Early Exercise of American Put Options: Investor Rationality on the Swedish Equity Options Market

MALIN ENGSTRÖM
LARS NORDÉN*
ANDERS STRÖMBERG

Using Swedish equity option data, this study investigates how well the actual exercise behavior of American put options corresponds to the early exercise rules. The optimal exercise strategy is established in two ways. First, the critical exercise price, above which a put option should be exercised early, is computed and compared to the actual exercise price. Second, the exercise value of the option is compared to its market bid price. The results show that most early exercise decisions conform to rational exercise behavior, even though a large number of failures to exercise are found. Most of the faulty exercises

*For correspondence, Department of Corporate Finance, School of Business, Stockholm University, S-106 91 Stockholm, Sweden.
This study has benefited from comments and suggestions made by Hans-Peter Bermin, Nuray Guner, Niclas Hagelin, Martin Godsk Hansen, Peter Hördahl, Mark Shackleton, Vesa Puttonen, and two anonymous referees. We have also received valuable input from seminar participants at Lund University, Stockholm University, the 1997 meeting of the European Financial Management Association, the 1998 Nordic Symposium on Contingent Claims, the 1998 meeting of the Multinational Finance Society, and the 1998 meeting of the European Finance Association. Financial support from *Bankforskningsinstitutet* is gratefully acknowledged.

- *Malin Engström is a Ph.D. student at School of Business, Stockholm University.*
- *Lars Nordén is an Assistant Professor, School of Business, Stockholm University.*
- *Anders Strömberg is currently working at OM Stockholm AB.*

The Journal of Futures Markets, Vol. 20, No. 2, 167–188 (2000)

can also be discarded after a sensitivity analysis, although several failures to exercise are considered irrational, even after taking transaction costs into account. Jrl Fut Mark 20:167–188, 2000

INTRODUCTION

Equity option trading in Sweden began in 1985. Standardized call and put options on 38 individual stocks are at the present time listed on the Swedish exchange for options and other derivative securities, OM. The underlying stocks are in turn listed on the Stockholm Stock Exchange. Because the options are American, they provide the option holder with the right to exercise the contract any time prior to expiration.

For an American put option written on a dividend-paying stock, early exercise is most likely to occur immediately after the ex-dividend instant (see, for example, Geske & Shastri, 1985). However, early exercise may be optimal even if the underlying stock is not paying any dividends. In fact, an American put option should always be exercised early if it is sufficiently in-the-money. Optimal early exercise of an American call option, on the other hand, is only conceivable immediately prior to the dividend payment. Consequently, American calls on non–dividend-paying stocks should never be exercised early, and could therefore be valued as if they were European options.

In recent decades, many competing approaches to the theoretical valuation of the early exercise opportunity have been developed. These approaches are either analytical approximations or numerical methods. Examples of the first category are Roll (1977), Geske (1979), and Whaley (1981) for call options and Johnson (1983), Blomeyer (1986), Geske and Johnson (1984), and MacMillan (1986) for put options. In addition, Barone-Adesi and Whaley (1987) analyze both calls and puts. The second category can be further divided into lattice approaches (following Cox et al., 1979), finite difference methods (Brennan & Schwartz, 1977), and Monte Carlo simulation (Boyle, 1977).

Studies of the actual exercise behavior are interesting both for valuation purposes and for assessment of the assignment risk borne by the option writer. The purpose of this study is to empirically investigate the early exercise behavior of American put options written on non–dividend-paying stocks. Using Swedish data, the rationality of the actual exercise behavior is examined by comparing the observed behavior to the early exercise rules. The study relates to previous studies by Gay et al. (1989), Diz and Finucane (1993), Overdahl and Martin (1994), and Finucane

(1997). However, these studies are all performed using United States (U.S.) data, and the differences in size and liquidity between the U.S. and the Swedish markets might result in different exercise patterns. If, for instance, the market exhibits low liquidity, the bid/ask spreads could be very high. Early exercise might then be the least expensive possibility to close an option position. Because the liquidity at the Swedish options market is lower than at the U.S. options markets, early exercise is expected to occur more frequently in Sweden than in the U.S.

The early exercise rules are established in two ways. First, the Cox et al. (1979) binomial model is used to compute the critical exercise price for each option series. The critical exercise price is the exercise price at which an option holder is indifferent between exercising and keeping the option, and put options with exercise prices above (below) the critical exercise price should (not) be exercised early. The computed critical exercise prices are then compared to the actual exercise prices of the option series to establish if contracts have failed to be exercised or been faulty exercised. This is a new approach, because usually a critical stock price is computed to determine whether early exercise is optimal.[1] Second, a direct test using actual market prices is performed. In this model-independent test, the closing exercise values of the options are compared to their closing bid prices. Option exercise is here considered faulty if the value of the option if it is sold exceeds the exercise value, whereas a contract that is not exercised despite the exercise value being higher than the market bid price indicates a failure to exercise.

Inefficiencies or even irrationalities could explain the deviation of the market behavior from the theoretical behavior. However, market frictions such as transaction costs or incorrect input in the models, most importantly the estimated volatility, are also possible explanations. A sensitivity analysis is therefore performed to examine whether the deviations persist after taking these factors into account.

The model-dependent and model-independent tests give similar results in that numerous failures to exercise early, and a few examples of faulty exercise, are detected. Most of the faulty exercises could be discarded after a sensitivity analysis, whereas a large number of failures to exercise still may be genuine mistakes, even after transaction costs are taken into account. The economic significance of the results is also assessed, and, although occurring frequently, the failures to exercise do not cost very much on average. The faulty exercises are more rare, but are also more costly.

[1]However, the critical exercise price approach and the critical stock price approach give identical results, and the critical exercise price is used merely for its time-saving features.

PREVIOUS RESEARCH

Zivney (1991) argues that the value of early exercise is an empirical question. In order to measure the value, he examines deviations from put-call parity of the American style S&P 100 index options traded at the Chicago Board Options Exchange (CBOE). The results show that the value of early exercise is substantial, and that the decision to exercise early is not appropriately described by the existing theoretical pricing models for American options, in particular for puts. Using prices of both European and American options written on the FTSE 100 stock index traded at the London International Financial Futures and Options Exchange (LIFFE), McMurray and Yadav (1993) estimate the early exercise premium directly.[2] They find evidence of significant early exercise premiums for both call and put options. These premiums are higher than predicted by the deviations from put-call parity according to Zivney (1991).

Gay et al. (1989) examine the exercise of Treasury bond futures options traded at the Chicago Board of Trade (CBOT). The results show that over 50% of the exercises occurred before expiration, and that these exercises were dispersed over many different days. The Barone-Adesi and Whaley (1987) model (henceforth the BAW model) is used to compute the critical futures price below (above) which an American put (call) option should be exercised immediately. Both failures to exercise and exercises that should not have occurred are documented, at and before expiration, but the conclusion is that the exercise behavior in general is rational. The most frequent error type is failing to exercise, which the authors claim implies that investors do not sufficiently monitor their positions. The average loss per failure is also computed, and it is concluded that failures to exercise early occur frequently but cost little. Exercises that should not have occurred are more rare, but they cost more.

Diz and Finucane (1993) investigate the rationality of early option exercise by comparing observed exercise values of the S&P 100 index options to market bids, quoted at the CBOE. The idea is that early exercise is irrational if occurring when the bid quote exceeds the exercise value of the option contract. In that case, the optimal way to dispose of the option position would be to sell it. The authors provide evidence of inefficiencies and irrational exercise behavior. Furthermore, they find that dividend payments might induce early exercise, much in accordance with previous research.

[2]Unfortunately, for pairs of European and American options on the FTSE 100 index, with the same time left to expiration, there is never a common exercise price. In other words, the analysis of McMurray and Yadav (1993) is not entirely ceteris paribus with respect to the early exercise premium.

Overdahl and Martin (1994) analyze the observed exercise behavior of the equity options traded at the CBOE. Their results show that approximately 30% of all exercised calls and 60% of all exercised puts were exercised early. In some cases, options are exercised despite the fact that their market prices exceed the intrinsic values, or they are not exercised when they should be.[3] Nevertheless, most exercises conform to the theoretical exercise boundaries. In a study of equity call exercises using U.S. data, Finucane (1997) finds a number of early exercises at times other than ex-dividend dates. Most of these non–dividend-related early exercised calls might be attributed to transaction costs. However, the conclusion is that several contracts have been exercised irrationally.

INSTITUTIONAL SETTING

At OM, series of call and put option contracts with an initial time to expiration of six months are listed every three months.[4] This means that option series with two different expiration days are listed for trading at the same time.[5] The expiration day is the third Friday of the expiration month, and options on different individual stocks are divided into groups depending on expiration months.[6]

Option trading at OM normally takes place between 10:00 AM and 4:00 PM CET, but a holder of an equity option may exercise the contract until 5:00 PM on any trading day. On the expiration day exercise is possible until 6:00 PM.[7] After receipt of exercise notices from the option holders, OM assigns the exercises among the option writers randomly. On the expiration day, automatic exercise is applied for contracts with a minimum intrinsic value. A put option will be subject to automatic exercise if the exercise price is more than 1% above the average transaction stock price on that day. At the close of trading, OM sends to each member

[3]Unfortunately, Overdahl and Martin (1994) do not have access to the option bid/ask quotes as in Diz and Finucane (1993). To test whether it is optimal to exercise early, they compare the intrinsic value of an option to the observed option transaction price. One of their hypotheses is that early exercise will occur if the intrinsic value exceeds this transaction price. This is not necessarily the case because it is unknown whether the transactions were carried out on the bid or ask side.

[4]One put (call) option contract gives the holder the right to sell (buy) 100 shares of the underlying stock. A trading block in turn comprises 10 contracts.

[5]Long options—that is, contracts with an initial time to expiration of two years—are also listed each year for some of the underlying stocks. However, these are not included in the analysis because they are seldom traded.

[6]During the analyzed period the expiration months were distributed as follows: Group 1: January, April, July, and October; Group 2: February, May, August, and November; and Group 3: March, June, September, and December.

[7]This section gives a description of the institutional setting that prevailed during the analyzed period. Since then, the trading hours and the last possible exercise times have been changed.

of the exchange a list of the series of contracts with enough intrinsic value to be exercised automatically, and a written objection from the option holder is then necessary to prevent the option contracts from being automatically exercised.[8]

The total number of equity option contracts traded on OM during 1995 was 12,807,778, which corresponds to an average daily trading volume of around 51,000 contracts. Roughly 14,500 of these were put options, whereas the remaining 36,500 were call options.[9] A comparison with CBOE shows that the CBOE equity option trading volume during 1995 amounted to 77,040,466 contracts. This corresponds to an average daily volume figure of around 305,000 contracts, of which about 90,000 were put options.[10]

DATA

The data used in this study consist of the complete record of exercised put options, written on 23 underlying stocks, during the sample period 1 July 1995 to 1 February 1996.[11] The exercise data is obtained directly from OM and consists of information on the option contract code, the exercised volume of each option contract, and the day during which the exercise occurred. Hence, the analysis of option exercise is made on a daily basis. Daily closing transactions prices, trading volume, open interest as well as the closing bid/ask quotes of the options are also obtained from OM.

Daily closing bid/ask quotes and transaction prices, as well as high/low transaction prices, of the underlying stocks are acquired from the Stockholm Stock Exchange. Cash dividends are paid only once a year, and most dividend payouts occur around May. Because no dividend payments are made during the sample period, only puts written on non–dividend-paying stocks are analyzed in this study. Finally, daily rates of the Swedish one-month Treasury bills, obtained from the Swedish central bank, are used as a proxy for the riskless rate of interest in the binomial option valuation framework.

Summary statistics for the 273,292 exercised put option contracts are presented in Table I. According to the table most exercises occur at

[8]Unlike the U.S. options markets, the OM functions both as the exchange and the clearinghouse.
[9]During the analyzed period the average daily trading volume on the Stockholm Stock Exchange for the underlying stocks was SEK 2.6 billion for the on average 10,470 transactions per day. The market capitalization for the 228 listed Swedish stocks at the end of year 1995 was SEK 1,179 billion (approximately $147 billion).
[10]Source: OM Factbook and Futures and Options World, February 1996, respectively.
[11]The selection of the sample is made entirely on the basis of the availability of the data.

TABLE I

Summary of Put Option Exercises with Respect to Time to Expiration

				Contract days on which exercise occurred		
					% of OI exercised	
Weeks to expiration	*# Early exercised put option contracts*	*% of total early exercise*	*Average moneyness*	*Total*	*100%*	*<100%*
≤ 1	16,462	16.39	1.16	71	6	65
1–2	24,547	24.45	1.18	78	7	71
2–3	10,418	10.38	1.16	69	5	64
3–4	5,372	5.35	1.18	36	6	30
4–5	10,922	10.88	1.23	50	1	49
> 5	32,690	32.56	1.25	239	20	219
Total early exercise	100,411	100.0	1.21	543	45	498
Exercise at expiration	172,881		1.11	197	189	8
Total exercise	273,292		1.19	740	234	506

The table presents summary statistics for put options exercised at and before expiration, during the period 1 July 1995 to 1 February 1996. The early exercised contracts are displayed with respect to the time left to expiration. The average moneyness (which is defined as the ratio of the exercise price to the closing transaction stock price; X/S_{close}) weighted by contract days is also reported for each group, as well as the number of contract days on which exercise occurred. The last two columns indicate how many contract days the entire open interest (OI) for the contract is exercised (the 100% column) and the number of contract days when less than 100% of the open interest is exercised (the < 100% column).

expiration, but around one third of the contracts are exercised early.[12] This figure is rather close to the results of Diz and Finucane (1993), where 38% of the exercised puts are exercised early, but is considerably lower than the results of Overdahl and Martin (1994), where 60% of all exercised puts are exercised before expiration.[13] Furthermore, the summary statistics suggest that early exercise is quite common for a variety of times left to expiration, and that options exercised with a relatively long remaining time to expiration on average are deeper in-the-money.

To be able to examine the dispersion of exercises on different contracts, the number of contract days that exhibit exercises is reported. A contract day, as defined in Gay et al. (1989), is any trading day for an option with a specific exercise price and expiration day.[14] Table I shows that the early exercises occur on many different contract days. During

[12]For a more detailed analysis of the exercise pattern on the Swedish options market, the reader is referred to Engström et al. (1997).

[13]Brennan and Schwartz (1977) examine a group of 55 put options traded in the New York dealer market in 1966–1969. Only one of the 15 exercised options was exercised prior to expiration, resulting in an early exercise percentage of less than 7%. The popularity of early exercise thereby seems to have increased during the years.

[14]For example, a calendar day when options with five different strike prices and two expiration days are listed corresponds to ten contract days.

most of these contract days (498 out of 543) less than the entire open interest is exercised, suggesting that either those who exercise or those who do not exercise might make a mistake. This is true also for 8 expiration days, indicating that some option holders have objected to automatic exercise or that some option holders have exercised options that are less than 1% in-the-money.[15]

METHODOLOGY

Gay et al. (1989) use the BAW model to establish the theoretical early exercise rules, whereas model-independent tests are used in the other previous studies on early exercise. In this study, both model-dependent and model-independent tests are used to infer the rationality of the exercise behavior. However, the Cox et al. (1979) binomial model is the choice of model here. Although it is more computationally inefficient than the BAW model, the binomial model is known to produce reasonably accurate option values. In fact, Carr and Faguet (1994) find that the binomial model is more accurate than the BAW model in terms of option valuation.

The Binomial Model

In order to establish the optimal early exercise behavior, the binomial model is used to compute the critical exercise price at which an option holder is indifferent between exercising and keeping the option, given a certain stock price.[16] If a put option has a higher (lower) exercise price than the critical exercise price, the option should (not) be exercised early. The observed exercise behavior is then evaluated by comparing the computed critical exercise prices to the exercise prices of the option series, to determine whether early exercise was rational or not during each specific day. The critical exercise price gives the same exercise decision as a critical stock price, but the advantage with the critical exercise price is its time-saving features. If, at any time, several options on the same underlying stock are traded with a range of exercise prices, the single critical exercise price can be used to determine whether the whole range of options should be exercised, given the prevailing stock price. Otherwise, a critical stock price would have to be calculated for each option.

A problem concerning studies of exercise behavior is that it is impossible to know exactly which stock price within the day influenced the

[15]For all 8 expiration days, the moneyness measure used in this study is very close to 1.

[16]The procedure for finding the critical exercise price is discussed in the appendix.

option holder to make the exercise decision. Therefore, to avoid measurement errors, the daily critical exercise price is computed using the stock price during the day that gives the most conservative results. When examining failures to exercise puts, the highest stock price will yield the most conservative results. If a failure to exercise is observed when the highest stock price is used, it is also a failure using all other stock prices during that day. In this way, there will be a downward bias rather than an exaggeration of the number of failures. For the analysis of faulty exercised contracts, the lowest stock price gives the most conservative results. Exercised contracts that are considered to be faulty according to the lowest stock price will consequently be considered faulty also for every other stock price during that day.

The implied volatility, which is used as an estimate of the volatility of the underlying stock price, is calculated using the average of the same day's observed closing bid and ask quotes of the option nearest-the-money.[17] The implied volatility is found using a binomial tree with 100 time steps and the closing transaction stock price. For each day in the sample, volatility estimates are made for the two different expiration days traded at that time. An identical binomial tree is then used to compute the critical exercise price.

To analyze the rationality of option exercise, the difference between the critical exercise price X^* and the actual exercise price X, is divided by the actual exercise price to form an exercise ratio:

$$\text{Exercise ratio} = \frac{X^* - X}{X}. \tag{1}$$

If this ratio is negative, the put option should have been exercised early. The open interest at the end of a day with a negative ratio therefore represents contracts that failed to be exercised. If the exercise ratio is negative also on the following day, and the contracts still are not exercised, they are regarded as additional failures. Furthermore, exercised contracts with a positive exercise ratio are considered to be faulty, because the model value of the put is higher than the exercise value.

A Model-Independent Test

A test based on the exercise boundaries according to some pricing model, as above, is unfortunately a joint test of rationality and the model in

[17]The same day's average is used because the option holder may postpone exercise until after the close of the exchange and thereby have access to the prices of that day. If no bid or ask quotes exist, the option quotes for the day before are used.

question. A direct test, comparing directly the actual market value of the option to its value if exercised (as in, for example, Diz & Finucane, 1993) might be more indisputable. A condition for this comparison to be accurate, though, is that the two values can be found simultaneously and at the precise instant of the exercise decision. However, because exercise data only can be found on a daily basis and because no time-stamped intradaily price data is available, the only time during each day when simultaneous prices can be found is at the close of trading. Hence, as an additional test of rationality, the closing exercise values of the options are compared to their actual closing bid prices. When using this direct test, it is investigated whether there are any unexploited exercise possibilities, or if contracts are faulty exercised, at the close of trading for each day. However, because the options written on some of the stocks are rather infrequently traded, there could be problems with stale prices.

In the direct test it is assumed that the exercise of a put option is initiated by the purchase of the underlying stock. Hence, the exercise value is calculated as the exercise price less the price at which the underlying stock could be bought just before the close of trading. The option holder then has plenty of time to submit an exercise order because exercise may be postponed until an hour after the close of the options exchange. The difference between the exercise value and the bid price of the option is calculated as:

$$\text{Diff} = (X - S_{ask}) - P_{bid}, \tag{2}$$

where S_{ask} is the closing ask price of the underlying stock and P_{bid} is the closing bid price of the option. If Diff is positive, it is argued that profitable exercise possibilities indeed existed at the close of trading. The open interest at that time then represents contracts that failed to be exercised. If Diff instead is negative, there were no profitable exercise opportunities at that time because the optimal way to dispose of the option would be to sell it. Contracts exercised when Diff is negative are therefore considered to be faulty exercised. However, because profitable exercise possibilities may have existed during the day, it cannot be established whether the exercises actually are irrational.[18]

[18]A possible way to extend the analysis could be to compute the maximum difference during the day, to see if profitable exercise opportunities existed at all during that trading day. The maximum difference could be found by comparing the highest intrinsic value during the day to the lowest option price. If exercises were observed during a day with a negative maximum difference—that is, a day completely lacking profitable exercise opportunities—it would be certain that the observed exercised contracts actually are faulty. However, this analysis requires transaction prices and, because some of the options are traded very infrequently, the results of this test could be misleading.

Sensitivity Analysis

Deviations from the early exercise rules do not have to be irrational. The transaction costs facing an option holder may affect the exercise decision because the exercise ratio will have to exceed the transaction costs for early exercise to be optimal. OM charges an exercise fee of 0.225% of the exercise value for exercises amounting to less than SEK 500,000 and 0.075% for the exercise values above that amount. In addition, an option holder who wants to exercise will have to pay the usual brokerage fees of around 0.50%. A representative transaction cost could therefore amount to 0.725% of the exercise price.[19] Because the exact transaction costs facing each option holder is unknown, it is reasoned that a failure to exercise when the absolute value of the exercise ratio exceeds 0.725% is a mistake. If the transaction costs involved when selling an option are, due to the prevailing option bid/ask spread, much higher than the exercise costs, it could explain faulty exercises—because this then might be the most inexpensive way to dispose of the option. Otherwise, the transaction costs could not explain the faulty exercised puts, because they only add to the costs of a faulty exercise. In the direct test, the transaction costs are accounted for by incorporating them in the Diff calculation:

$$\text{Diff} = X(1 - t_{ex}) - S_{ask}(1 + t_s) - P_{bid}(1 - t_p), \tag{3}$$

where t_{ex} is the transaction costs associated with exercise, t_s is the transaction cost for the stock purchase and t_p is the cost for the sale of the option.

The results according to the binomial model could to some extent also be attributed to the choices of parameters in the option valuation framework. To investigate this possibility, a sensitivity analysis is carried out. Because the stock price that gives the most conservative results is used in the analysis, this could only lead to a downward bias in the results. Further, the time to expiration is a known parameter and can be discarded in a sensitivity analysis. The interest rate is observed in the market, and the estimation error is therefore not likely to be very large. Option prices are in general not very sensitive to interest rates, at reasonable levels, and this is true also for the results of this study. However, the robustness of

[19]The same transaction costs apply for early exercise as well as exercise on the expiration day. However, when the exercise is effected on a member's own account (and not on behalf of a customer), the exercise fee is considerably less than that mentioned above. Because there are also minimum absolute transaction costs, for a very small option position the relative transaction costs could amount to more than 0.725%. Because it is impossible to assess the actual cost for every position, it is assumed that the mentioned transaction costs are representative.

the results with respect to the volatility estimate used when calculating the critical exercise price has to be checked.

In this study, the implied volatility of the nearest-the-money option is used as the volatility estimate. If the volatility exhibits some kind of a smile pattern—that is, is not constant over different degrees of moneyness—it is possible to overestimate or underestimate the correct volatility, and thus to overestimate or underestimate the critical exercise price. If an option holder uses any other estimate of the volatility, the assessment of the critical exercise price is likely to be different, and the exercise decision may be optimal in the holder's view. To take this problem into account, whenever a faulty exercised contract or a failure to exercise is observed, the implied volatility is calculated, given that the option holder is indifferent between keeping and exercising the option. This critical volatility is then compared to the volatility estimate used in the study. If the deviation between the two estimates is large, it is argued that the observed exercise behavior indeed is incorrect and not only due to the volatility estimate.

Economic Significance

To assess the economic significance of incorrect exercise behavior, the actual loss due to a failure to exercise is estimated. Because failures are counted on a daily basis, it is reasoned that a failure to exercise the put one day (instead of on the following day) leads to a loss of interest on the exercise value for that day.[20] The following day, the option holder could make the appropriate decision, otherwise another day's interest is added to the loss for the previous day. Because transaction costs have to be paid regardless of which day the contract actually is exercised, the transaction costs are not included in the loss calculation. The exercise value is computed using the highest stock price during the day for the binomial model, and the closing ask price of the stock for the direct test.

The loss due to a faulty exercise could be much higher than for a failure to exercise. Because the option is now "dead," the mistake is irreparable. The incurred loss is the difference between the market value of the option and its value if exercised. In addition, the option holder must pay transaction costs for the exercise, which makes the loss even higher.

[20]The loss might very well be calculated differently. The intention here is just to get an approximation of the magnitude of the loss, not to calculate the exact loss facing each option holder that fails to exercise.

TABLE II

Summary of Failures to Exercise Put Options Early According to the Binomial Model

$\frac{X^* - X}{X}$	*Number of failures to exercise*	*%*	*% of acc. open interest*	*Contract days*	*Average moneyness*	*Average volatility difference*	*Average days left to expiration*	*Total loss*	*Average loss*
0.00–0.25	104,536	5.89	0.42	127	1.14	0.0041[a]	53.76	53,522	0.51
0.25–0.50	79,503	4.48	0.32	116	1.14	0.0086[a]	59.67	41,919	0.53
0.50–0.75	96,567	5.44	0.38	129	1.14	0.0160[a]	56.33	51,215	0.53
0.75–1.00	80,765	4.55	0.32	125	1.14	0.0244[a]	55.42	38,775	0.48
1.00–2.00	307,209	17.31	1.22	428	1.15	0.0361[a]	55.53	146,346	0.48
2.00–3.00	201,114	11.33	0.80	384	1.17	0.0605[a]	53.62	101,919	0.51
3.00–4.00	188,695	10.63	0.75	341	1.18	0.0768[a]	54.48	98,588	0.52
4.00–5.00	138,693	7.81	0.55	322	1.19	0.1041[a]	54.48	74,651	0.54
5.00–6.00	100,307	5.65	0.40	273	1.21	0.1362[a]	55.10	61,216	0.61
6.00–7.00	112,791	6.35	0.45	267	1.23	0.1412[a]	51.70	70,323	0.62
7.00–8.00	85,315	4.81	0.34	226	1.24	0.1496[a]	53.18	62,540	0.73
8.00–9.00	58,710	3.31	0.23	204	1.25	0.1726[a]	49.70	38,162	0.65
9.00–10.0	53,590	3.02	0.21	166	1.26	0.1821[a]	45.00	34,533	0.64
> 10.0	167,220	9.42	0.66	947	1.35	0.2412[a]	47.87	143,155	0.86
Total	1,775,015	100.0	7.05	4,055	1.23	0.1194[a]	52.34	1,016,863	0.57

The table shows the number of failures to exercise according to the exercise ratio (which is negative for failures to exercise) expressed in percentage terms. The failures are also reported as percentages of all failures and of the accumulated open interest for days exhibiting failures respectively. Reported also is the number of contract days that exhibit failures to exercise as well as the average moneyness of the options in each group, where the moneyness is defined as X/S_{close}. In addition, the average difference between the critical volatility and the implied volatility according to the binomial model is presented for each group. The symbol [a] indicates that the average volatility difference is significantly different from zero at a p-value < 0.001. Finally, the average time to expiration and the loss associated with failures to exercise are reported.

EMPIRICAL RESULTS

This section contains the empirical analysis of the study. First, the actual exercise behavior is compared to the theoretically optimal exercise behavior according to the binomial model. Then the observed behavior is compared to the optimal behavior according to the direct test. Both subsections contain a sensitivity analysis and an evaluation of the economic significance of an incorrect exercise decision.

Actual Exercise Behavior Compared to Optimal According to the Binomial Model

The failures to exercise put options during the sample period are reported in Table II.[21] Failures to exercise are very common, both when the num-

[21]The results in the table are reported for all underlying stocks together. The results for the individual stocks are available on request, and vary substantially for the different stocks. Stocks with a higher number of failures tend to have the failures dispersed over more contract days and also tend to have

ber of contracts and the number of contract days are considered. The total number of option contracts that failed to be exercised according to the binomial model is 1,775,015 and these failures occur on 4,055 different contract days. To be able to put the number of failures into perspective, the total open interest accumulated over trading days exhibiting failures is also computed, and amounts to 25,160,405 contracts. On average, the failures account for around 7% of the accumulated open interest.[22]

The failures to exercise are presented in Table II according to the size of the exercise ratio. Most failures belong to the groups with the lowest exercise ratios, and the group with the lowest ratio also reports four times as many failures per contract day as the group with the highest ratio.[23] However, almost 10% of the failures have an exercise ratio in excess of 10%. The average moneyness for each group is also computed, and as expected it increases with the exercise ratio, whereas the average time to expiration is about the same in each group.

A reported failure to exercise is considered to be a real mistake if the absolute value of the exercise ratio exceeds the transaction cost of 0.725%. More than 85% of the failures in Table II correspond to a ratio higher than 0.725%, and it is therefore concluded that most of the failures cannot be explained by transaction costs. The results might also depend on the volatility estimates used in the study. In Table II, the average volatility difference for each exercise ratio group is reported. If the deviation between the critical volatility and the implied volatility according to the binomial model is large, it is reasoned that the observed failures to exercise are actual mistakes and not only due to the volatility estimate. The average volatility difference is significantly different from zero for all groups, and ranges from less than half a percentage point to almost 25 percentage points. As expected, the volatility difference increases with the exercise ratio. For failures with an exercise ratio above 4%, the average volatility difference is over 10 percentage points, and it is therefore improbable that these results are caused by measurement errors in the volatility estimates only. Consequently, these failures, in total 716,626, are considered as genuine mistakes made by the option holders.

The loss from a failure to exercise is computed as one day's interest on the difference between the exercise price and the highest stock price during the day. The total loss during the sample period is SEK 1,016,863,

a higher accumulated open interest, but there is no clear pattern in the differences. Astra and Ericsson, which are very frequently traded, show comparatively few failures, whereas Trelleborg, SCA, and Volvo, which are also rather frequently traded, show the highest number of failures.

[22]The average is significantly different from zero at a p-value < 0.001.

[23]The exercise ratio is negative for failures, but for ease of exposition the absolute values are reported.

TABLE III

Summary of Faulty Early Exercised Put Options According to the Binomial Model

$\frac{X^* - X}{X}$	Number of faulty exercised contracts	% of all faulty exercised contracts	% of all exercised contracts	Contract days	Average moneyness	Average volatility difference	Average days left to expiration
0.00–0.25	40	0.24	0.04	2	1.17	−0.0013	77.00
0.25–0.50	8,620	52.40	8.58	10	1.15	−0.0063	39.10
0.50–0.75	985	5.99	0.98	9	1.12	−0.0162	54.22
0.75–1.00	315	1.91	0.31	7	1.11	−0.0168[a]	47.71
1.00–2.00	1,170	7.11	1.17	11	1.14	−0.0295[a]	55.36
2.00–3.00	1,120	6.81	1.12	12	1.13	−0.0375[a]	59.17
3.00–4.00	1,350	8.21	1.34	9	1.09	−0.0591[a]	42.33
4.00–5.00	1,200	7.29	1.20	5	1.16	−0.0540[a]	77.00
> 5.00	1,650	10.03	1.64	22	1.14	−0.1523[a]	45.00
Total	16,450	100.0	16.38	87	1.13	−0.0604[a]	51.06

The table shows the number of faulty exercised contracts according to the exercise ratio (which is positive for faulty exercises) expressed in percentage terms. The faulty exercises are also reported as percentages of all faulty exercised contracts as well as of all exercised contracts respectively. Presented also is the number of contract days that exhibit faulty exercise and the average moneyness of the options in each group, where the moneyness is defined as X/S_{close}. In addition, the average difference between the critical volatility and the implied volatility according to the binomial model is reported for each group. The symbol [a] indicates that the average volatility difference is significantly different from zero at a p-value < 0.001. Finally, the average time to expiration is presented for each group.

with an average loss of SEK 0.57 per contract. The conclusion is therefore that although failures occur frequently, they do not cost very much per contract and day.

Table III reports the faulty exercised contracts. According to the binomial model, a total of 16,450 (16.38% of all 100,411 early exercised puts) contracts are exercised when exercise should not have occurred. These contracts are exercised on 87 different contract days. Hence, the faulty exercised contracts do not occur nearly as frequently as the failures to exercise. The faulty exercises are displayed in Table III with respect to the exercise ratio. Over 50% of the faulty exercises correspond to an exercise ratio less than 0.50%, which implies that they are very sensitive to measurement errors. Contrary to the results for the failures to exercise, there is no clear pattern between moneyness and exercise ratio. There is also more dispersion in the time to expiration of the exercised options.

In Table III, it can be seen that the faulty exercises are very sensitive to the volatility, because the deviation between the two estimates most of the time is within five percentage points. However, for the group with an exercise ratio above 5%, the average volatility difference is over 15 percentage points, which is significantly different from zero.[24] This group, consisting of around a tenth of the faulty exercises, are therefore considered as genuine mistakes, whereas the results concerning the other faulty exercises have to be interpreted with more care.

A reasonable explanation to the faulty exercised puts could be that the options market exhibits low liquidity. In that case a contract may be exercised instead of sold, because this might be the most inexpensive way to dispose of the option.[25] The Swedish market exhibits some rather inactively traded options, and the average daily trading volume is therefore calculated for all contract days and contract days that exhibit faulty exercised puts, respectively. The calculations are repeated three times; for all options and for two subgroups consisting of the 10 most frequently traded options and the rest of the options, respectively. The results in all three groups show that the average daily trading volume is higher for days exhibiting faulty exercises than for all contract days. Hence, low liquidity is unlikely to be the sole explanation for the faulty exercised contracts.[26]

Actual Exercise Behavior Compared to Optimal According to the Direct Test

The results of the direct test, which are presented in Panel A of Table IV, show a total of 3,400,654 contracts that failed to be exercised.[27] These failures occur on 4,078 different contract days, and account for on average 12% of the open interest accumulated over days exhibiting failures, which amounts to 28,150,219 contracts.[28] However, for many of the detected failures to exercise the differences between exercise values and option prices are very small. The average difference is SEK 0.74, which is significantly different from zero at a p-value < 0.001.

The failures to exercise are in Panel A of Table IV presented according to the size of the difference between exercise value and bid price for the options. Over 50% of the failures correspond to a difference below

[24]Because the deviation between the critical volatility and the volatility estimate is negative when faulty exercises are observed, the absolute value of the deviation is discussed in this section.

[25]Low liquidity therefore tends to make a failure to exercise more serious.

[26]To test whether the results depend on the choice of theoretical model, the BAW model is also used to analyze the exercise rationality. The two models give very similar numbers of failures to exercise and faulty exercises, which enhances the robustness of the results.

[27]The results are reported for all of the underlying stocks together, but the results for the individual stocks are available on request. The results are very similar to the results of the model-dependent test in that Trelleborg, SCA, and Volvo show the highest number of failures, whereas Astra and Ericsson show comparatively few failures. No clear pattern between the number of failures and, for example, liquidity is discernible.

[28]The average is significantly different from zero at a p-value < 0.001.

TABLE IV

Summary of Failures to Exercise Put Options Early According to the Direct Test

Difference in SEK	*Number of failures to exercise*	*%*	*% of acc. open interest*	*Contract days*	*Average days left to expiration*	*Total loss*	*Average loss*
Panel A: No Transactions Costs							
0.00–0.25	874,903	25.73	3.11	948	60.69	272,477	0.31
0.25–0.50	896,114	26.35	3.18	1,036	60.35	340,088	0.38
0.50–0.75	661,856	19.46	2.35	737	60.42	262,842	0.40
0.75–1.00	400,927	11.79	1.42	555	59.12	186,552	0.47
1.00–2.00	471,651	13.87	1.68	683	57.63	238,384	0.51
2.00–3.00	58,308	1.71	0.21	88	50.82	46,199	0.79
3.00–4.00	27,974	0.82	0.10	17	50.94	22,188	0.79
> 4.00	8,921	0.26	0.03	14	39.43	9,623	1.08
Total	3,400,654	100.0	12.08	4,078	59.50	1,378,353	0.41
Panel B: Including Transactions Costs							
0.00–0.25	317,190	47.23	1.13	333	66.23	128,259	0.40
0.25–0.50	173,379	25.82	0.62	180	66.37	70,310	0.41
0.50–0.75	81,329	12.11	0.29	66	61.39	35,038	0.43
0.75–1.00	24,470	3.64	0.09	36	51.61	8,915	0.36
1.00–2.00	53,978	8.04	0.19	58	48.03	29,280	0.54
2.00–3.00	13,430	2.00	0.05	7	61.29	7,832	0.58
3.00–4.00	410	0.06	0.00	2	94.50	232	0.57
> 4.00	7,341	1.09	0.03	11	41.00	7,693	1.05
Total	671,527	100.0	2.39	693	63.15	287,560	0.43

The tables show the number of failures to exercise according to the difference in SEK between the exercise value and the market price of the option. The failures are also reported as percentages of the accumulated open interest for days exhibiting failures. In addition, the number of contract days when failures to exercise occur is presented. Finally, the average time to expiration and the loss associated with failures to exercise are reported in each panel.

SEK 0.50, and the small differences imply that the results are sensitive to transaction costs. In Panel B of Table IV, where the results are reported after transaction costs, as in eq. (3), the number of failures decreases drastically compared to the results in Panel A. The failures remaining after transaction costs amount to 671,527 (19.75% of the original figure) contracts and occur on 693 (17%) contract days.[29] Hence, although the number of failures is greatly diminished, a substantial number of failures to exercise could still be regarded as actual mistakes made by the option holders.

[29]Because the transaction costs for an individual investor are used in the test, these results are conservative.

TABLE V

Summary of Faulty Early Exercised Put Options According to the Direct Test

Difference in SEK	*Number of faulty exercised contracts*	*% of all faulty exercised contracts*	*% of all exercised contracts*	*Contract days*	*Average days left to expiration*
0.00–0.25	2,450	15.39	2.44	16	32.38
0.25–0.50	9,840	61.80	9.80	65	40.82
0.50–0.75	60	0.38	0.06	3	23.00
0.75–1.00	1,890	11.87	1.88	28	32.36
1.00–2.00	1,582	9.94	1.58	10	20.90
2.00–3.00	60	0.38	0.06	2	63.00
3.00–4.00	10	0.06	0.01	1	15.00
> 4.00	30	0.19	0.03	2	55.50
Total	15,922	100.00	15.86	127	36.28

The table shows the number of faulty exercised contracts according to the absolute difference in SEK between the exercise value and the market price of the option. (The difference is negative for a faulty exercised contract, but for ease of exposition the absolute difference is reported.) The faulty exercises are also reported as percentages of all faulty exercised contracts as well as of all exercised contracts respectively. Presented also is the number of contract days that exhibit faulty exercises and the average time to expiration for each group.

In the direct test, the loss incurred by a failure to exercise is computed as one day's interest on the difference between the exercise price and the closing ask price of the stock. The total loss, SEK 1,378,353, is a bit higher than for the binomial model due to the higher number of failures reported. However, the average loss per contract and day, SEK 0.41, is lower.

The faulty exercises are in Table V presented according to the size of the difference between exercise values and option bid prices.[30] The 15,922 faulty exercised contracts (15.86% of all 100,411 early exercised puts) occur on 127 different contract days. However, the differences between exercise values and market prices are in most cases very small. About 77% of the contracts correspond to a difference below SEK 0.50, and several of the apparent irrationalities might therefore be explained by measurement errors. The average difference between exercise value and market price is SEK 0.65, which is significantly different from zero at a p-value < 0.001.

The differences between the option market prices and corresponding exercise values at the close of trading are used to get an indication of the loss for the faulty exercised contracts. The calculated losses range from SEK 2 to SEK 775 per contract, with an average loss of around SEK 65.

[30]The difference is negative for faulty exercised contracts, but for ease of exposition the absolute values are reported.

The direct test gives results that are of the same magnitude as the results according to the binomial model, and, regardless of which of the tests is used, numerous failures to exercise are found, as are some examples of faulty exercise. For the majority of the contracts, the model-dependent and model-independent tests also indicate the same deviations from the early exercise rules. This is especially the case when the contracts with the largest deviations (that is, highest exercise ratio or largest difference between exercise value and option bid price) are analyzed. For these options, the overlapping is almost total. In other words, both tests detect roughly the same failures to exercise and faulty exercises, and the similarity between the outcomes enhances the robustness of the results.

CONCLUDING REMARKS

This study investigates the exercise behavior of American equity put options using Swedish data. Both model-dependent and model-independent tests are used to establish the optimal exercise behavior. Regardless of which of the tests is used, two types of incorrect exercise behavior are found: faulty exercise and failures to exercise, where the latter ones are much more common. A substantial amount of failures to exercise still remain after transaction costs and differences in volatility estimation have been taken into consideration, whereas the results concerning faulty exercises are very sensitive to the volatility estimate used in the theoretical model. Thus, many of the failures to exercise may be genuine mistakes, whereas most of the faulty exercises could be considered rational.

Comparisons with the results of the earlier studies show that several failures to exercise, as well as some cases of faulty exercise, are found both in this study and in the U.S. studies. However, a comparatively higher number of failures to exercise are detected in this study. This is somewhat surprising, because the relatively low liquidity at the Swedish options market is inclined to make early exercise more favorable relative to closing an option position by selling it. The faulty exercised contracts, on the other hand, are fewer in this study than in, for example, Diz and Finucane (1993), where about 35% of the exercises of index puts are inefficient. Although many of these can be explained, the authors find instances of clearly irrational exercise. Overdahl and Martin (1994) analyze the exercise of equity options, and, compared to their closing prices results, this study reports more failures to exercise and fewer cases of faulty exercise. This could imply that Swedish option holders are less inclined to exercise contracts early than option holders in the U.S. Therefore, according to the results, liquidity does not have the expected impact on the exercise behavior, and other factors that can explain the differences have to be found.

Taxation might be a candidate to explain the observed exercise behavior. During the analyzed period, capital gains and losses had differential tax treatment, in that losses only were deductible to 70%. This could induce investors to match capital gains and losses, because a loss could be deducted to 100% against a capital gain in the same year. Hence, the exercise of an in-the-money put could be postponed to a year when the investor has an offsetting loss, which could explain a failure to exercise. Furthermore, the option holder's decision may be correct according to the holder's preexisting position. If the option is held as part of a hedge, it may be optimal not to exercise it early even if it should be exercised early according to the optimal exercise rules. Hedging could clearly explain some of the failures to exercise, but, because no information on the proportion of hedging transactions in the Swedish option market is available, it is not possible to assess the impact this explanation could have on the results.

The numerous examples of failures to exercise could also imply that the option holders do not monitor their positions with sufficient care. However, because each contract that failed to be exercised did not cost very much per day, the monitoring cost could be higher than the loss caused by the negligence. Nevertheless, it is not possible to rule out that the exercise behavior to some extent actually is irrational. The faulty actual exercises are more rare than the failures to exercise, but they are also more costly.

The results of the study are potentially good news for those interested in assessing the assignment risk borne by the option writer, because it shows that options rarely are exercised early when the theory says that they should not.

APPENDIX: CRITICAL EXERCISE PRICES FOR AMERICAN PUT OPTIONS

Consider a non–dividend-paying stock having a market price equal to S. The change in the stock price dS can be described by the stochastic differential equation:

$$dS = \mu S dt + \sigma S dW, \tag{A.1}$$

where μ is the instantaneous expected rate of stock return, σ is the instantaneous volatility of the stock returns, and dW is the change in a standard Wiener process. Given (A.1), a binomial tree can be constructed for the evolution of the stock price with the parameters:

$$u = e^{\sigma\sqrt{\Delta t}}, \; d = e^{-\sigma\sqrt{\Delta t}}, \tag{A.2}$$

where Δt is the time evolving between changes in the stock price, u is the multiplicative upward movement in the stock price, and d is the corresponding downward movement.

In the binomial framework, the value of an American put option can be obtained from the following recursive procedure:

$$P_{i,j} = \max[e^{-r\Delta t}(pP_{i+1,j+1} + (1 - p)P_{i+1,j}); X - S_{i,j}], \tag{A.3}$$

where $P_{i,j}$ is the value of an American put option at a time $i\Delta t$, after j upward movements in the stock price. $p = (e^{r\Delta t} - d)/(u - d)$ can be interpreted as the risk-neutral probability of an upward movement in the stock price, $X - S_{i,j}$ is the exercise value of the option, and r is the riskless rate of interest.

Suppose that several put options exist at time $i\Delta t$ with a range of exercise prices. Then it would be interesting to know at which exercise price an option holder is indifferent between exercising and keeping the option, given a certain stock price. This critical exercise price X^* can be obtained numerically by equating the terms within the brackets in equation (A.3):

$$e^{-r\Delta t}(pP^*_{i+1,j+1} + (1 - p)P^*_{i+1,j}) = X^* - S_{i,j}. \tag{A.4}$$

The put option prices $P^*_{i+1,j+1}$ and $P^*_{i+1,j}$ are functions of X^*, and the critical exercise price is found by iterating through the binomial tree for different exercise prices while keeping the stock price fixed at $S_{i,j}$ throughout the iterations.

BIBLIOGRAPHY

Barone-Adesi, G., & Whaley, R. (1987). Efficient analytic approximation of American option values. Journal of Finance, 42, 301–320.

Blomeyer, E. (1986). An analytical approximation for the American put price for options on stocks with dividends. Journal of Financial and Quantitative Analysis, 21, 229–233.

Boyle, P. (1977). Options: A Monte Carlo approach. Journal of Financial Economics, 4, 323–338.

Brennan, M., & Schwartz, E. (1977). The valuation of American put options. Journal of Finance, 32, 449–462.

Carr, P. & Faguet, D. (1994). Fast accurate valuation of American options. Working paper, Cornell University, USA.

Cox, J., Ross, S., & Rubinstein, M. (1979). Option pricing: A simplified approach. Journal of Financial Economics, 13, 145–166.

Diz, F., & Finucane, T. (1993). The rationality of early exercise decisions: Evidence from the S&P 100 index options market. Review of Financial Studies, 6, 765–798.

Engström, M., Nordén, L., & Strömberg, A. (1997). Early exercise of American equity options: A study of the exercise pattern at the Swedish options market. Working paper, Stockholm University.

Finucane, T. (1997). An empirical analysis of common stock call exercise: A note. Journal of Banking and Finance, 21, 563–571.

Gay, G., Kolb, R., & Yung, K. (1989). Trader rationality in the exercise of futures options. Journal of Financial Economics, 23, 339–361.

Geske, R. (1979). A note on an analytical formula for unprotected American call options on stocks with known dividends. Journal of Financial Economics, 7, 375–380.

Geske, R., & Johnson, H. (1984). The American put valued analytically. Journal of Finance, 39, 1511–1524.

Geske, R., & Shastri, K. (1985). The early exercise of American puts. Journal of Banking and Finance, 9, 207–219.

Johnson, H. (1983). An analytical approximation for the American put price. Journal of Financial and Quantitative Analysis, 18, 141–148.

MacMillan, L. (1986). Analytic approximation for the American put option. Advances in Futures and Options Research, 1, 119–139.

McMurray, L., & Yadav, P. (1993). The early exercise premium in American option prices: Direct empirical evidence. Paper presented at the European Finance Association 1993 annual meeting.

Overdahl, J., & Martin, P. (1994). The exercise of equity options: Theory and empirical tests. Journal of Derivatives, 2, 38–51.

Roll, R. (1977). An analytical valuation formula for unprotected American options on stocks with known dividends. Journal of Financial Economics, 5, 251–258.

Whaley, R. (1981). On the valuation of American call options on stocks with known dividends. Journal of Financial Economics, 9, 207–211.

Zivney, T. (1991). The value of early exercise in option prices: An empirical investigation. Journal of Financial and Quantitative Analysis, 26, 129–138.