. There is some debate, however, whether investors can actually locate and, equally important, maintain the short position when the stock is special, i.e., when its rebate rate spread is negative. The evidence in Section 3.2 above suggests this possibility may be empirically relevant. Nevertheless, it seems worthwhile taking the view that the equity lending market is a competitive market, and that the rebate rate represents the market rate all investors can obtain. In other words, there is limited arbitrage only to the extent that the rebate rate spread is negative, i.e., short selling, and therefore arbitrage, is attainable but at a cost.

Including the cost of shorting stocks when they are special implies a revision of Eq. (3), and therefore an adjustment to Eq. (6) above, namely

$$S^{\mathrm{A}} \equiv S(1-v) = \mathrm{PV}(K) + C - P + \mathrm{EEP}, \tag{7}$$

where v measures the spread between the rebate rate and the market rate. In theory, v represents the cost of shorting the stock over the life of the option, which may or may not equal the current rebate rate spread. For our purposes, we employ the three-state autoregressive model for rebate rates described in Section 3.2 above and documented in Table 4.

Table 5B looks at put–call parity violations assuming both that the rebate rate spread is a cost and that transactions take place at the bid and ask prices in the options market. Violations on the short sell side for negative rebate rate spread stocks are still more numerous, with 13.63% of the observations exceeding S^{L} versus only 7.82% for zero rebate rate stocks. While the fall from 19.51% to 13.63% once rebate rates are incorporated is clearly significant, it also shows that even with all transactions costs taken into account, violations of put–call parity remain. Moreover, the mean of these violations is 2.84%. We feel this provides further evidence that in practice the rebate rate spread represents not only a cost of transacting, but also the difficulty of shorting. For intuition, take the extreme case in which it is almost impossible to locate a short, i.e., search costs are close to infinite. The rebate rate is obviously not negative infinity in this case.

As a final look at the interaction between put–call parity deviations and transactions costs, we conduct the following volatility decomposition experiment. We take our measure of put–call parity deviations, R, without the adjustment for the early exercise premium, rebate rate spreads, and transactions costs in the options market. Conditional on negative spreads, how much of the variation in R is due to these various factors? Individually, the rebate rate, early exercise premium, and call and put spreads explain 10.8%, 1.1%, and 1.2% of the variation, respectively. For brevity, the regressions that yield these results are not reported. Collectively, they explain 14.1% of this variation. Dropping the rebate rate, early exercise premium and spreads from the regression in turn reduces the 14.1% to 3.4%, 13.2%, and 12.3%, respectively.

These results imply that shorting costs play a far more important role than the other factors. This result is economically intuitive. Negative rebate rate spreads are consistent with the stock being difficult to short. Shorting arises endogenously, possibly because of divergent opinions in the stock market, although shorting might also result from hedging needs. If this is the case and there is market segmentation between equities and options, for whatever reason, then put–call parity violations will result (e.g., Ofek and Richardson, 2003). In contrast, the presence of transactions costs yields no such prediction. It is a mistake to think that higher transactions costs imply larger put–call parity deviations. Asset pricing theory still implies that assets should be priced relative to their underlying payoffs. In fact, in our sample, put–call parity violations are lower in the presence of higher transactions costs.

4. Explaining the put–call parity violations: empirical analysis

Several important conclusions can be drawn from the stylized facts of Section 3. First, there is substantial evidence that across the universe of stocks, there are limits to arbitrage. A significant percentage of these stocks face short sales restrictions (e.g., over 10% of the observations are associated with negative rebate rate spreads of -1% or larger), which have an effect on the ability to conduct arbitrage between the equity and options markets. Second, and related, these limits to arbitrage lead to violations of put–call parity. Third, transactions costs, whether the shorting cost or the bid–ask spread in the options market, seem to limit the magnitude of these deviations in many cases.

However, even with transactions costs, the question of why the stock and options markets deviate in the first place remains. There are a few theories in the finance literature that might help answer this question. For example, Duffie et al. (2002) argue that stock prices can deviate from "fundamental value" because the stock price should also include the benefits derived from being able to lend out the stock to short-sellers. Of course, not all shares can be lent out, so the magnitude of this effect might be small. This point aside, put–call parity could be violated because the added benefit from the cash flow stream of possible share loans is similar to a stream of dividend payments. Dividends, if not accounted for, will lead to violations of Eq. (3).

We have also ignored frictions such as taxes and differences between borrowing and lending rates, although it is not clear exactly how these factors will affect put–call parity violations, especially in relation to the presence of short sales constraints. Finally, fluctuations in the value of the control rights associated with the equity, but not with the synthetic position in the options market, might also generate put–call parity violations under specific circumstances. This control rights effect also acts like a dividend if the value declines prior to option expiration. Moreover, there is anecdotal evidence that stocks go special during corporate events associated with changes in control, such as takeovers. Nevertheless, it is difficult to believe that declines in the value of control rights are pervasive enough to explain the observed results.

Alternatively, the growing literature in behavioral finance also suggests a possible explanation. A number of papers (e.g., Miller, 1977; Chen et al., 2002; Ofek and Richardson, 2003; among others) show that when investors with diverse beliefs face short sales constraints, prices can drift from fundamental values. Suppose there exist periods in which there are both overly optimistic investors and rational investors. The overly optimistic investors bid the prices of stocks up, but, due to short sales constraints, the rational investors do not simultaneously bid the shares back down. Thus, the stock price tends to drift above the value associated with aggregate beliefs.

Of course, the fact that stock prices drift from fundamental value does not necessarily lead to put–call parity violations. Why would these overly optimistic investors buy shares in the equity market when they could achieve the same payoffs at lower costs using options? One must also be willing to argue that the equity and options markets are sometimes segmented in terms of their investor classes; that is, these overly optimistic investors choose not to invest in the options market. One potential justification for this segmentation is that investors in the equity market trade frequently enough and in large enough volume that transactions costs and lack of depth in the options market prevent them from duplicating these trading patterns. Cochrane (2002) provides empirical evidence that these characteristics were present in numerous stocks during our sample period. Rational investors enter the options market, but, given short sales restrictions, cannot arbitrage between the two markets. The remainder of this paper focuses for the most part on building implications from this behavioral theory and then bringing evidence to bear on its validity.

Note that even in the above world with segmented markets, there still may not be put–call parity violations. Because option payoffs are based on the underlying share price, both the likelihood and magnitude of put–call parity violations depend on the probability and degree to which stock prices will eventually revert to fundamental value. Consider the extreme case in which prices never revert to their fundamental value. In this case, put–call parity will not be violated because options, as derivatives on the underlying stock, will reflect the expected stochastic process of the stock price. More generally, as long as the mispricing in the stock market is not expected to be corrected over the life of the option, there will not be put–call parity violations. Below, we describe three implications of the behavioral theory and the corresponding empirical evidence.

4.1. The maturity effect of put–call parity violations

Under the behavioral theory outlined above, and with short sales restrictions, put–call parity violations can occur if options investors believe that the stock price will revert, at least in part, to fundamental value over the life of the option. Thus, ceteris paribus, the put–call parity violation should increase in the maturity of the option as the expected magnitude of reversion to fundamental value increases.

Alternatively, the cost and difficulty of shorting may increase with the horizon length, as investors must pay the rebate rate spread over longer periods and short positions are more likely to be recalled. This alternative story also falls under the behavioral theory, and the implications for the maturity effect are the same. In this case, however, the cost of shorting replaces the speed of reversion to fundamental value. While the presence of a maturity effect cannot distinguish between these two alternatives, the risk-adjusted return on the stock over the life of the option will provide additional information, as discussed later in Section 4.3. In any case, the maturity effect will provide important information on the magnitude of mispricing and either the speed at which this mispricing is corrected or the cost of exploiting it.

Table 6 reports results on the relation between put–call parity violations and the maturity of the options. Specifically, whereas previous tables focused on intermediate-term options with a median expiration of 135 days, we now look at options with three different ranges of maturities: (i) short (i.e., 30–90 days, median 51 days), (ii) intermediate (i.e., 91–182 days, median 135 days), and (iii) long (i.e., 183–365 days, median 206 days) (see the data description in Section 2.3 and the appendix for further details). Note that the results for the intermediate maturity options are identical to those reported in Table 2 and are provided again for ease of comparison.

As can be seen from Table 6, the magnitudes of violations increase for longer maturity options. For stocks with negative rebate rate spreads, the mean violation for long maturities is 0.86%, versus 0.61% and 0.37% for medium and short maturity options, respectively. Violations increase less than linearly in maturity, but this result is to be expected under either the reversion to fundamental value or short sales costs explanations above. Mean reversion in either prices or rebate rate spreads will generate effects that attenuate over longer horizons. Interestingly, violations are still increasing past the intermediate maturity options, which extend to horizons of approximately six months. Thus, the results in Tables 3 and 5 that focus on these options may be understating the magnitudes of these effects. Of equal importance perhaps, the magnitudes of violations in the tails of the distribution are larger for long maturity options but only on the asymmetric side associated with shorting. For example, the 99th percentiles of the stock price ratios are 4.85%, 7.68%, and 9.07% for the short, intermediate, and long maturity options, respectively. In contrast, the 1st percentiles are very similar at −2.51%, −3.04%, and −2.91%, respectively. This evidence is consistent with limits to arbitrage (i.e., short sales constraints) mattering, but only to the extent that there is mispricing and the possibility that prices will revert to fundamental values (as measured by the maturity of the option).

Table 6
Put–call parity and option expiration
The table reports the distribution of the ratio $R \equiv 100\ln(S/S^*)$ for at-the-money, short (30 to 90 days), intermediate (91 to 181 days) and long (182 to 365 days) maturity options (see Table 2). The four test statistics are described in Table 2.

	Short			Intermediate			Long		
	All	Reb = 0	Reb<0	All	Reb = 0	Reb<0	All	Reb = 0	Reb<0
Obs.	75,771	52,439	23,332	80,614	56,072	24,542	32,652	22,891	9761
Mean	0.21	0.14	0.37	0.3	0.16	0.61	0.38	0.17	0.86
Percentiles									
1	−2.50	−2.49	−2.51	−2.93	−2.87	−3.04	−2.90	−2.88	−2.91
5	−1.03	−1.01	−1.09	−1.22	−1.19	−1.27	−1.13	−1.12	−1.15
10	−0.60	−0.58	−0.62	−0.68	−0.67	−0.69	−0.66	−0.67	−0.62
25	−0.14	−0.14	−0.13	−0.16	−0.18	−0.12	−0.18	−0.21	−0.08
50	0.16	0.13	0.24	0.20	0.16	0.35	0.22	0.15	0.46
75	0.51	0.44	0.70	0.65	0.53	1.02	0.69	0.53	1.31
90	1.03	0.88	1.40	1.33	1.04	2.04	1.54	1.04	2.70
95	1.53	1.28	2.12	1.97	1.49	2.97	2.36	1.56	4.01
99	3.34	2.54	4.85	4.42	2.82	7.68	5.37	2.97	9.07
$R<0$ (%)	35.21	36.41	32.52	34.90	36.83	30.50	35.49	38.51	28.40
$R>0$ (%)	64.79	63.59	67.48	65.10	63.17	69.50	64.51	61.49	71.60
Test		Stat	*P*-value		Stat	*P*-value		Stat	*P*-value
$E[R\|\text{Reb}=0] = E[R\|\text{Reb}<0]$		8.26	0.00		9.08	0.00		9.72	0.00
$\Pr(R>0\|\text{Reb}=0) = 50\%$		43.31	0.00		28.92	0.00		16.79	0.00
$\Pr(R>0\|\text{Reb}<0) = 50\%$		30.92	0.00		25.92	0.00		22.42	0.00
$\Pr(R>0\|\text{Reb}=0) = \Pr(R>0\|\text{Reb}<0)$		6.02	0.00		7.19	0.00		8.56	0.00

Under the behavioral theory, maturity should affect the magnitude of put–call parity violations but not necessarily the number of violations. Across all rebate rate spreads, the percentage of positive stock price ratios, R, is approximately 65% for all three maturity samples. Conditional on a negative rebate rate spread, the percent of positive violations does increase slightly in maturity, from 67.5% for short maturities to 71.6% for long maturities. However, this increase is minimal relative to the increase in the magnitudes of the violations.

In order to capture these differences while controlling for the size of the rebate rate spread, Fig. 2 provides a graphical representation of the regression of R on the rebate rate spread for the three partitions of the data—short, medium, and long maturity options. The coefficients on R in these three regressions are −0.13, −0.24, and −0.36, respectively. Again, they decrease less than linearly with maturity, as expected, but implied violations are still close to three times greater for long maturity versus short maturity options. These results are clearly consistent with behavioral biases among some investors in the equity market. Moreover, they suggest that these

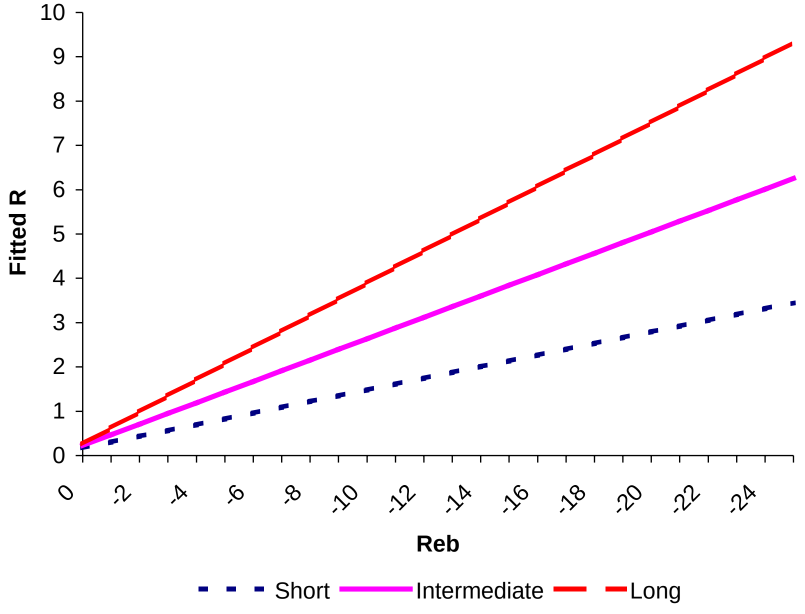

Fig. 2. Stock price ratios and maturity. The figure shows the fitted stock price ratio R from regressions for short, intermediate, and long maturity at-the-money options on the rebate rate spread (Reb). The sample period is July 1999 to November 2001. Table 6 reports information on the maturity-sorted samples.

biases and/or the costs of attempting to exploit them are quite persistent, with effects increasing out to horizons well past six months.

4.2. Structural shifts in mispricing

Under the behavioral theory, the magnitude of the put–call parity violation is related not only to maturity (as in Section 4.1 above) but also to the size of the disparity between the stock price and its fundamental value. If the put–call parity violation is small, it could be because the maturity of the option is short (i.e., a low probability of reversion or low short sales costs) or that the mispricing is small (i.e., the stock price reflects fundamental value).

To get at this latter point, it is worthwhile to condition on periods of possible equity mispricing and then look for violations of put–call parity in the options market. Of course, the difficulty with implementing such a test is that we do not know ex ante when these periods occur, if ever. The regression in Table 3B presents some suggestive evidence in that the admittedly noisy proxy for mispricing, the earnings-price ratio, enters with a negative and significant coefficient. Table 7 reports two additional tests. We choose the so-called crash of the NASDAQ as the structural shift in mispricing. From its peak in March 2000, the NASDAQ fell by approximately two-thirds over the subsequent year. The market declined further

Table 7
Structural change
Panel A reports the distribution of the ratio $R \equiv 100\ln(S/S^{*})$ (see Table 2) for two separate subperiods for at-the-money, intermediate maturity options (for negative rebate rate spread stocks only). The sample is divided by the technology crash into the subperiods July 1999 to February 2000 and May 2000 to November 2001. The two test statistics test for equality of means and the percentage of stocks with ratios greater than zero across the two subperiods. The test statistics have an asymptotic N(0,1) distribution under the null hypotheses. Panel B reports regressions of R on the rebate rate spread for negative rebate spread stocks for the two subperiods. The test statistic tests the equality of the coefficients on the rebate spread across the subperiods. Standard errors are in parentheses.

Panel A: distribution of unadjusted stock price ratios				
Sample	Mean	Median	$R>0$(%)	Obs.
7/99–2/00	0.687	0.490	74.151	7304
3/01–11/01	0.470	0.199	64.643	7068
Stat	2.410		5.942	
P-value	0.008		0.000	
Panel B: regressions				
Sample	Const.	Reb	R^2	Obs.
7/99–2/00	0.235[a]	−0.313[a]	0.110	7304
	(0.053)	(0.038)		
3/01–11/01	0.214[a]	−0.182[a]	0.121	7068
	(0.056)	(0.038)		
Stat		2.448[a]		
P-value		0.007		

[a] Significant at the 1% level.

thereafter, but at a much slower rate. Consequently, we define the pre- and post-crash periods as pre-March 2000 and post-March 2001, respectively. We first calculate both the percentage and magnitude of put–call parity violations during the two periods; the results are reported in Table 7A. Specifically, conditional on negative rebate rate spreads, the mean and median levels of R are 0.69% versus 0.47% and 0.49% versus 0.20% for the pre- and post-crash periods, respectively. These differences are statistically significant at the 1% level, and the results suggest put–call parity violations were affected by the NASDAQ crash. If the reader believes the crash was partly due to a correction in market mispricings, then these results are consistent with the aforementioned story of segmented markets, limited arbitrage, and put–call parity violations.

Second, using the pre- and post-crash periods, we test formally for the relation between put–call parity violations and the rebate rate spread. In other words, controlling for the level of short sales constraints, did the stock price ratio decline? Table 7B provides results from regressions of the violations, R, on rebate rate spreads, the same specification as estimated in Table 3A, as well as a formal test for the difference in the coefficients across the sample periods. The key result is that the slope coefficient is larger in magnitude pre-crash, which suggests that these violations are more sensitive to the existence of limits to arbitrage. That is, short sales restrictions are only relevant if mispricings do exist. The test for a structural change is statistically significant at the 1% level, and the difference is also economically significant.

Finally, in order to avoid specifying a particular date for the structural shift, we look at the relation between put–call parity violations and a continuous measure of mispricing, namely the P/E ratio of the S&P500. While the P/E ratio reflects the present value of growth opportunities and therefore can vary for quite rational reasons, we treat high (low) P/E ratios as reflective of overpricing (underpricing) for our purposes. Fig. 3 graphs the median put–call parity violation magnitude for stocks with and without negative rebate rate spreads and the S&P500 P/E ratio on a quarterly basis.

Several observations are in order. First, and perhaps most interesting, the time-series pattern in violations appears to match closely that of the P/E ratio of the S&P500, our measure of overvaluation. When the P/E ratio is high, at the beginning of the sample period, put–call parity violations are relatively large in magnitude. As the P/E ratio falls, the magnitude of violations also drops. The figure presents the data on a quarterly basis in order to smooth out some of the noise for presentation purposes, but, on a monthly basis, the correlation between the P/E ratio and the median violation for negative rebate rate spread stocks is an astonishing 0.76. This somewhat casual evidence clearly suggests a strong and positive relation between valuation levels in the market and the magnitude of put–call parity violations. Second, consistent with Table 7, there appears to be a structural shift in the magnitude of these violations in mid-2000. Anecdotally, this time frame is associated

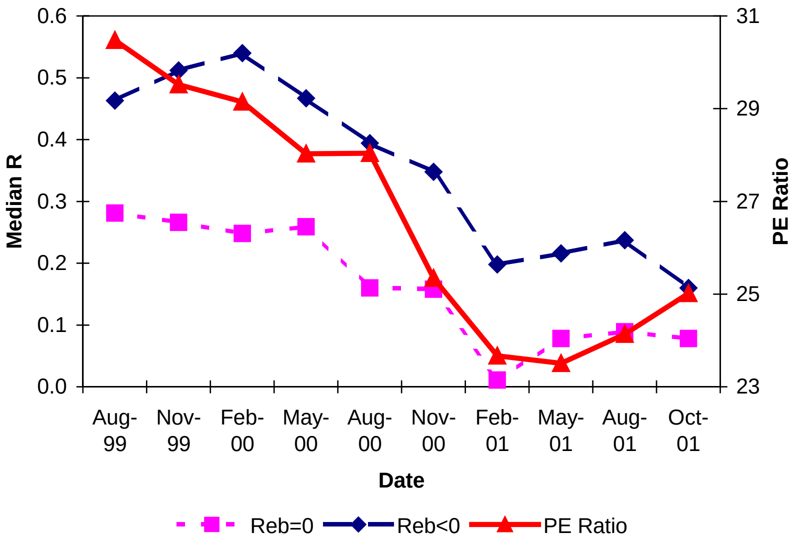

Fig. 3. Stock price ratios and PE ratios over time. The figure shows the median stock price ratio R for both zero (Reb $= 0$) and negative (Reb < 0) rebate rate spread stocks for the at-the-money, intermediate maturity sample (left axis), and the average PE ratio of the S&P500 (right axis) on a quarterly basis. The sample period is July 1999 to November 2001.

with the so-called bursting of the tech bubble, which many researchers consider a period of mass overvaluation. Of course, the P/E ratio also falls dramatically during this period. Third, before mid-2000, and after early 2001, the magnitudes of violations are fairly stable. The magnitudes, however, are at completely different levels. Again, this is consistent with the earlier period being governed by greater mispricings, and it also parallels the behavior of the P/E ratio. Fourth, the difference in magnitudes between the groups conditioned on rebate rate spreads is interesting. There is always a substantial difference, which is consistent with the rebate rate spread proxying for limited arbitrage conditions. Interestingly, after early 2001, there are few violations for normal rebate rate stocks, which is consistent with the forces of arbitrage. However, during the so-called bubble period, substantial violations still take place for stocks with normal rebate rates (albeit less than for stocks with negative spreads). Recall that the stocks in our sample do not pay dividends, which generally puts many of our stocks in the technology sector (e.g., technology, electronic equipment, semiconductor, and internet firms account for about 40% of the sample). Even if the rebate rate is normal, and this suggests (though not definitively) that one can short the stock today, there might be an expectation that shorting will be difficult in the future. Thus, violations can still occur over the life of the option.

4.3. Forecasting returns

Consider the behavioral model outlined above. In that world, option prices deviate from equity prices because rational investors price the assets in the options markets, and irrational investors price assets in the equity market. Arbitrage is not possible because investors cannot short in the equity market. Two factors limit the magnitude of the divergence between these markets: (i) some shorting (albeit at a cost) can take place, and (ii) there must be an expected convergence of these markets during the life of the option. With respect to this latter factor, this convergence suggests some form of predictability in stock returns. That is, assuming the rational investors accurately reflect the "truth" on average, we would expect stock returns to fall over the life of the option conditional on a put–call parity violation and/or a negative rebate rate spread. Our analysis is similar in spirit to that of Jones and Lamont (2002), who also look at the ability of short-selling costs to predict future returns. The key differences are that they examine a smaller cross-section of stocks (90 on average) for the period 1926 through 1933, and they condition only on short-selling costs and not on information from the options market. Nevertheless, their conclusions are similar.

One way to assess predictability is to examine the average excess stock return over the life of the option, conditional on available information such as the current put–call parity violation, rebate rate spread, and combinations of these variables.[3] For

[3] The theory implies that the difference between the option-implied stock price and the market price reflects the excess risk-adjusted return. We measure this excess return on each stock by subtracting out the corresponding industry return over the life of the option.

example, conditional on a rebate rate spread of less than −0.5%, the mean excess return over the life of the option is −9.96%, versus 0.70% for zero rebate rate stocks. Similarly, conditioning on put–call parity violations of greater than 1.0%, the mean excess return over the life of the option is −4.49% versus 0.13% for $R<0\%$. Combining these signals produces an average excess return of −12.57%, which illustrates that the rebate rate and the violation contain independent information about future stock price movements.

These returns are much larger in magnitude than both the estimated shorting costs over the life of the option and the put–call parity violation.[4] This result has several possible interpretations. First, it could be that over our sample period, corrections of mispricing occurred much faster than anticipated by the traders in the options market. Thus put–call parity violations underestimated future negative returns. Second, it is also consistent with short sales costs limiting the distance that options prices can deviate from stock prices. In other words, even if stocks are significantly overpriced and expected to revert to fundamental value quickly, the magnitudes of the put–call parity violations are limited by the cost of implementing the arbitrage between the two markets. Finally, interpreting the adjusted rebate rate spread as the income (dividend) that can be generated by lending out the stock is consistent with the direction but not the magnitude of the results (Duffie et al., 2002).

From a statistical standpoint, the mean excess returns should be interpreted with some caution for two reasons. First, the returns are calculated over the life of the option; we are therefore averaging returns across horizons ranging from 91 to 182 days. Second, we select stocks on every date; thus, the same stock may be selected on consecutive dates. The expiration date of the option may or may not be the same for these two observations, but in either case we include both returns in the sample. Clearly these returns will have a substantial overlap, and as a consequence we do not attempt to assess the statistical significance of these results.

Another way to evaluate the forecastability of returns that gets around these statistical issues is to evaluate a trading strategy that takes all the relevant costs into account. In particular, let us assume that shorting can take place albeit at the rebate rate spread. We form five different zero investment portfolios and follow their performance from week-to-week. In particular, we form a long portfolio of the relevant industry returns and a short portfolio of stocks that satisfy one of five different criteria: (i) stocks with a negative rebate rate spread less than −0.5%, (ii) stocks with a negative rebate rate spread less than −1.0%, (iii) stocks with put–call parity violations, (iv) stocks with put–call parity violations greater than 1%, and (v) stocks with both (i) and (iii). The portfolio has equal weights on all stocks satisfying the relevant criteria, and stocks are held until the expiration of the corresponding option.[5] Each week, the return on the portfolio is adjusted for the costs of shorting as described by the *actual* rebate rate spreads on the stocks in the portfolio.

[4] The former result is consistent with that of Jones and Lamont (2002), who also find that returns exceed the associated borrowing costs of the stock in their sample.

[5] Our nondividend paying stock sample includes a variety of technology, pharmaceutical, electronic equipment, semiconductor, and internet firms (with each of these industries accounting for approximately

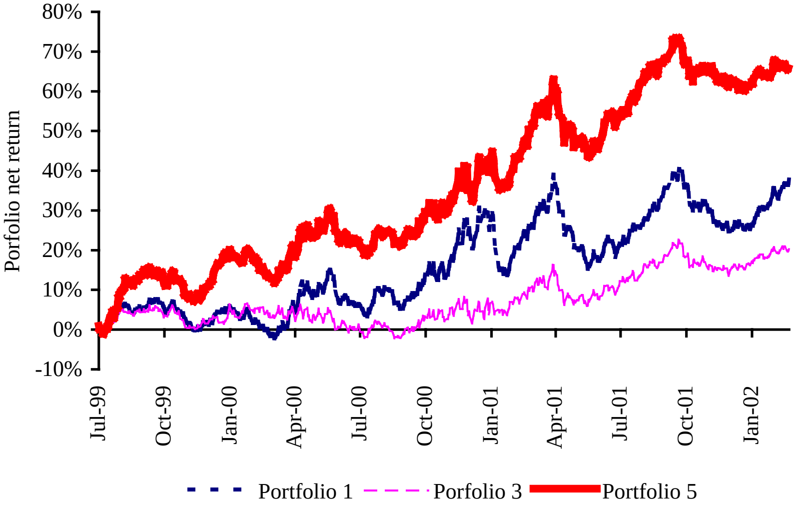

Fig. 4. Portfolio returns. The figure shows cumulative portfolio returns over the July 1999 to February 2002 period for portfolios 1, 3, and 5 from Table 9. These three strategies have short positions in stocks based on rebate rate spread and stock price ratio signals and long positions in the corresponding industry portfolios. Returns are net of shorting costs as measured by the rebate rate spread.

Fig. 4 graphs the returns on portfolios 1, 3, and 5 over the sample period. Irrespective of the criteria, the portfolios of stocks (with short signals) perform miserably relative to the weighted portfolio of corresponding industry returns. Thus, the zero investment portfolio produces large excess returns. For example, the cumulative returns on portfolios 1, 3, and 5 are approximately 38%, 20%, and 66%, respectively. As can be seen from the figure, the performance of the portfolio over the sample period suggests pervasive, and fairly consistent, poor returns on stocks that are subject to arbitrage constraints. We take this as evidence that there exist binding arbitrage constraints for a reason. Even if the above strategy is not implementable (i.e., the rebate rate represents more than just the cost of shorting), it presents a considerable puzzle to financial economists. Specifically, who is buying these arbitrage-constrained stocks at these inflated prices?

Table 8A documents the statistical properties of all five portfolios. While all the portfolios produce positive mean excess returns, the returns are higher the greater the arbitrage constraint. Changing the rebate rate criteria from -0.5% to -1.0% changes daily mean excess returns from 0.066% to 0.092%. If we adjust these returns for the daily cost of shorting (as defined by the actual rebate rate spread), the

(*footnote continued*)
10% of the sample), among 30 other industries. However, the portfolios, e.g., consider portfolio 5, are more concentrated in internet firms (25%) and pharmaceutical firms (20%). Young public companies are also somewhat over represented in the portfolios. For example, in portfolio 5, the median age is 3.2 years and 25% of the companies are within 1.2 years of their IPOs. Of course, if one's view is that these industries and firms are overpriced and subject to arbitrage limits, there is nothing surprising about this.

Table 8
Portfolio returns
The table reports returns characteristics of portfolios formed based on trading signals relating to the rebate spread and the unadjusted stock price ratio $R \equiv 100 \ln(S/S^*)$ (see Table 2). All portfolios start on July 1999 and close on February 2002 for a total of 666 trading days. The portfolios have zero net investment and stocks are equally weighted each day. All portfolios short stocks with the relevant signal and go long an equal amount in a matched industry portfolio. Daily return is the average daily return on the portfolio; Daily net return is the average daily return, net of the daily borrowing cost (rebate spread); STD net return is the daily standard deviation of the return on the portfolio; Short obs. is the average number of firms in the short portfolio per day. Panel B reports the intercept of the Fama-French three-factor model for daily portfolio returns. Gross α is the return on the portfolio. Net α is net of the rebate cost on the short position. t-statistics are in parentheses.

Panel A: *portfolio daily-return characteristics*

Portfolio	Filter	Daily gross return (%)	Daily net return (%)	STD net return (%)	Short obs.
1	Reb < −0.5%	0.066[c]	0.057	1.00	221
2	Reb < −1.0%	0.092[b]	0.081[b]	1.01	167
3	$R > 0\%$	0.034	0.030	0.81	318
4	$R > 1\%$	0.094[b]	0.085[b]	0.95	90
5	Reb < −0.5% and $R > 0\%$	0.113[a]	0.100[b]	1.05	93

Panel B: *intercept* (α) *of Fama-French three-factor model for daily portfolio returns*

Portfolio	Filter	Gross α	Net α
1	Reb < −0.5%	0.042%[c]	0.033%
		(1.79)	(1.41)
2	Reb < −1.0%	0.074%[a]	0.063%[b]
		(2.83)	(2.41)
3	$R > 0\%$	0.010%	0.006%
		(0.59)	(0.37)
4	$R > 1\%$	0.077%[a]	0.068%[b]
		(2.86)	(2.51)
5	Reb < −0.5% and $R > 0\%$	0.090%[a]	0.077%[a]
		(3.22)	(2.76)

[a] Significant at the 1% level.
[b] Significant at the 5% level.
[c] Significant at the 10% level.

corresponding net mean returns are 0.057% to 0.081%, representing only a slight drop. Interestingly, the volatilities across the portfolios are very similar. Thus, the standard risk-return tradeoff is not the source of these differences. While the means increase, the volatilities are stable at 1.00% and 1.01%, respectively.

The results above are adjusted for industry effects, but it is now fairly standard in the literature to also adjust returns for the three Fama and French (1992) factors, i.e., the market return, the return on a high-minus-low book-to-market portfolio, and the return on a small-minus-large firm portfolio. Estimating the coefficients on these factors using our five portfolio returns, we can estimate αs for each portfolio. Table 8B shows that on the whole, the αs tend to drop uniformly across our various portfolios relative to the industry adjustment alone, though only slightly. Moreover, because the variance of the residual has been reduced, the statistical significance actually increases for some of the portfolios. For example, for $R > 1\%$, though the gross mean αs drop from 0.94% to 0.77% when we include the Fama-French factors, the significance is below the 1% level, versus the 5% level before. The general conclusion can be drawn that the substantial gross and net returns documented in Table 8B are not driven by movements in aggregate factors over this period.

5. Conclusion

Shleifer (2000) argues that there are two necessary conditions for behavioral finance to have some chance of explaining financial asset prices, that is, for prices to deviate from fundamental value. The first is that some investors must be irrational, namely, they must ignore fundamental information or process irrelevant information in forming their trading decisions. The second is that there must be some limits to arbitrage such that this irrationality cannot get priced out of the market. In this paper, we look at a unique experiment that gets at these conditions. Specifically, by investigating the relation between equities and their corresponding options both under conditions of severe arbitrage constraints and little or no constraints, we are able to investigate this issue directly. The power of the analysis is greatly increased by looking across a large sample of stocks over a three-year period.

We provide empirical evidence that poses considerable problems for rational asset pricing models. Specifically, we show a strong relation between the rebate rate spread, which is a measure of short sales constraints, and the magnitude of put–call parity violations. This suggests a degree of mispricing across markets, although it is perhaps not arbitrageable. These results are consistent with a behavioral explanation to the extent that both the number and magnitude of these violations seem related to periods of mispricing and expectations that these mispricings will eventually be reversed.

One might conclude that the results in this paper support the foundations of behavioral finance, i.e., that there are enough irrational investors to matter for pricing assets. Researchers should find it heartening, however, that the forces of arbitrage do appear to limit the relative mispricing of assets. That is, there is a clear relation between arbitrage constraints (e.g., transactions costs, rebate rates and specialness in general) and the level of mispricing. On a more discouraging note, it remains a puzzle why any investor would ever wish to purchase such poorly performing stocks. We hypothesize that any explanation based on options completing the market will be a difficult story to swallow.

Appendix

All options data come from the Ivy DB database provided by OptionMetrics. This database contains option prices and related data "for the entire U.S. listed index and equity options markets" (IVY DB File and Data Reference Manual). The pricing data are compiled from raw end-of-day pricing information provided by Interactive Data Corporation. Other than contract-specific information (e.g., strike price, expiration date), our analysis uses two primary pieces of data:

1. Daily option (put and call) quotes (bid and ask prices), i.e., the best, or highest, closing bid price and the best, or lowest, closing ask price across all exchanges on which the option trades.
2. Daily continuously compounded zero-coupon interest rates whose maturities match the expiration dates on the options. These rates are calculated using interpolation from a zero curve generated using LIBOR rates and settlement prices of CME Eurodollar futures. (See the IVY DB File and Data Reference Manual for details).

Some of our analysis uses the option prices at the midpoint of the spread, i.e., the average of the bid and ask prices. We also calculate the option spread, i.e., the difference between the ask and bid prices as a percentage of the midpoint, to measure liquidity.

The rebate rate data come from a large dealer–broker and cover essentially all the stocks in the options database. Quotes for a given stock are sometimes missing, but we can detect no systematic pattern to these missing observations, and the number of missing observations is small. For each day and stock, we calculate the rebate rate spread (short selling cost) as the deviation of the rebate rate on that stock from the median rebate rate for that day, i.e., the cold rate. On every day, the majority of stocks have a rebate rate equal to the cold rate. Over time, the cold rate moves with prevailing market interest rates.

Starting with the above datasets, we select 118 dates between July 1999 and November 2001 that are approximately five business days apart. We then apply the following filters:

1. We eliminate all dividend-paying stocks. Thus, American call options can be treated as European call options and no dividend adjustments are necessary to compute option values and implied volatilities.
2. On each date, we eliminate options that have zero open interest. We use open interest as a proxy for liquidity in the options market. (Many of these options would also be eliminated by the moneyness filter discussed below since deep in- or out-of-the-money options tend to be the least liquid.)
3. On each date, we eliminate stocks (and the corresponding options) for which we do not have rebate rate data. While the rebate rate database is comprehensive, there are sometimes missing quotes.
4. On each date, we eliminate call and put options that do not have a corresponding put or call option with the same maturity and exercise price.

These filters leave us with pairs of matched call and put options on stocks with rebate rate data. Table 1A provides descriptive statistics on the options in this sample.

In order to maximize the quality of the data, we then apply a second set of filters:

1. On each date, we eliminate stocks (and the corresponding options) with prices less than $5.
2. We eliminate option pairs with maturities of less than 30 days or greater than 365 days.
3. We eliminate option pairs that are either deep in- or out-of-the-money ($|\ln(S/K)| > 0.3$).
4. We eliminate option pairs if either the put or the call has a bid–ask spread that is greater than 50% of the option price (at the midpoint). This filter catches both recording errors and options with very low liquidity.
5. We eliminate stocks (and the corresponding options) if the stock price ratio R exceeds 40.5 in absolute value. This filter also catches recording errors.
6. We eliminate option pairs if it is impossible to calculate the implied volatility of the call option because the option price (at the bid–ask midpoint) exceeds the stock price less the present value of the exercise price.

Finally we sort the option pairs into five moneyness/expiration groups as follows:

1. At-the-money, short maturity ($-0.1 < \ln(S/K) < 0.1$, 30–90 days)
2. At-the-money, intermediate maturity ($-0.1 < \ln(S/K) < 0.1$, 91–182 days)
3. At-the-money, long maturity ($-0.1 < \ln(S/K) < 0.1$, 183–365 days)
4. In-the-money, intermediate maturity ($0.1 < \ln(S/K) < 0.3$, 91–182 days)
5. Out-of-the-money, intermediate maturity ($-0.3 < \ln(S/K) < -0.1$, 91–182 days)

On any given date and for any given stock there may be multiple pairs that satisfy the moneyness and expiration criteria. If this is the case, we select the option pair that is closest to the middle of the range. Thus, there is only a single option pair per stock per date in the final sample. The reduction in the sample size from the full sample of over one million pairs to, for example, 80,614 pairs for the at-the-money, intermediate maturity sample is primarily due to the elimination of multiple option pairs for a stock on a given date and the moneyness/expiration grouping. The other filters eliminate relatively few observations.

The majority of the analysis is conducted with the at-the-money, intermediate maturity sample. The maturity effect (Table 6) is studied using the other two at-the-money samples. The effect of moneyness is studied using the other two intermediate maturity samples. Since moneyness has no apparent effect on the results, these results are neither reported nor discussed in the paper.

References

Almazan, A., Brown, K., Carlson, M., Chapman, D., 2002. Why constrain your mutual fund manager? Unpublished working paper, University of Texas, Austin.

Barberis, N., Thaler, R., 2003. A survey of behavioral finance. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. North-Holland, Amsterdam.

Basak, S., Croitoru, B., 2000. Equilibrium mispricing in a capital market with portfolio constraints. Review of Financial Studies 13, 715–748.

Bodurtha, J., Courtadon, G., 1986. Efficiency tests of the foreign currency options market. Journal of Finance 41, 151–162.
Chen, J., Hong, H., Stein, J., 2002. Breadth of ownership and stock returns. Journal of Financial Economics 66, 171–205.
Cochrane, J., 2002, Stocks as money: convenience yield and the tech-stock bubble. Unpublished working paper, University of Chicago.
D'Avolio, G., 2002. The market for borrowing stock. Journal of Financial Economics 66, 271–306.
Danielson, B., Sorescu, S., 2001. Why do options introductions depress stock prices? A study of diminishing short sales constraints. Journal of Financial and Quantitative Analysis 36, 451–484.
Detemple, J., Jorion, P., 1990. Option listing and stock returns: an empirical analysis. Journal of Banking and Finance 14, 781–802.
Detemple, J., Murthy, S., 1997. Equilibrium asset prices and no-arbitrage with portfolio constraints. Review of Financial Studies 10, 1133–1174.
Detemple, J., Selden, L., 1991. A general equilibrium analysis of option and stock market interactions. International Economic Review 32, 279–303.
Duffie, D., Garleanu, N., Pedersen, L., 2002. Securities lending, shorting and pricing. Journal of Financial Economics 66, 307–339.
Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.
Figlewski, S., 1981. The informational effects of restrictions on short sales: some empirical evidence. Journal of Financial and Quantitative Analysis 16, 463–476.
Figlewski, S., Webb, G., 1993. Options, short sales, and market completeness. Journal of Finance 48, 761–777.
Geczy, C., Musto, D., Reed, A., 2002. Stocks are special too: an analysis of the equity lending market. Journal of Financial Economics 66, 241–269.
Geske, R., Johnson, H., 1984. The American put option valued analytically. Journal of Finance 34, 1511–1524.
Gould, J., Galai, D., 1974. Transaction costs and the relationship between put and call prices. Journal of Financial Economics 1, 105–129.
Harrison, J., Kreps, D., 1978. Speculative investor behavior in a stock market with heterogenous expectations. Quarterly Journal of Economics 92, 323–336.
Ho, T., Stapleton, R., Subrahmanyam, M., 1994. A simple technique for the valuation and hedging of American options. Journal of Derivatives 1, 52–66.
Hong, H., Stein, J., 2002. Differences of opinion, short sales constraints and market crashes. Review of Financial Studies 16, 487–525.
Jarrow, R., 1981. Heterogeneous expectations, restrictions on short sales, and equilibrium asset prices. Journal of Finance 35, 1105–1113.
Johnson, H., 1983. An analytic approximation of the American put price. Journal of Financial and Quantitative Analysis 18, 141–148.
Jones, C., Lamont, O., 2002. Short sales constraints and stock returns. Journal of Financial Economics 66, 207–239.
Kamara, A., Miller Jr., T., 1995. Daily and intradaily tests of European put–call parity. Journal of Financial and Quantitative Analysis 30, 519–539.
Klemkosky, R., Resnick, B., 1979. Put–call parity and market efficiency. Journal of Finance 34, 1141–1155.
Lamont, O., Thaler, R., 2003. Can the market add and subtract? Mispricing in tech stock carve-outs. Journal of Political Economy 111, 227–268.
Lintner, J., 1969. The aggregation of investor's diverse jugdements and preferences in purely competitive strategy markets. Journal of Financial and Quantitative Analysis 4, 347–400.
Longstaff, F., 1995. Option pricing and the martingale restriction. Review of Financial Studies 8, 1091–1124.
McDonald, R., Shimko, D., 1998. The convenience yield of gold. Unpublished working paper, Northwestern University.
Merton, R., 1973. The theory of rational option pricing. Bell Journal of Economics 4, 141–183.

Miller, E., 1977. Risk, uncertainty, and divergence of opinion. Journal of Finance 32, 1151–1168.

Mitchell, M., Pulvino, T., Stafford, E., 2002. Limited arbitrage in equity markets. Journal of Finance 57, 551–584.

Nisbet, M., 1992. Put–call parity theory and an empirical test of the efficiency of the London traded options market. Journal of Banking and Finance 16, 381–403.

Ofek, E., Richardson, M., 2003. DotCom mania: the rise and fall of internet stock prices. Journal of Finance 58, 1113–1137.

Shleifer, A., 2000. Clarendon Lectures: Inefficient Markets. Oxford University Press, Oxford.

Unni, S., Yadav, P., 1999. Market value of early exercise: direct empirical evidence from American index option prices. Unpublished working paper, University of Strathclyde.