# The Implications of Early Exercise Policies for Option and Stock Returns

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities

2021

**Adnan Gazi**

**Alliance Manchester Business School**

**Accounting and Finance Division**

Dedicated to my little daughter, Armita . . .

# Acknowledgements

Firstly, I would like to express my deepest gratitude to my PhD supervisors, Prof. Kevin Aretz and Prof. Ian Garrett, for their continuous support, help and insightful advices throughout my PhD. Working with them, I have not only learned the techniques and rigour of conducting research but also grasped the idea of staying calm and focused even when ‘everything else is falling apart’. Without their motivation and encouragement, it would not have been possible for me to come this far with my thesis.

I am also thankful to Prof. Carol Alexandar, Prof. Michael Brennan, Dr. Eirini Konstantinidi, Prof. Stephen Figlewski, Prof. Massimo Guidolin, Prof. Kris Jacobs, Dr. Olga Kolokolova, Prof. Hening Liu, Prof. Tyler Shumway, Prof. Christian Wagner, Prof. Robert Whaley and participants from a number of local and international seminars, workshops and conferences for valuable comments and suggestions on chapters of my thesis. My thesis has come to its current state from their discerning instructions and directions.

I have had a wonderful and memorable four-year journey in Manchester and it would have been impossible without my PhD colleagues and friends. I am indebted to Mohammed Hassan, Tasmima Muniya, Shadab Mashuk, Kashfia Mashuk, Ahmed Ameya Prapan, Xinyu Cui, Xiangshang Cai, Shuwen Yang, Kevin Schneider, Wei Liu and Manar Elgazzar for always being there for me. I am specially grateful to my fellow batch mate and good friend Nestor Romero with whom I have shared many ups and downs throughout the PhD. I wish I could share these final few steps with him though I know he would always be watching us from heaven, smiling and cheering, like he always used to do.

Finally, I am specially grateful for the unconditional love and support from my wife Paromita and daughter Armita. My journey so far would have been unimaginable without them by my side at all the time. I am also thankful to my younger sibling Asif and to my parents for their continuous guidance.

Adnan Gazi
Manchester, UK

# Declaration

I, Adnan Gazi, declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

# Copyright Statement

# Abstract

In the first chapter of my thesis, I study the asset pricing implications of being able to optimally early exercise a plain-vanilla put option, contrasting the expected returns of equivalent American and European put options. Standard pricing models with stochastic volatility and asset-value jumps suggest that the expected return spread between them is positive, can be economically sizable, and widens with a higher optimal early exercise probability, as induced through a higher moneyness, shorter time-to-maturity, or lower underlying-asset volatility. Studying single-stock American put options and equivalent synthetic European options formed from applying put-call parity to American call options on zero-dividend stocks, my empirical work supports the theoretical predictions. My results, therefore, indicate that the early exercise feature can have a strong effect on option returns.

In the second chapter, I introduce a dynamic trading strategy based on a theoretical proposition of Shreve (2004). Many studies report that American option investors often exercise their positions suboptimally late. Yet, when that can happen in case of puts, there is an arbitrage opportunity in perfect markets, mentioned in Shreve (2004), exploitable by longing the asset-and-riskfree-asset portfolio replicating the put and shorting the put. Using early exercise data, I show that the arbitrage strategy also earns a highly significant mean return with low risk in real single-stock put markets, in which exactly replicating options is impossible. In line with theory, the strategy performs particularly well on high strike-price puts in high interest-rate regimes. It further performs well on short time-to-maturity puts on low volatility stocks, consistent with evidence that investors do not correctly incorporate those characteristics into their exercise decisions. The strategy survives accounting for trading and short-selling costs, at least when executed on liquid assets.

In the third chapter, I revisit the value-weighted stock return predictability of Black-Scholes (1973) option implied volatility spreads. Studies so far have explained this predictability using investors' informed trading activities in options ahead of the stock market and/or frictions in the underlying stock. Nevertheless, for single-stock American options, I show that the ability of implied volatility spreads to predict cross sectional stock returns is primarily driven by the friction-induced optimal early exercise of put options that is not accounted for in calculating implied volatility. The contribution of other factors to the predictive ability of implied volatility spreads are largely insignificant. Further evidence suggests that the predictability cannot be solely explained by the trading activities of informed option investors.

# Table of Contents

**Acknowledgements** **2**

**Declaration** **3**

**Copyright Statement** **4**

**Abstract** **5**

**List of Figures** **8**

**List of Tables** **9**

**1 The Early Exercise Risk Premium** **11**

1.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

1.2 Theory . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

1.2.1 A Monte Carlo Simulation Exercise . . . . . . . . . . . . . 17

1.3 Data and Methodology . . . . . . . . . . . . . . . . . . . . . . . 24

1.3.1 Data Sources and Filters . . . . . . . . . . . . . . . . . . . 24

1.3.2 Calculating Single-Stock Option Returns . . . . . . . . . . 25

1.4 Cross-Sectional Evidence . . . . . . . . . . . . . . . . . . . . . . 27

1.4.1 The Early Exercise Risk Premium in the Data . . . . . . . 28

1.4.2 Relations with Stock and Option Characteristics . . . . . . 30

1.5 Robustness Tests . . . . . . . . . . . . . . . . . . . . . . . . . . 35

1.5.1 Violations of Early Exercise Rules and Put-Call Parity . . 36

1.5.2 Identification of Zero-Dividend Stocks . . . . . . . . . . . . 38

1.5.3 Bid-Ask Transaction Costs . . . . . . . . . . . . . . . . . . 39

1.6 Time-Series Evidence . . . . . . . . . . . . . . . . . . . . . . . . 40

1.7 Concluding Remarks . . . . . . . . . . . . . . . . . . . . . . . . 42

**2 Taking Money Off the Table: Suboptimal Early Exercises, Risky Arbitrage, and American Put Returns** **60**

2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

2.2 Theory . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 65

2.3 Methodology and Data . . . . . . . . . . . . . . . . . . . . . . . 70

2.3.1 Calculating the American Put Return . . . . . . . . . . . . 70
2.3.2 Calculating the Replication Portfolio Return . . . . . . . . 72
2.3.3 Calculating Theoretical Early Exercise Probabilities . . . . 74
2.3.4 Data and Data Sources . . . . . . . . . . . . . . . . . . . . 75
2.4 The Performance of the Risky Arbitrage Strategy . . . . . . . . . . . . . . . . . . . . . . . . . . . 77
2.4.1 Mean Return and Volatility of the Strategy . . . . . . . . 77
2.4.2 The Factor Model Alphas of the Strategy . . . . . . . . . . 79
2.5 Conditioning the Strategy Return . . . . . . . . . . . . . . . . . 81
2.5.1 Conditioning on the Strike Price and the Interest Rate . . 81
2.5.2 Conditioning on Moneyness, Maturity Time, and Volatility 84
2.5.3 Jointly Conditioning on All Factors . . . . . . . . . . . . . 87
2.5.4 Why Do Moneyness, Time-to-Maturity, and Stock Volatility Condition the Success of the Risky Arbitrage Strategy? . . 88
2.6 Adjusting for Transaction Costs . . . . . . . . . . . . . . . . . . 91
2.7 Concluding Remarks . . . . . . . . . . . . . . . . . . . . . . . . 93

**3 Why Does the Implied Volatility Spread Predict Future Stock Returns? 110**
3.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . 110
3.2 Data and Methodology . . . . . . . . . . . . . . . . . . . . . . . 116
3.2.1 Data Sources and Filters . . . . . . . . . . . . . . . . . . . 116
3.2.2 Calculating The Early Exercise Premium and Implied Volatility Spread . . . . . . . . . . . . . . . . . . . . . . . . 117
3.3 Empirical Results . . . . . . . . . . . . . . . . . . . . . . . . . . 119
3.3.1 Option Data Characteristics . . . . . . . . . . . . . . . . . 119
3.3.2 Pre-formation Stock Portfolio Characteristics . . . . . . . 121
3.3.3 Does The Early Exercise Premium Predict Future Stock Returns? . . . . . . . . . . . . . . . . . . . . . . . . . . . 122
3.4 Robustness Tests . . . . . . . . . . . . . . . . . . . . . . . . . . 127
3.4.1 Violations of Early Exercise Rules and Put-Call Parity . . . . 128
3.4.2 Option Moneyness and Time-to-Maturity . . . . . . . . . . 130
3.4.3 Further Checks . . . . . . . . . . . . . . . . . . . . . . . . 131
3.5 The Role of Informed Trading . . . . . . . . . . . . . . . . . . . 133
3.6 Concluding Remarks . . . . . . . . . . . . . . . . . . . . . . . . 134

# List of Figures

Figure 1.1: The Early Exercise Risk Premium in a GBM World . . . . . . 22
Figure 1.2: Cumulative Profits From Selling American and European Put Options . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29

Figure 2.1: Monthly Returns From Replicating Portfolio and Traded Put Positions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 79
Figure 2.2: Cumulative Profits From Replicating Portfolio and Traded Put Positions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 79

# List of Tables

1.1 The Early Exercise Risk Premium in a GBM World . . . . . . . 47
1.2 The Early Exercise Risk Premium, Stochastic Volatility, and Jumps 48
1.3 Descriptive Statistics . . . . . . . . . . . . . . . . . . . . . . . . . . 49
1.4 Portfolios Double-Sorted on Moneyness and Maturity Time . . . . 50
1.5 Portfolios Univariately-Sorted on Idiosyncratic Volatility . . . . . 51
1.6 Subperiod Tests . . . . . . . . . . . . . . . . . . . . . . . . . . 52
1.7 Fama-MacBeth (1973) Regressions . . . . . . . . . . . . . . . . . 53
1.8 Controlling for Stock Short-Selling Constraints . . . . . . . . . . 54
1.9 Controlling for Stock and Option Illiquidity . . . . . . . . . . . . 55
1.10 Stocks That Never Paid Dividends . . . . . . . . . . . . . . . . . 56
1.11 Accounting for Bid-Ask Transaction Costs . . . . . . . . . . . . . 57
1.12 Time-Series Asset Pricing Tests . . . . . . . . . . . . . . . . . . . 59

2.1 The Performance of the Risky Arbitrage Strategy . . . . . . . . . 97
2.2 Adjusting the Strategy's Performance for Systematic Risk . . . . 98
2.3 Conditioning the Strategy on the Strike Price . . . . . . . . . . . 99
2.4 Conditioning the Strategy on the Strike Price and Interest Rate . 100
2.5 Conditioning the Strategy on Moneyness and Maturity Time . . . 101
2.6 Conditioning the Strategy on Idiosyncratic Stock Volatility . . . . 102
2.7 Fama-MacBeth (1973) Regressions . . . . . . . . . . . . . . . . . 103
2.8 Mean Black-Scholes Vs. Real-World Early Exercise Probabilities . 104
2.9 Mean Difference Between Black-Scholes and Real-World Early Exercise Probabilities By Moneyness, Maturity Time, and Volatility 105
2.10 Fama-MacBeth Regressions Explaining the Difference Between Black-Scholes and Real-World Early Exercise Probabilities . . . . 106
2.11 Adjusting the Strategy for Transaction Costs . . . . . . . . . . . . 107
2.12 Adjusting the Strategy for Transaction Costs Under Different Replication Portfolio Rebalancing Schemes . . . . . . . . . . . . . . . . 109

3.1 Descriptive Statistics for the Option Data . . . . . . . . . . . . . 138
3.2 Average Preformation Characteristics of Portfolios Sorted on the Early Exercise Premium . . . . . . . . . . . . . . . . . . . . . . . 139

3.3 Decomposing the Implied Volatility Spread . . . . . . . . . . . . . 140
3.4 Univariate Portfolio Sorts . . . . . . . . . . . . . . . . . . . . . . . 141
3.5 Portfolio Alphas . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 142
3.6 Portfolios Double-Sorted on the Early Exercise Premium and Residual Frictions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 143
3.7 Fama-MacBeth (1973) Regressions . . . . . . . . . . . . . . . . . . 144
3.8 Controlling for Stock Short-Selling Constraints . . . . . . . . . . . 145
3.9 Controlling for Stock Short-Selling Constraints and Option Illiquidity 146
3.10 Controlling for Option Moneyness and Time-to-Maturity . . . . . 147
3.11 Additional Robustness Tests . . . . . . . . . . . . . . . . . . . . . . 148
3.12 Only Including Observations on Options With Zero Trading Volume 149

This thesis contains 43,312 words, including title page, tables, and footnotes.

# Chapter 1

# The Early Exercise Risk Premium

***Keywords***: Empirical asset pricing; cross-sectional option pricing; put options; early exercise.

## 1.1 Introduction

Although starting from Black and Scholes (1973) an abundance of studies consider the valuation of plain-vanilla and exotic options, only recently a much smaller number have started looking into the systematic risk of these assets. Among those latter studies, Coval and Shumway (2001) study how the strike price affects the expected return of a European option, while Hu and Jacobs (2018) and Aretz et al. (2019) study how underlying-asset volatility does. Surprisingly, however, there is so far no research into how the ability to optimally early exercise an option influences the expected option return, perhaps due to the widely-held view that the early exercise feature of American options has only minimal implications.[1] In accordance with that view, a large empirical literature uses American option data in European option

[1] This view is often credited to the theoretical studies of Brennan and Schwartz (1977) and Broadie et al. (2007). Using a wide variety of stochastic processes to model the evolution of the underlying asset's value, these studies show that the values of American options are generally close to those of European options.

pricing models, effectively assuming that American and European options are "almost identical" (see, e.g., Bakshi et al. (2003), Carr and Wu (2008), Hu and Jacobs (2018), Martin and Wagner (2019), and others).

In our paper, we present theoretical and empirical evidence on how being able to optimally early exercise a plain-vanilla put option affects the option's expected return. On the theoretical front, we calculate expected American and European put option returns through simulating the value of the underlying asset from popular stochastic processes, showing that the expected return spread between an American option and its equivalent European option (which we label the "early exercise risk premium") is positive and reaches up to 8.14% *per month*. We further show that the spread increases with factors positively related to the early exercise probability. On the empirical front, we study the mean return spread between single-stock American put options and equivalent *synthetic* European options, formed from applying put-call parity to American call options written on zero-dividend stocks. In line with our theory, our estimate of the early exercise risk premium is positive, reaches up to 11.74% per month ($t$-statistic: 12.04), and significantly increases with factors positively conditioning the early exercise probability.

We rely on Longstaff and Schwartz's (2001) least-squares method to calculate the expected returns of American and European options from simulations. More specifically, we simulate a large number of daily underlying-asset-value paths under the physical and risk-neutral measure from Geometric Brownian motion, a stochastic volatility, and a stochastic-volatility jump model (see Bates (1996)). Dividing the mean maturity option payoff under the physical measure by the mean discounted maturity payoff under the risk-neutral measure, we obtain the expected European option return. To calculate the expected American option return, we move backward from maturity, on each date and for each path comparing the early exercise payoff with the expected discounted risk-neutral option value from a least-squares regression. Doing so enables us to delineate the optimal early exercise boundary (i.e., the underlying-asset value for which an early exercise is as attractive as keeping the option alive). Dividing the mean compounded-up earliest option payoff under the physical measure by the mean discounted earliest option payoff under the risk-neutral measure, we obtain the expected American option return.

The simulations show that while, in line with Merton (1973), American options have higher values than the equivalent European options, they have — in proportional terms — even higher expected payoffs, leading them to have higher (i.e., less negative) expected returns than the European options. The intuition behind this result, as first pointed out by Barraclough and Whaley (2012), is that early exercising an American put option is equivalent to converting a risky asset into a risk-free bond, skewing systematic option risk toward zero. The simulations further show that the magnitude of the early exercise risk premium can meaningfully depend on factors positively conditioning the early exercise probability. The premium, for example, increases significantly with a higher option moneyness, a shorter time-to-maturity, and a lower underlying-asset volatility, while, however, only being mildly affected by stochastic volatility and asset-value jumps.

We use exchange-traded single-stock American put options and their equivalent synthetic European options to determine the sign and magnitude of the early exercise risk premium and its relations with factors conditioning the chance of an early exercise in the data. We calculate the monthly American put option return as the compounded early exercise payoff (if there is an optimal early exercise) or the end-of-month option value (if there is none) to the start-of-month option value, comparing market option values with early exercise payoffs to identify optimal early exercises (Barraclough and Whaley's (2012) "market rule"). To calculate the monthly European put option return, we first create synthetic European put options. To achieve that aim, we start from Merton's (1973) insight that it is never optimal to early exercise an American call option written on a zero-dividend asset, allowing us to treat these options as *quasi-European* options. Prompted by Zivney (1991), we next recognize that a European put option can be replicated using a portfolio long the equivalent European call option, long an investment of the discounted strike price into a money market account, and short the underlying asset ("put-call parity"). We finally calculate the monthly synthetic European put option return as the end-of-month value of the replication portfolio to its start-of-month value.

In accordance with our theory, portfolio sorts and Fama-MacBeth (FM; 1973) regressions run on spread portfolios long an American put option and short its equivalent synthetic European option suggest that the early exercise risk premium is generally positive and highly significant in the data. In the pooled data, an equally-weighted portfolio of the above spread portfolios, for example, yields a mean return of 3.66% per month ($t$-statistic: 8.67, greatly above the bootstrap upper limit of the 95% confidence interval of 2.11). The mean return of the portfolio of portfolios, however, significantly increases with a higher option moneyness (calculated as the strike-to-stock price ratio), a shorter time-to-maturity, and a lower idiosyncratic underlying-asset volatility derived from the market or Fama-French-Carhart (FFC) model. Among short (i.e., 30-60 days) time-to-maturity options, the mean return is, for example, a highly significant 11.74% per month ($t$-statistic: 12.04) for in-the-money (ITM) options (moneyness $> 1.05$) but an only insignificant –0.68% ($t$-statistic: –0.90) for out-of-the-money (OTM) options (moneyness $< 0.95$).

Our empirical evidence relies crucially on it never being optimal to early exercise an American call option on a zero-dividend asset and put-call parity holding. Cremers and Weinbaum (2010), Jensen and Pedersen (2016), and Figlewski (2018), however, report that both these conditions can be violated due to stock and options market illiquidity and/or short-sales constraints on the underlying asset. To establish whether such violations confound our early exercise risk premium estimates, we condition our tests on Amihud's (2002) mean absolute return-to-dollar trading volume measure (a stock illiquidity measure), the option bid-ask spread or open interest (two option illiquidity measures), and the "Daily-Cost-to-Borrowing" score (DCBS) from Markit (a short-sale constraints measure). Results suggest that while short-sales constraints do not exert a meaningful effect, controlling for stock and options illiquidity slightly decreases our early exercise risk premium estimates, without, however, rendering them insignificant.

We finally study whether real investors can earn the early exercise risk premium by longing American put options and shorting the equivalent European option replication portfolios. To do so, we assume real investors always buy at the midpoint price plus a fraction of the bid-ask spread, whereas they always sell at the midpoint minus the same fraction of that spread. Our evidence suggests that real investors can earn a significantly positive premium for short time-to-maturity deep-ITM options under reasonably high transaction costs, but not for other options.

Our work adds to an empirical literature studying the spreads in prices between American options and equivalent European options (labelled the "early exercise premium"). Zivney (1991) compares the prices of traded American S&P 500 call or put options with those of equivalent synthetic European options obtained from put-call parity. Closer to us, de Roon and Veld (1996) and Engström and Nordén (2000) conduct the same exercise on U.S. stock index options and Swedish single-stock options for which early exercising the call options is never optimal. Conversely, McMurray and Yadav (2000) compare the prices of traded American and European FTSE 100 options with identical times-to-maturity, but slightly different strike prices. Supporting Merton (1973), all these studies find a significantly positive early exercise premium. In contrast to them, we study the spread in expected returns — and not prices — between American and European options. Given that the ability to early exercise an option affects the expected option payoff *and* the option price, our conclusions do not follow mechanically from theirs. In fact, if the expected return spread between the two types of options were only driven by an effect on the option price, we would reach exactly the opposite conclusions of those reported in our paper.

We further add to a literature studying stock and option characteristics pricing the cross-section of option returns. Coval and Shumway (2001) show that, in a stochastic discount factor model, the expected European call (put) option return decreases (increases) with moneyness, confirming their predictions using S&P 500 option data. Hu and Jacobs (2018) report that, in a Black and Scholes (1973) framework, the same return decreases (increases) with underlying-asset volatility. Re-examing those relations in a stochastic discount factor model, Aretz et al. (2019) establish that Hu and Jacobs' (2018) results only hold for idiosyncratic volatility, showing that the signs of the relations with systematic volatility are ambiguous and depend on moneyness. Goyal and Saretto (2009) report that the delta-hedged call or put option return increases with the realized-minus-implied volatility of the underlying asset. Cao and Han (2013) show that the same return decreases with idiosyncratic underlying-asset volatility. We contribute to these studies by identifying another option characteristic pricing options: the ability to early exercise.

We finally add to studies examining investors' early exercise policies. Overdahl and Martin (1994) show that the majority of early exercises of single-stock call and put options fall within theoretically optimal early exercise boundaries, suggesting rational exercise policies. Conversely, Brennan and Schwartz (1977) find that American put options are often exercised significantly earlier or later than advocated by the Black and Scholes (1973) model. Finucane (1997) shows that investors often early exercise call options written on zero-dividend underlying assets, conflicting with Merton (1973). Extending Finucane's (1997) analysis, Poteshman and Serbin (2003) show that only individual — but not institutional — investors early exercise those options. Pool et al. (2008) estimate that total profits lost from failing to optimally early exercise single-stock call options on ex-dividend dates amount to $491 million over the 1996-2006 period. Barraclough and Whaley (2012) show that total profits lost from failing to optimally early exercise single-stock put options are similarly large. Eickholt et al. (2018) report that suboptimal early exercise policies can be explained using investor irrationality, transaction costs, and a demand for liquidity and financial flexibility. Given the widespread evidence on how blatantly investors violate optimal early exercise policies, it is perhaps surprising that we find that the ability to early exercise is priced in accordance with neoclassical models assuming optimal policies.

We proceed as follows. Section 1.2 studies the early exercise risk premium inherent in standard neoclassical finance models. In Section 1.3, we discuss our data and methodology. Section 1.4 presents our main cross-sectional evidence, while Section 1.5 presents the results from related robustness tests. Section 1.6 offers our time-series evidence. Section 1.7 sums up and concludes.

## 1.2 Theory

In this section, we study the early exercise risk premium in plain-vanilla American put options in neoclassical finance models. To that end, we conduct a Monte Carlo simulation exercise, using the Longstaff and Schwartz (2001) least-squares method together with popular stochastic processes to obtain the expected returns of American and European put options. We finally study the simulated expected return spread between equivalent American and European put options, allowing moneyness, time-to-maturity, and underlying-asset volatility to vary across our simulations.

### 1.2.1 A Monte Carlo Simulation Exercise

#### A. Calculating Simulated Expected Option Returns

We conduct a Monte Carlo simulation exercise to gain some broader insights into the early exercise risk premium. To do so, we use Longstaff and Schwartz's (2001) least-squares approach to compute the expected returns of American and European put options written on a zero-dividend underlying asset, relying on alternative stochastic processes to model the evolution of the asset's value. The alternative stochastic processes studied by us are Geometric Brownian motion (GBM), a stochastic volatility (SV) process, and a stochastic-volatility jump (SVJ) process (see Bates (1996), Andersen et al. (2002), and others). We can compactly write these processes as:

$$dS(t) = \alpha S(t)dt + S(t)\sqrt{V(t)}dW^S(t) + d\left(\sum_{j=1}^{N(t)} S(\tau_{j-})\left[e^{Z_j^s} - 1\right]\right) \tag{1.1}$$

$$- \lambda\bar{\mu}S(t)dt,$$

$$dV(t) = \kappa_v(\theta_v - V(t))dt + \sigma_v\sqrt{V(t)}dW^v(t), \tag{1.2}$$

where $S(t)$ and $V(t)$ are, respectively, the asset value and variance at time $t$, $\alpha$ the asset value drift rate, $\kappa_v$ the variance mean reversion parameter, $\theta_v$ the long-run variance, and $\sigma_v$ the volatility of variance. $W^S(t)$ and $W^v(t)$ are Brownian motions, with $\text{Corr}(W^S(t), W^v(t)) = \rho$. Finally, $N(t)$ is a Poisson process with intensity $\lambda$, $S(\tau_{j-})$ is the asset value one instant before a jump, $Z_j^s \sim N(\mu_z, \sigma_z^2)$, and $\bar{\mu} = e^{\mu_z + \sigma_z^2/2} - 1$. To rule out jumps, the SV model imposes $\lambda = 0$. To ensure that asset-value variance is constant, the GBM model further imposes $\kappa_v = \sigma_v = 0$.

We simulate asset values from Equations (1.1) and (1.2) under both the physical, $\mathcal{P}$, and the risk-neutral, $\mathcal{Q}$, probability measure. As stressed by Cheredito et al. (2007) and Broadie et al. (2009), the asset value drift rate $\alpha$, the variance mean reversion parameter $\kappa_v$, the jump intensity $\lambda$, and the expected jump size $\mu_z$ can all be different under the $\mathcal{P}$ and $\mathcal{Q}$ measures. To indicate that, we let $\alpha \in \{\alpha^P, \alpha^Q = r_f\}$, $\kappa_v \in \{\kappa_v^P, \kappa_v^Q\}$, $\lambda \in \{\lambda^P, \lambda^Q\}$, and $\mu_z \in \{\mu_z^P, \mu_z^Q\}$, where the first entry is the parameter value under the $\mathcal{P}$ measure, the second that under the $\mathcal{Q}$ measure, and $r_f$ the risk-free rate of return. In line with Broadie et al. (2009), we however restrict the mean reversion parameters to be identical across physical and risk-neutral measure (i.e., $\kappa_v^P = \kappa_v^Q$).

Having simulated the evolution of the underlying asset value under the $\mathcal{P}$ and the $\mathcal{Q}$ measure using one of the three stochastic processes multiple times, we next compute the expected European put option return as follows. Separately for each simulated asset-value path and the two measures, we first calculate the maturity payoff of the option, $\max(K - S(T), 0)$, where $K$ is the strike price of the option and $S(T)$ the maturity value of the underlying asset under either measure. We next calculate the expected option payoff as the simple mean of the asset-value-path specific maturity payoffs of the option under the $\mathcal{P}$ measure. Conversely, we calculate the option value as the simple mean of the same payoffs of the option under the $\mathcal{Q}$ measure, discounted to the option initiation date at the risk-free rate. We finally scale the expected option payoff by the option value to obtain the expected gross European put option return over the time-to-maturity.

To compute the expected American put option return, we first use the simulated asset-value paths to delineate the optimal early exercise threshold (i.e., the highest underlying asset value for which an exercise is optimal) over the time-to-maturity. To do so, we compute the maturity payoff of the option under the $\mathcal{Q}$ measure. Moving back from the maturity date to the date directly before ($T - T/n$, where $n$ is the number of time steps), we estimate the value of the option conditional on the underlying asset value using a regression. In particular, we regress the maturity option payoff per path discounted to time $T - T/n$ on a higher-order polynomial of the underlying asset value per path, using, however, only observations for which the option is ITM at time $T - T/n$. We then assume that the option is early exercised if the early exercise payoff, $\max(K - S(T - T/n), 0)$, is above the fitted regression value. Continuing in that way, we always move back one period, regress the discounted earliest (early exercise or maturity) payoff on the same higher-order polynomial, and determine the optimal early exercise policy. Doing so until the option initiation date, we are able to estimate the entire early exercise threshold.

Analogous to our European put option calculations, we calculate the expected payoff of an American put option as the simple mean of the asset-value-path specific payoffs of the option under the $\mathcal{P}$ measure, while we calculate its value as the discounted simple mean of the asset-value-path specific payoffs of the option under the $\mathcal{Q}$ measure. In case of the American option, the payoff is, however, either the earliest early exercise payoff compounded to maturity at the risk-free rate of return (if the underlying asset value drops below the estimated early exercise threshold) or the maturity payoff (if it does not). We finally again scale the expected option payoff by the option value to obtain the expected gross American put option return over the time-to-maturity.

In our simulations, we study an underlying asset whose initial value ($S(0)$) is 100. The basecase drift rate, $\alpha$, and (initial) volatility, $\sqrt{V(0)}$, of the underlying asset are 12% and 20% per annum, respectively. We always set the long-run variance, $\theta_v$, equal to the initial asset-value variance. The basecase option-contract parameters, the strike price ($K$) and time-to-maturity ($T$), are 100 and 60 days, respectively. We assume a risk-free rate of return of 2.5% per annum ($r_f$). We select the basecase parameter values for the other stochastic process parameters in line with the literature (see, e.g., Broadie et al. (2009)). In particular, we set the mean reversion in variance parameters under both measures, $\kappa_v^P$ and $\kappa_v^Q$, to 5.33, the volatility of variance, $\sigma_v$, to 14% per annum, and the correlation between the asset value and asset variance shocks, $\rho$, to –0.52. Turning to the jump parameters, we choose a jump intensity of 0.91 and an expected jump size of –0.0325 under the physical measure ($\lambda^P$ and $\mu_z^P$, respectively) and of 1.51 and –0.0685 under the risk-neutral measure ($\lambda^Q$ and $\mu_z^Q$, respectively). We assume a jump volatility, $\sigma_z$, of 6% under both measures.

Our simulations rely on one million asset-value paths sampled at a daily frequency to calculate expected option returns. In the Longstaff and Schwartz (2001) regressions, we use a third-order polynomial to estimate option values, regressing the discounted earliest option payoff on the underlying-asset value, its squared value, its cubed value, and a constant.

## B. The Early Exercise Risk Premium in Neoclassical Models

Table 1.1 studies the early exercise risk premium inherent in American put options assuming that the underlying asset value follows GBM. To do so, the table presents the expected payoffs, values, and expected returns of equivalent American (columns (1) to (3), respectively) and European (columns (4) to (6), respectively) put options as well as the differences in these statistics across those types of options.[2] Panels A, B, and C

[2]The N/A entries in Panel C of Table 1.1 arise in situations in which the option is never optimally exercised over the simulated underlying-asset-value paths under the $\mathcal{Q}$ measure, yielding a zero option value.

consider ITM (strike-to-stock price ratio = 1.10), at-the-money (ATM; 1.00), and OTM (0.90) options, respectively.[3] Within each moneyness class, we further distinguish between options with a short (30 days), medium (60), and long (120) time-to-maturity. Within each maturity class, we distinguish between options written on assets with a low (10%), medium (20%), and high (30%) annualized idiosyncratic volatility. To ease comparability, we consistently report the monthly expected return, computed by dividing the expected return of the medium-term 60-day (long-term 120-day) options over their entire time-to-maturity by two (four).

Table 1.1 About Here

The table suggests that the early exercise risk premium in put options (i.e., the expected return spread between the equivalent American and European put options shown in the final column) is positive but varies significantly with option and underlying-asset characteristics. To be specific, the premium ranges from a minimum of 0.75% per month (see row 7 in Panel C) to a maximum of 8.14% (see row 4 in Panel A). Variations in the premium can be traced to variations in moneyness, time-to-maturity, and underlying-asset volatility, with the premium strongly increasing with moneyness and more weakly decreasing with time-to-maturity and underlying-asset volatility. Looking at 30 day-to-maturity options written on underlying assets with an annualized volatility of 20%, the early exercise risk premium, for example, rises from 1.17% per month for the OTM option to 5.65% for the ITM option (compare the second rows in Panels A and C). The relations with the option and underlying-asset characteristics originate from the characteristics conditioning the optimal early exercise probability, with that probability also increasing in moneyness but decreasing in time-to-maturity and underlying-asset volatility.[4] Figure 1.1 graphically displays the relations between the early exercise

[3] We consistently vary option moneyness through varying the option's strike price $K$.

[4] While it is easy to grasp why the optimal early exercise probability increases with moneyness, it is perhaps harder to understand why it decreases with both time-to-maturity and underlying-asset volatility. The negative relation with time-to-maturity arises from the optimal early exercise threshold being a monotonically increasing convex function over the time period to the maturity

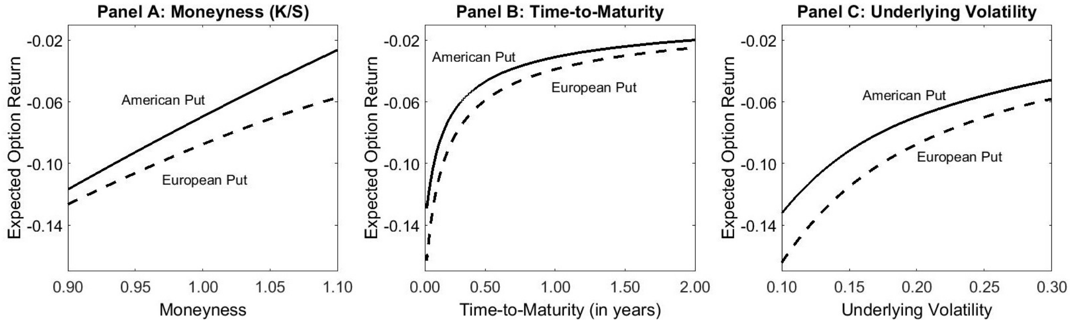


**Figure 1.1: The Early Exercise Risk Premium in a GBM World**

The figure plots the expected returns of American put options (solid line) and equivalent European put options (broken line) against moneyness (Panel A), time-to-maturity (Panel B), and underlying-asset volatility (Panel C) under the assumption that the underlying-asset value follows GBM. Defining moneyness as the ratio of the strike price to the underlying asset value, we induce variations in moneyness through variations in the strike price in Panel A. We describe the basecase parameter values used to produce the figure in the main text.

risk premium and the option and underlying-asset characteristics.

Separately studying the expected returns of the American and European put options underlying the early exercise risk premium, we find that they are both below the risk-free rate of return (which is 0.21% per month) and typically negative, in line with Coval and Shumway's (2001) theoretical work on European put options (see columns (3) and (6)). Turning to the components making up the expected return, which are the expected payoff and value, we find that the American put options consistently have higher values than the equivalent European options (compare columns (2) and (5)). The higher values of the American options agree with Merton's (1973) insight that the value of an American option equals the value of the equivalent European option plus the value of the right to early exercise. Notwithstanding, the American options also have consistently higher expected payoffs than the equivalent European options, suggesting that being able to optimally early exercise an option makes the option more profitable in expectation. Comparing the effects of the ability to optimally early exercise a put option on the option's expected payoff and value,

date (see Shreve (2004)). The probability that the underlying-asset value hits the optimal early exercise threshold over some fixed time period is thus lower the further away an option is from its maturity date. The negative relation with underlying-asset volatility arises since an option is optimally early exercised when the early exercise payoff exceeds the value of the alive option. Raising underlying-asset volatility, only the value of the alive option — but not the early exercise payoff — increases, making it less likely that the early exercise payoff exceeds the value of the alive option and lowering the early exercise probability.

the effect on the expected payoff is *in proportional terms* consistently larger than the effect on value, explaining why the early exercise risk premium is positive.

Table 1.2 studies how stochastic volatility and asset-value jumps affect the early exercise risk premium. To do so, the table contrasts the premium under a GBM underlying-asset value process (repeated in Panel A) with those under a SV (Panel B) or SVJ (Panel C) process. Comparing Panels A and B, we find that stochastic volatility has a mild but ambiguous effect on the early exercise risk premium. Allowing for stochastic volatility, the premium of the 30-day ITM option on a 20% volatility asset, for example, *rises* by 7.6% (from 5.65% to 6.08%), while that of the otherwise identical ATM option *drops* by 3.6% (from 3.31% to 3.19%). The ambiguous stochastic-volatility effect originates from our choice of a negative correlation between asset value and volatility (i.e., $\rho = -0.52$), possibly inflating the left tail of the future asset-value distribution. The fatter left tail, however, has a mixed effect. On one hand, the underlying asset is now able to reach lower values, making an early exercise more likely. On the other, low underlying-asset values now tend to come with high underlying-asset volatility, making an early exercise less likely. Depending on the option and underlying-asset characteristics, either of the channels can dominate, leading the effect of stochastic volatility on the early exercise risk premium to be ambiguous.

<u>TABLE 1.2 ABOUT HERE</u>

A comparison of Panels B and C in Table 1.2 suggests that asset-value jumps typically decrease the early exercise risk premium. The premium of the 60-day ATM option written on a 10% volatility asset, for example, drops by 61% (from 3.96% to 1.54%) upon allowing for jumps in addition to stochastic volatility. The reason is that jumps often contribute to the volatility of the underlying asset's value, making investors more reluctant to early exercise. The negative effect of jumps on the early exercise risk premium does, however, not materialize in situations characterized by a high moneyness, short time-to-maturity, and high underlying-asset volatility (see, e.g., the first entry in the final row in Panel C).

In those situations, jumps significantly affect the underlying asset's payoff, but mostly through them decreasing its expectation without boosting volatility. In turn, option moneyness increases, making investors more likely to early exercise.

Overall, this section offers strong evidence that, under realistic parameter values, neoclassical asset pricing models predict a positive early exercise risk premium, defined as the expected return spread between American and equivalent European put options. The premium increases strongly with moneyness and decreases slightly less strongly with time-to-maturity and underlying-asset volatility. Conversely, it is only mildly affected by stochastic volatility, while the effect of asset-value jumps on it comes mostly through those jumps raising underlying-asset volatility. Due to its strong dependence on several option and underlying-asset characteristics, the premium can easily become sizable, with it reaching a maximum of 8.14% *per month* in our simulations.

## 1.3 Data and Methodology

In this section, we explain how we calculate the returns of equivalent single-stock American and European put options, to be used to study the early exercise risk premium in the data. We first describe our data sources and filters. We next elaborate on our return calculations.

### 1.3.1 Data Sources and Filters

We first introduce our data sources and filters. We obtain daily data on American call and put options written on stocks with zero payouts over the options' maturity time ("zero-dividend stocks"), on the stocks underlying the options, and on the term structure of the risk-free rate of return from Optionmetrics. We source additional market data on the underlying stocks from CRSP, while we source firm-fundamental data on them from Compustat. We retrieve data on stock short-sale constraints from Markit. We finally obtain data on the Fama-French benchmark factors, the VIX index, the TED spread, and a liquidity factor from Kenneth French's website, the

CBOE website, the Fred Database, and Lubos Pastor's website, respectively.[5]

We impose standard filters on our options data (see Goyal and Saretto (2009) and Cao and Han (2013)). To be specific, we exclude option-day observations for which the option price violates standard arbitrage bounds (as, e.g., the bound that an American call option's price must lie between the maximum of zero and the value of the equivalent long forward, and the stock price), lies below \$1, or is less than one-half the option bid-ask spread. We further exclude observations for which the option bid-ask spread is negative or the underlying stock's price is missing.

### 1.3.2 Calculating Single-Stock Option Returns

We next explain how we calculate American and synthetic European put option returns over calendar month $t$. We start with the American options. To calculate the return on such an option, we need to determine whether the option is optimally early exercised over month $t$. To do so, we use Barraclough and Whaley's (2012) "market rule," comparing the option's early exercise payoff at the end of each trading day with its traded price at the same time and assuming that an early exercise occurs if the payoff is equal to or exceeds the price. The upside to using that approach is that the market rule does not depend on an option value estimate obtained from a (perhaps misspecified) model. The downside, however, is that, in the absence of arbitrage opportunities, the early exercise payoff cannot exceed the traded option price, while, in the presence of minimum tick size rules in stock and options markets, it also cannot be exactly equal to that price. To overcome that problem, we assume that an early exercise occurs if the early exercise payoff is within 1% of the traded option price, that is, if it holds that:

$$\frac{P^A(t) - \max(K - S(t), 0)}{\max(K - S(t), 0)} \leq 0.01, \tag{1.3}$$

[5] The URL address of Ken French's website is: <https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/>, while that of Lubos Pastor's website is: <https://faculty.chicagobooth.edu/lubos.pasto r/research/>. The VIX data can be retrieved from: <http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data)>, and the TED spread from: <https://fred.stlouisfed.org/series/TEDRATE>.

where $\max(K - S(t), 0)$ is the early exercise payoff and $P^A(t)$ the option price, both measured at the end of day $t$. While our strategy to identify optimal early exercises may lead us to overlook some of those, that oversight works against us finding a positive early exercise risk premium since optimal early exercises raise the expected option return (see Section 1.2).[6]

Having identified the trading day on whose end an option is optimally early exercised (if any), we calculate the American put option return as the ratio of the early exercise payoff compounded to the end of the month at the risk-free rate (if there is an early exercise) or the option price at the end of the month (if there is none) to the option price at the start of the month. Although a large empirical literature studies the returns of American single-stock options, we are, to the best of our knowledge, first in incorporating optimal early exercises into those calculations.

Turning to calculating the returns of European put options written on single stocks, we face the problem that options exchanges exclusively trade in American (and not European) single-stock options.[7] To circumvent that problem, we synthetically create single-stock European put options by trading in American options, the underlying stock, and the money markets. To do so, we start from Merton's (1973) insight that it is never optimal to early exercise an American call option written on an underlying asset not paying cash. Since we restrict our sample to options written on stocks not paying out dividends over their time-to-maturity, our sample American call options are effectively European call options. We next recognize that a European put option can be replicated by longing the equivalent European call option, shorting the underlying stock, and investing the discounted

[6]Using a threshold level of 2% or 5% in Equation (1.3), we continue to find a highly significant positive early exercise risk premium displaying the same relations with the stock and option characteristics as reported later. In line with our theoretical result that optimally early exercising an option raises the option's expected profitability, we, however, also find that American put option returns tend to increase with the threshold level.

[7]While there are European single-stock options traded over-the-counter (OTC), we neither have data on those nor are their contract terms comparable across one another or with those of exchange-traded options.

strike price in the money market, allowing us to write:

$$P_{i,K,T}^{synE} = C_{i,K,T}^{A} - S_i + Ke^{-r_f T}, \tag{1.4}$$

where $P_{i,K,T}^{synE}$ is the price of a synthetic European put option written on stock $i$ and with strike price $K$ and time-to-maturity $T$, $C_{i,K,T}^{A}$ is the price of the exchange-traded American call option written on the same stock and with the same strike price and time-to-maturity, $S_i$ is stock $i$'s price, and $r_f$ is the risk-free rate of return over the maturity time ("put-call parity").

To ensure that the synthetic European put option prices are comparable to those of the traded American put options, we impose the same data filters on them as on the American options whenever possible. To be specific, we again exclude option-month observations for which the synthetic European put option price violates standard arbitrage bounds (as, e.g., the bound that an European put option's price must lie between the maximum of zero and the value of the equivalent short forward contract, and the strike price) or lies below \$1. Since we do not have bid or ask prices for the synthetic options, we are however unable to impose any of the restrictions based on those. We finally exclude option-month observations for which the synthetic European put option price exceeds the price of the traded American put option.

Having derived synthetic European put option prices, we calculate those options' returns as the ratio of the end-of-month option price to the start-of-month option price.

## 1.4 Cross-Sectional Evidence

In this section, we offer our cross-sectional evidence on the early exercise risk premium and its relations with stock and option characteristics. We first present descriptive statistics on the mean return spread between equivalent American and European put options, an estimate of the early exercise risk premium. We next provide the results from portfolio sorts and FM regressions studying the relations between the premium and stock and option characteristics.

### 1.4.1 The Early Exercise Risk Premium in the Data

In Table 1.3, we present descriptive statistics for the monthly returns of single-stock American and synthetic-European put options (columns (1) to (2), respectively), the difference in their returns (column (1)–(2)), and their moneyness and days-to-maturity (columns (3) to (4), respectively). The option-month observations in columns (1) and (2) are matched along the moneyness and time-to-maturity dimensions, so that each observation in column (1) is associated with exactly one observation in column (2) with identical moneyness and time-to-maturity. The descriptive statistics include the mean, the standard deviation (StDev), the mean's $t$-statistic (Mean/StError), a 95% confidence interval for the $t$-statistic, several percentiles, and the number of observations. With the exception of the $t$-statistic and the confidence interval, the descriptive statistics are calculated by sample month and then averaged over time. As a result, we can interpret the means in columns (1) to (2) as the mean returns of equally-weighted portfolios of the American and European options, respectively. Again defining moneyness as the strike-to-stock price ratio, we measure both moneyness and days-to-maturity at the start of the return month.

TABLE 1.3 ABOUT HERE

Since option returns are non-normally distributed, as, for example, shown in Broadie et al. (2009), it is possible that standard asymptotic inference techniques yield biased conclusions in their case. To guard against that possibility, we follow Vasquez (2017) and use a boostrap to construct the 95% confidence interval in Table 1.3. We do so as follows. Separately for the American and European options and the spread portfolio in columns (1) to (3), respectively, we first impose the null hypothesis of a zero mean on the time-series of cross-sectional average returns. We next draw with replacement and an equal probability of being drawn 244 cross-sectional average returns from that time-series, where 244 is the number of months in our sample period. We then compute a bootstrap $t$-statistic using the drawn returns.

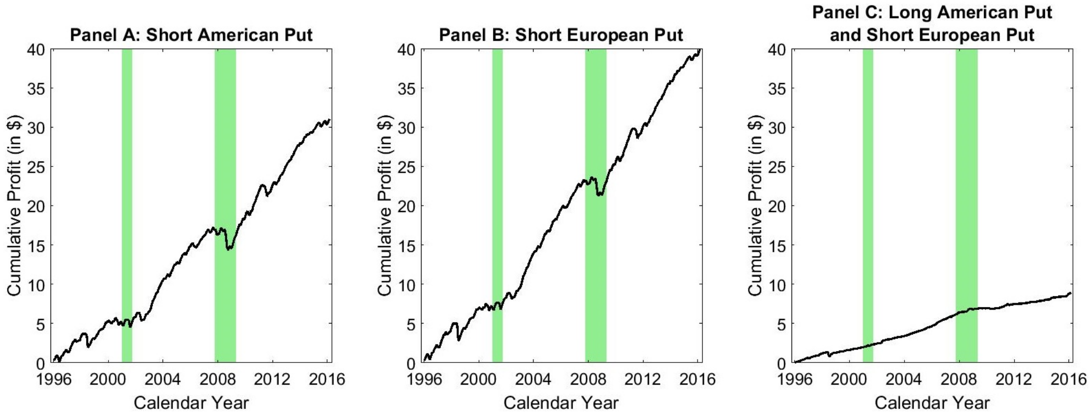


**Figure 1.2: Cumulative Profits From Selling American and European Put Options**

The figure plots the cumulative profits from shorting one dollar of the equally-weighted American (Panel A) or European (Panel B) put option portfolio or from longing one dollar of the American and shorting the same number of dollars of the European put option portfolio (Panel C) at the start of each sample month and holding that position over the month. The green areas are NBER recession periods.

Repeating those steps 20,000 times, we are able to generate a distribution of the bootstrap $t$-statistic. The lower and upper limits of the confidence interval are the $2.5^{th}$ and $97.5^{th}$ percentiles of that distribution, respectively.

Table 1.3 strongly suggests that the early exercise risk premium is positive. While both the American and European put options yield significantly negative mean returns in columns (1) and (2), the mean American option return is a less negative –12.76% per month ($t$-statistic: –5.36) compared to the mean European option return of –16.42% ($t$-statistic: –7.06). Thus, the mean spread return across the American and European options is 3.66% per month ($t$-statistic: 8.67) in column (3). The $t$-statistic of the mean spread return is not only remarkable because it greatly exceeds the upper limit of the bootstrap confidence interval of 2.11, but perhaps more so because it exceeds the $t$-statistics of the underlying options' mean returns in absolute terms (compare columns (1) to (3)). The reason for the high $t$-statistic is that the spread return is far less volatile than the underlying options' returns, as can be seen from its standard deviation and percentiles. The moneyness and days-to-maturity statistics in columns (3) and (4) suggest that the average option pair is close to ATM and has slightly more than two months to maturity.

In Figure 1.2, we contrast the returns of our American and European put options over our sample period. To do so, we plot the cumulative profits from shorting an equally-weighted portfolio of the American (Panel A) or the European (Panel B) options or from longing the equally-weighted American and shorting the equally-weighted European option portfolio (Panel C) over each sample month. The green shaded areas are NBER recession periods. Panels A and B confirm that shorting put options is profitable except in recessions in which most options are exercised against their sellers (see the 2000-2001 and 2007-2008 periods). Notwithstanding, Panel C suggests that shorting European put options is more profitable than shorting American put options in almost every month. Interestingly, however, the long-short strategy is far more profitable over the earlier sample period, with it earning close to 70% of its full-sample-period cumulative profits until 2008 (about half of our sample period). We will return to this observation later on.

### 1.4.2 Relations with Stock and Option Characteristics

Recalling that our theoretical work in Section B. relates the early exercise risk premium to several stock and option characteristics, we next use portfolio sorts and FM regressions to investigate whether we can find these relations in the data. We start with portfolio sorts based on moneyness and time-to-maturity. At the end of each sample month $t-1$, we thus split the universe of American-European put-option pairs into independently double-sorted portfolios according to these sorting variables. To be specific, we first sort the option pairs into an ITM (strike-to-stock price above 1.05), an ATM (0.95 to 1.05), and an OTM (below 0.95) portfolio. We then split them into three portfolios according to whether their maturity time is below 60, between 60 and 90, or above 90 days. The intersection yields the double-sorted portfolios. We equally-weight the double-sorted portfolio constituents and hold them over month $t$.

Table 1.4 shows the results from the double-sorted portfolio exercise based on moneyness and time-to-maturity. The plain numbers are the mean monthly returns of equally-weighted portfolios of American (column (1)) or European (column (2)) options or spread portfolios long an American and short the equivalent European option (column (1)–(2)). The numbers in square parentheses are $t$-statistics calculated from Newey and West's (1987) formula with a twelve-month lag length, while those in curly parentheses are 95% bootstrap $t$-statistic confidence bounds. Panels A, B, and C focus on the ITM, ATM, and OTM options, while the upper, middle, and lower rows in each panel focus on the below 60, 60 to 90, and above 90 day options, all respectively.

Table 1.4 About Here

Supporting our theory, the table offers strong evidence that the early exercise risk premium relates positively to moneyness and negatively to time-to-maturity. The mean spread return of 30-60 day options, for example, rises from an insignificant –0.68% per month for OTM options in Panel C to a highly significant 11.74% ($t$-statistic: 12.04, greatly above the bootstrap upper confidence bound) for ITM options in Panel A. Conversely, the mean spread return of ATM options rises from a mildly significant but economically small –0.31% ($t$-statistic: –2.16) for 90-120 day options to 4.55% ($t$-statistics: 7.20, also above the upper confidence bound) for 30-60 day options in Panel B. Turning to the underlying American and European put options, the mean returns of both types of options become more negative the closer an option is to maturity, broadly consistent with our theoretical work (compare with the simulated expected option returns in Table 1.1). Deviating from both Coval and Shumway's (2001) and our theoretical work, the same returns, however, also become sharply more negative with moneyness, which is our single empirical result entirely inconsistent with neoclassical option pricing theory.[8]

We next conduct portfolio sorts based on the idiosyncratic volatility of the underlying stock. At the end of each sample month $t-1$, we thus split the universe of option pairs into quintile portfolios according to idiosyncratic volatility estimates

[8] The negative relation between mean put option return and moneyness is all the more remarkable since Coval and Shumway (2001) prove that, in the absence of arbitrage opportunities, the only condition necessary to produce a positive relation between expected European put option return and strike price is that the underlying asset's value and the stochastic discount factor are negatively correlated, hardly a heroic assumption.

obtained from either the market model or Fama-French-Carhart (FFC; 1997) model. We can write the market model as:

$$R_{i,\tau} = \alpha_i + \beta_i^{mkt}(R_\tau^{mkt} - R_{f\tau}) + \epsilon_{i,\tau}, \tag{1.5}$$

where $R_{i,\tau}$ is stock $i$'s return over month $\tau$, $R_\tau^{mkt} - R_{f\tau}$ is the excess market return, $\alpha_i$ and $\beta_i^{mkt}$ are parameters, and $\epsilon_{i,\tau}$ is the residual. We can write the FFC model as:

$$R_{i,\tau} = \alpha_i + \beta_i^{mkt}(R_\tau^{mkt} - R_{f\tau}) + \beta_i^{smb}R_\tau^{smb} + \beta_i^{hml}R_\tau^{hml} + \beta_i^{mom}R_\tau^{mom} + \epsilon_{i,\tau}, \tag{1.6}$$

where $R_\tau^{smb}$, $R_\tau^{hml}$, and $R_\tau^{mom}$ are the returns of spread portfolios on size, the book-to-market ratio, and the eleven-month (momentum) past return, respectively, and $\beta_i^{smb}$, $\beta_i^{hml}$, and $\beta_i^{mom}$ are additional parameters. We estimate both models over the prior 60 months of monthly data, calculating idiosyncratic volatility as the standard deviation of the residual, $\epsilon_{i,\tau}$. The bottom quintile contains option pairs written on low idiosyncratic volatility stocks ("Low"), while the top contains option pairs written on high idiosyncratic volatility stocks ("High"). We also form a spread portfolio long the top quintile and short the bottom quintile ("H–L"). We equally-weight the quintile portfolios and the spread portfolios and hold them over month $t$.

Table 1.5 presents the results from the univariate portfolio exercise based on idiosyncratic stock volatility, using a table design identical to that of Table 1.4. In Panels A and B, we sort into portfolios based on market- and FFC-model volatility estimates, respectively. Again consistent with our theoretical work, the table offers strong evidence that the early exercise risk premium decreases with idiosyncratic stock volatility. Using market model estimates, Panel A, for example, suggests that the mean spread return between American and equivalent European put options drops from 5.65% ($t$-statistic: 6.60) for options written on low-volatility stocks to 1.26% ($t$-statistic: 3.56) for options written on high-volatility stocks. The difference is a highly significant –4.39% ($t$-statistic: –6.37, greatly below the lower bootstrap confidence bound of –2.31). Looking at the underlying American and European options, their mean returns become significantly less negative with stock volatility, in agreement with our theory results (compare with Table 1.1).

Table 1.5 About Here

While the portfolio sorts above effectively slice our data along the cross-sectional dimension, we next also slice the data along the time-series dimension, repeating our double portfolio sorts on moneyness and time-to-maturity in Table 1.4 separately for the subsample periods until start-2008 and from start-2008. We do so because Barraclough and Whaley (2012) argue that the incentive to early exercise a put option becomes more pronounced the higher the risk-free rate of return at which the early exercise proceeds can be invested. The huge drop in the risk-free rate over our sample period, from an average of 3.76% until start-2008 to an average of 0.05% after, thus enables us to conduct a quasi-natural experiment of how an exogenous shock to the incentive to early exercise a put option affects that option's early exercise risk premium.

Table 1.6 presents the results from the subperiod tests, repeating the double-sorted portfolio exercise in Table 1.4 separately by subperiod. While the design of the table is similar to that of Table 1.4, the table only reports mean spread returns. In complete agreement with theory, the table suggests that the early exercise risk premium is generally higher over the earlier (high risk-free rate) than later (low risk-free rate) period, with the effect, however, only being economically and statistically significant for options with a meaningful early exercise probability (e.g., high-moneyness options with a short time-to-maturity; "treated options"). The mean spread return of 30-60 day ITM options in the first row of Panel A, for example, drops from 13.36% ($t$-statistics: 12.49) over the earlier period to 8.88% ($t$-statistic: 5.89) over the later. The difference is a highly significant –4.49% ($t$-statistic: –3.26, greatly below the lower bootstrap confidence bound). Conversely, the mean spread return of 90-120 day OTM options in the last row of Panel C drops from an economically small –0.60% to an economically small –0.96% from the earlier to the later subperiod.

Table 1.6 About Here

Table 1.7 switches to FM regressions to find out whether our portfolio sort results are robust to variations in methodology. In those regressions, we project the return spread between an American and the equivalent European put option (Panel A), the American option's return (Panel B), or the European option's return (Panel C) over month $t$ on moneyness, time-to-maturity (as fraction of a year), and annualized FFC idiosyncratic stock volatility measured at or until the start of that month. Plain numbers are monthly premium estimates, the numbers in square parentheses are $t$-statistics obtained from Newey and West's (1987) formula with a twelve-month lag length, and the numbers in curly parentheses are 95% bootstrap $t$-statistic confidence bounds.

We construct the bootstrap confidence bounds as follow. Separately for each regression model and independent variable, we first estimate all cross-sectional regressions and then impose the null hypothesis by recreating the dependent variable through adding the fitted regression value excluding the summand involving the independent variable of interest and the residual. We next resample each cross-section, drawing with replacement and an equal probability of being drawn a number of observations for the recreated dependent variable and its related independent variables equal to that in the original cross-section. We finally run the FM estimation on the resampled data, yielding a bootstrap $t$-statistic for the estimate of the independent variable of interest. Repeating those steps 1,000 times, we are again able to generate a bootstrap $t$-statistic distribution from which we can calculate the lower and upper limits of the confidence interval.

Table 1.7 About Here

The FM regressions in Table 1.7 yield results completely consistent with those obtained from the portfolio sorts. In particular, while column (1) in Panel A reproduces the mean spread return from column (1)–(2) in Table 1.4, columns (2) to (4) confirm that the spread return is significantly positively related to moneyness and significantly negatively to time-to-maturity and idiosyncratic stock volatility. Looking at the most comprehensive model in column (4), the mean spread return attracts a moneyness coefficient of 0.348 ($t$-statistic: 19.17), a time-to-maturity coefficient of –0.399 ($t$-statistic: –12.57), and a volatility coefficient of –0.052 ($t$-statistic: –5.50), with all $t$-statistics lying greatly outside of their bootstrap confidence bounds. Conversely, Panels B and C confirm the relations between the returns on either the American or European put options, respectively, and the stock and option characteristics established in the portfolio sorts.

Taken together, this section offers empirical evidence suggesting a positive and economically meaningful early exercise risk premium, as predicted by theory. Further in accordance with theory, it suggests that the premium rises with moneyness and falls with both time-to-maturity and underlying stock volatility, and that it is more pronounced in high rather than low interest-rate regimes. The section finally suggests that the expected returns of both American and European put options have the theoretically-anticipated relations with both time-to-maturity and underlying stock volatility, but that their relations with moneyness significantly deviate from theory.

## 1.5 Robustness Tests

In this section, we offer robustness test results. Since our evidence crucially hinges on it never being optimal to early exercise American call options on zero-dividend stocks and on put-call parity, we first rerun our tests on options for which these conditions are more likely to be fulfilled. Given that our evidence also relies on investors being able to correctly anticipate which stocks do not pay out dividends over an option's time-to-maturity, we next rerun our tests on options written on stocks with a virtually zero probability of doing so. We finally implement tests incorporating bid-ask transaction costs to find out whether real investors benefit from our results.

### 1.5.1 Violations of Early Exercise Rules and Put-Call Parity

In our empirical work, we assume that it is never optimal to early exercise American call options written on zero-dividend stocks, allowing us to use these *quasi-European* options in combination with put-call parity to synthetically create European put options. A problem with that approach is that Jensen and Pedersen (2016) and Figlewski (2018) show that short-selling constraints in stock markets and transaction (i.e., bid-ask spread) costs in stock and options markets can lift the early exercise payoff of an American call option written on a zero-dividend stock above the option's value, casting doubt on our assumption that such American options are always equivalent to European options. Conversely, Cremers and Weinbaum (2010) show that, under exactly the same conditions, put-call parity can break down, casting doubt on whether we are always able to convert European call option prices into meaningful European put option prices.

To study whether violations of the rule to never early exercise an American call option written on a zero-dividend stock or of put-call parity lead us to produce biased results, we condition our tests on popular proxies used to measure stock short-selling constraints and stock and option illiquidity in prior studies. We start with looking at the effect of stock short-selling constraints on our results. To measure such constraints, we use the Daily-Cost-to-Borrowing score (DCBS), as also employed in Jensen and Pedersen (2016). The DCBS value is an integer ranging from one to ten, with a higher value indicating greater stock short-selling constraints and stocks with a score below five generally considered as easy to short. Interesting for our purposes, Jensen and Pedersen (2016) report that far less than one percent of all deep-ITM American call options written on zero-dividend stocks with a DCBS value below five are early exercised, whereas almost ten percent of those same options on stocks with a DCBS value of ten are early exercised.

Table 1.8 reports the results from double portfolio sorts on moneyness and time-to-maturity conditional on the DCBS value at the start of the return period. While the table's design is similar to that of Table 1.4, the table only reports mean spread returns, indicating those lying outside of the bootstrap 95% confidence interval with an asterisk without showing the interval. Columns (1) to (4) focus on options written on stocks with a non-missing DCBS value and a value below eight, seven, and five, respectively. Since the DCBS value is only widely available from start-2004, we run the tests conditional on it starting from that date. The table suggests that, if anything, stock short-selling constraints work against us finding a positive early exercise risk premium, with the mean spread return increasing as we progressively exclude options on hard-to-short stocks. Focusing on 60-90 day ATM options in the middle row of Panel B, the mean spread return, for example, increases from 0.64% per month ($t$-statistic: 2.68) to 1.25% ($t$-statistic: 4.51) going from options on stocks with a non-missing DCBS value to those on stocks with a value below five.

TABLE 1.8 ABOUT HERE

We next study the effect of stock and options illiquidity on top of that of stock short-selling constraints on our estimates of the early exercise risk premium. To do so, we follow Amihud (2002) and measure stock illiquidity as the absolute daily stock return scaled by daily dollar trading volume averaged over the twelve months prior to the option return period. In line with Cao and Han (2013) and Christoffersen et al. (2018), we measure option illiquidity as either an option's bid-ask spread scaled by its price or the inverse of its open interest scaled by the underlying stock's dollar trading volume, both measured at the start of the option return period. Using only options written on stocks with a DCBS value below five at the start of the same period, we next sort our option pairs into three sets of univariate portfolios, the first based on the median of the American put option's illiquidity measure, the second based on the median of the American call option's illiquidity measure, and the third based on the median of the stock's illiquidity measure. The intersection of the three sets of univariate portfolios gives us independently triple-sorted portfolios.

Table 1.9 reports the results from double portfolio sorts on moneyness and time-to-maturity separately run on the option pairs in the top American put option, top

American call option, and top stock illiquidity portfolios (H-H-H; "high-illiquidity assets") or those in the corresponding bottom portfolios (L-L-L; "low-illiquidity assets"). While columns (1) to (2) use the bid-ask spread to measure option illiquidity and columns (3) to (4) option open interest, the table's design is else similar to that of Table 1.8. The table suggests that, in contrast to stock short-selling constraints, stock and options illiquidity can upward bias our early exercise risk premium estimates, with the mean spread return generated from high-illiquidity assets often exceeding that generated from low-illiquidity assets. While the mean spread return of 30-60 day ITM options is, for example, 15.01% per month ($t$-statistic: 8.98) among high-illiquidity assets when using bid-ask spreads to measure option illiquidity, it is 9.10% ($t$-statistic: 6.15) among low-illiquidity assets in that case (see the first row of Panel A). Notwithstanding, the mean spread returns of high moneyness and short time-to-maturity options are significantly positive among *both* high and low-illiquidity assets, suggesting stock and options market illiquidity does not invalidate our conclusions.

Table 1.9 About Here

### 1.5.2 Identification of Zero-Dividend Stocks

Another concern with our empirical strategy is that it requires real investors to be able to correctly identify underlying stocks not paying out dividends over an option's maturity time at the start of the option return period. The reason is that only American call options written on such stocks are equivalent to European call options. While the observations that the vast majority of firms pay out regular dividends at exactly the same points within a calendar year, that extraordinary dividends are rare, and that dividends tend to be announced 3-4 weeks in advance make it likely that investors are able to do so, we nonetheless run a robustness test to verify that assumption. In that test, we repeat the double-sorted portfolio exercise in Table 1.4 using only options written on stocks that never paid out a dividend until the start of the option return period. The idea is that it is extremely unlikely that the stocks start doing so over the option's maturity time.[9]

Table 1.10 gives the results from the double portfolio sorts on moneyness and time-to-maturity run using only options written on stocks that never paid out a dividend. The table's design is similar to that of Table 1.4. The table suggests that while the options written on stocks that never paid out a dividend tend to produce slightly less positive mean spread returns than the full sample, they interestingly also tend to produce equal or even higher inference levels (compare Tables 1.4 and 1.10). The mean spread return of the 30-60 day ATM options written on consistently-zero-dividend stocks is, for example, 3.69% per month, which is slightly lower than the full sample estimate of 4.55% (see the first row of Panel B in both tables). Despite that, the corresponding $t$-statistic is 9.00, which is slightly higher than the equivalent full-sample $t$-statistic of 7.20. Most importantly, however, the mean spread returns of high moneyness and short time-to-maturity options are significantly positive even among options written on consistently-zero-dividend stocks, suggesting that investors' ability to anticipate dividends does not drive our conclusions.

TABLE 1.10 ABOUT HERE

### 1.5.3 Bid-Ask Transaction Costs

We finally study whether real investors are able to earn the positive early exercise risk premium suggested by our tests. Studying that question is interesting since Goyal and Saretto (2009) and Cao and Han (2013) show that high bid-ask transaction costs in options markets greatly eat into the profits of the option trading strategies advocated by them, often rendering those profits insignificant. To find out whether bid-ask transaction costs in stock and options markets also render our estimate of the early exercise risk premium insignificant, we assume that investors consistently buy at the quoted stock and option price plus $S$ times the bid-ask spread and sell at the same price minus $S$ times the bid-ask spread, where $S$ is zero, 0.10, 0.25, or 0.50. When $S$ is

[9] Since CRSP is a more comprehensive stock market data source than Optionmetrics, we use CRSP data to identify stocks that never paid out a dividend to the current month.

equal to 0.50, investors consistently buy at the ask price and sell at the bid price.

Table 1.11 gives the results from the double portfolio sorts on moneyness and time-to-maturity incorporating bid-ask transaction costs, run on either all option pairs in Panel A or only those relying on low illiquidity assets and written on stocks with a DCBS value equal to or below five in Panel B (i.e., the option pairs in the L-L-L portfolio in Table 1.9). In columns (1) to (4), we report mean spread returns calculated from $S$ equal to zero, 0.10, 0.25, and 0.50, respectively. Panel A suggests that accounting for bid-ask transaction costs greatly reduces our early exercise risk premium estimates, which is perhaps unsurprising since setting up the spread portfolio involves us trading in two options and one underlying stock. Notwithstanding, the first row of the panel reveals that the mean return of the deepest ITM and shortest time-to-maturity spread portfolio remains significantly positive up until $S$ equal to 0.25, with it being 4.45% per month ($t$-statistic: 4.66) at that value. In case of all other types of options, the mean spread returns, however, become insignificant or even significantly negative already at $S$ equal to 0.10.

<u>TABLE 1.11 ABOUT HERE</u>

Importantly, Panel B suggests that we can improve the mean spread portfolio returns net of bid-ask transaction costs by using only highly liquid assets and easy-to-short stocks in the portfolio sorts. Doing so, the mean return of the 30-60 day-to-maturity ITM portfolio becomes, for example, significantly positive for all $S$ values, while those of the 60-90 day-to-maturity ITM and the 30-60 day-to-maturity ATM portfolios become significantly positive for $S$ equal to 0.10.

## 1.6 Time-Series Evidence

In Table 1.12, we finish our empirical analysis with time-series regressions of the spread in returns between the equally-weighted American and the equally-weighted synthetic European put option portfolio on several sets of factors. In column (1), we regress on only the excess market return and a constant. Column (2) adds Fama and French's (1993) benchmark factors SMB and HML, while column (3) also adds Carhart's (1997) MOM factor. Column (4) adds Fama and French's (2015) additional benchmark factors PRF and INV. Column (5) finally adds the change in the VIX index, the TED funding spread, and Pastor and Stambaugh's (2003) liquidity factor.[10] Plain numbers are monthly premium estimates, while numbers in parentheses are $t$-statistics calculated from Newey and West's (1987) formula with a twelve-month lag length. An asterisk indicates that the $t$-statistic of an estimate lies outside of its bootstrap 95% confidence bounds.

Table 1.12 About Here

We construct the bootstrap confidence bounds as follows. Separately for each regression model and independent variable, we first estimate the time-series regression and then impose the null hypothesis by recreating the dependent variable through adding the fitted regression value excluding the summand involving the independent variable of interest and the residual. We next draw with replacement and an equal probability of being drawn 244 observations for the recreated dependent variable and its related independent variables. We finally run the time-series regression on the resampled data, yielding a bootstrap $t$-statistic for the estimate of the independent variable of interest. Repeating those steps 1,000 times, we generate a bootstrap $t$-statistic distribution from which we can calculate the lower and upper limits of the confidence interval.

[10]The SMB factor is the return of a portfolio long small and short big stocks, controlling for book-to-market, whereas the HML factor is the return of a portfolio long high book-to-market ("value") and short low book-to-market ("growth") stocks, while controlling for size. The MOM factor is the return of a portfolio long stocks with high returns over the prior twelve months and short stocks with low returns over that period. The PRF factor is the return of a portfolio long more profitable and short less profitable stocks, while INV is the return of a portfolio long low-investment and short high-investment stocks, with both factors controlling for size. See Kenneth French's website for more details. The VIX index is a portfolio of options mimicking option-implied volatility, the TED spread is the difference between the interest rate on short-term U.S. government debt and the interest rate on interbank loans, and the systematic liquidity factor is the return of a portfolio long stocks with a high liquidity exposure and short stocks with a low exposure. See Lubos Pastor's website for more details.

Recalling that the time-series constant can be interpreted as an asset's alpha in models featuring only tradable factors (Black et al. (1972)), Table 1.12 suggests that the spread portfolio's alpha is consistently positive even when controlling for the factors. Looking at the most comprehensive model featuring all factors, column (5), for example, suggests that the alpha is a significant 3.77% per month ($t$-statistic: 9.01). Also interestingly, columns (1) to (4) suggest that, of the Fama-French and Carhart benchmark factors, the spread portfolio loads only significantly on the excess market return, with the coefficient implying that, in accordance with theory, the American options have higher (i.e., less negative) market betas than the European options. The higher market betas, however, come from a lower exposure to volatility risk, as shown in column (5). To be more specific, also controlling for the change in the VIX index, the market beta becomes insignificant, while the VIX beta is a significant –0.05 ($t$-statistic: –2.15). Conversely, neither of the other additional factors in column (5), the TED spread or the liquidity spread portfolio, are significant.

All in all, this section suggests that the American-European put-option spread return is not spanned by other well-known pricing factors and thus represents a factor in its own right.

## 1.7 Concluding Remarks

Spurred by the widely-held belief that American and European option returns are similar, we offer a theoretical and empirical analysis of the early exercise risk premium embedded in plain-vanilla put options, contrasting the expected returns of American put options with those of equivalent European put options. On the theoretical front, we show that standard neoclassical asset pricing models (including those with stochastic volatility and asset-value jumps) can produce a sizable premium, with that premium strongly varying with moneyness, time-to-maturity, and underlying stock volatility. On the empirical front, we compare the returns of single-stock American put options with those of equivalent synthetic European put options, formed from applying put-call parity to American call options written on zero-dividend stocks. The comparisons suggest a significantly positive early exercise risk premium, which increases with moneyness, decreases with time-to-maturity and underlying stock volatility, and is higher in high rather than low interest-rate regimes. Further tests reveal that our empirical conclusions are not driven by stock short-sale constraints, stock or options illiquidity, or the ability to identify zero-dividend stocks. Time-series regressions finally show that the spread return between American and European put options is not spanned by well-known pricing factors, as, for example, the Fama-French factors.

## References


Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5(1), pp.31-56.

Andersen, T.G., Benzoni, L., and Lund, J., 2002. An empirical investigation of continuous-time equity return models. *The Journal of Finance*, 57(3), 1239-1284.

Aretz, K., Lin, M.T., and Poon, S.H., 2019. Moneyness, underlying-asset volatility, and the cross-section of European option returns. *Working Paper*.

Bakshi, G., Kapadia, N., and Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1), pp.101-143.

Barraclough, K., and Whaley, R.E., 2012. Early exercise of put options on stocks. *The Journal of Finance*, 67(4), pp.1423-1456.

Bates, D.S., 1996. Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *The Review of Financial Studies*, 9(1), pp.69-107.

Black, F., Jensen, M.C., and Scholes, M.S., 1972. The capital asset pricing model: Some empirical tests. *Studies in the Theory of Capital Markets*, New York: Praeger, pp.79-121.

Black, F., and Scholes, M., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), pp.637-654.

Brennan, M.J., and Schwartz, E.S., 1977. The valuation of American put options. *The Journal of Finance*, 32(2), pp.449-462.

Broadie, M., Chernov, M., and Johannes, M., 2007. Model specification and risk premia: Evidence from futures options. *The Journal of Finance*, 62(3), pp.1453-1490.

Broadie, M., Chernov, M., and Johannes, M., 2009. Understanding index option returns. *The Review of Financial Studies*, 22(11), pp.4493-4529.

Cao, J., and Han, B., 2013. Cross section of option returns and idiosyncratic stock volatility. *Journal of Financial Economics*, 108(1), pp.231-249.

Carhart, M.M., 1997. On persistence in mutual fund performance. *The Journal of Finance*, 52(1), pp.57-82.

Carr, P., and Wu, L., 2008. Variance risk premiums. *The Review of Financial Studies*, 22(3), pp.1311-1341.

Cheridito, P., Filipović, D., and Kimmel, R.L., 2007. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics* 83(1), pp. 123-170.

Christoffersen, P., Goyenko, R., Jacobs, K., and Karoui, M., 2018. Illiquidity premia in the equity options market. *The Review of Financial Studies*, 31(3), pp.811-851.

Coval, J.D., and Shumway, T., 2001. Expected option returns. *The Journal of Finance*, 56(3), pp.983-1009.

Cox, J.C., Ross, S.A., and Rubinstein, M., 1979. Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3), pp.229-263.

Cremers, M., and Weinbaum, D., 2010. Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis*, 45(2), pp.335-367.

de Roon, F., and Veld, C., 1996. Put-call parities and the value of early exercise for put options on a performance index. *Journal of Futures Markets*, 16(1), pp.71-80.

Eickholt, M., Entrop, O., and Wilkens, M., 2018. What makes individual investors exercise early? Empirical evidence from non-tradable fixed-income products. *Journal of Banking & Finance*, 97, pp.318-334.

Engström, M., and Nordén, L., 2000. The early exercise premium in American put option prices. *Journal of Multinational Financial Management*, 10(3), pp.461-479.

Fama, E.F., and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), pp.3-56.

Fama, E.F., and French, K.R., 2015. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), pp.1-22.

Fama, E.F., and MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), pp.607-636.

Figlewski, S., 2018. An American call is worth more than a European call: The value of American exercise when the market is not perfectly liquid. *Working Paper*.

Finucane, T.J., 1997. An empirical analysis of common stock call exercise: A note. *Journal of Banking & Finance*, 21(4), pp.563-571.

Goyal, A., and Saretto, A., 2009. Cross-section of option returns and volatility. *Journal of Financial Economics*, 94(2), pp.310-326.

Hu, G., and Jacobs, K., 2018. Volatility and expected option returns. *Working Paper*.

Jensen, M. V., and Pedersen, L. H., 2016. Early option exercise: Never say never. *Journal of Financial Economics*, 121(2), pp.278-299.

Longstaff, F.A., and Schwartz, E.S., 2001. Valuing American options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 14(1), pp.113-147.

Martin, I. W., and Wagner, C., 2019. What is the expected return on a stock? *The Journal of Finance*, 74(4), pp.1887-1929.

McMurray, L., and Yadav, P.K., 2000. The early exercise premium in American option prices: Direct empirical evidence. *Working Paper*.

Merton, R.C., 1973. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1), pp.141-183.

Newey, W.K., and West, K.D., 1987. A simple positive-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.

Overdahl, J.A., and Martin, P.G., 1994. The exercise of equity options: Theory and empirical tests. *The Journal of Derivatives*, 2(1), pp.38-51.

Pástor, L., and Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3), pp.642-685.

Pool, V.K., Stoll, H.R., and Whaley, R.E., 2008. Failure to exercise call options: An anomaly and a trading game. *Journal of Financial Markets*, 11(1), pp.1-35.

Poteshman, A.M., and Serbin, V., 2003. Clearly irrational financial market behavior: Evidence from the early exercise of exchange traded stock options. *The Journal of Finance*, 58(1), pp.37-70.

Shreve, S.E., 2004. *Stochastic Calculus for Finance II: Continuous-Time Models* (Vol. 11). Springer Science & Business Media.

Vasquez, A., 2017. Equity volatility term structures and the cross section of option returns. *Journal of Financial and Quantitative Analysis*, 52(6), pp.2727-2754.

Zivney, T.L., 1991. The value of early exercise in option prices: An empirical investigation. *Journal of Financial and Quantitative Analysis*, 26(1), pp.129-138.

## Table 1.1: The Early Exercise Risk Premium in a GBM World

The table presents the expected payoffs, market values, and monthly expected returns of American (columns (1) to (3), respectively) and equivalent European (columns (4) to (6), respectively) put options plus the differences in these (remaining columns) across the two types of options. Panels A to C consider in-the-money (strike-to-stock price=1.10), at-the-money (1.00), and out-of-the-money (0.90) options. Within each moneyness class, we consider options with 30, 60, and 120 days-to-maturity. Within each maturity class, we finally consider options written on an underlying asset with an annualized volatility of 10%, 20%, and 30%. We calculate the options' expected payoffs and market values using Longstaff and Schwartz's (2001) least-squares method applied to simulated underlying-asset-value paths obtained from a Geometric Brownian motion (GBM) process under either the physical $\mathcal{P}$ (expected option payoff) or risk-neutral $\mathcal{Q}$ (option market value) probability measure. In case of each statistic, we rely on 1,000,000 underlying-asset-value paths featuring a number of time steps equal to an option's days-to-maturity. The monthly expected option return is the expected option payoff divided by its market value minus one, scaled by the months-to-maturity. We describe the basecase parameter values in Section A.

| | | American Put | | | European Put | | | American–European | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Days to Mat. | Vol. (%) | Exp. Pay-off | Market Value | Exp. Ret. (%) | Exp. Pay-off | Market Value | Exp. Ret. (%) | Exp. Pay-off | Market Value | Exp. Ret. (%) |
| | | (1) | (2) | (3) | (4) | (5) | (6) | (1)–(4) | (2)–(5) | (3)–(6) |
| Panel A: In-The-Money (Strike-to-Stock Price = 1.10) | | | | | | | | | | |
| 30 | 10 | 10.01 | 10.00 | 0.16 | 9.00 | 9.78 | −7.93 | 1.01 | 0.22 | 8.09 |
| | 20 | 9.87 | 10.03 | −1.67 | 9.19 | 9.91 | −7.31 | 0.68 | 0.12 | 5.65 |
| | 30 | 10.15 | 10.45 | −2.81 | 9.78 | 10.43 | −6.21 | 0.37 | 0.02 | 3.40 |
| 60 | 10 | 10.03 | 10.00 | 0.16 | 8.04 | 9.57 | −7.98 | 1.99 | 0.43 | 8.14 |
| | 20 | 9.76 | 10.24 | −2.36 | 8.82 | 10.12 | −6.42 | 0.94 | 0.13 | 4.06 |
| | 30 | 10.66 | 11.31 | −2.86 | 10.12 | 11.24 | −5.01 | 0.55 | 0.07 | 2.15 |
| 120 | 10 | 10.06 | 9.99 | 0.16 | 6.48 | 9.27 | −7.52 | 3.58 | 0.73 | 7.69 |
| | 20 | 9.70 | 10.87 | −2.70 | 8.44 | 10.66 | −5.20 | 1.25 | 0.21 | 2.50 |
| | 30 | 11.53 | 12.79 | −2.46 | 10.72 | 12.65 | −3.81 | 0.81 | 0.14 | 1.35 |
| Panel B: At-The-Money (Strike-to-Stock Price = 1.00) | | | | | | | | | | |
| 30 | 10 | 0.80 | 1.06 | −24.59 | 0.72 | 1.05 | −31.00 | 0.08 | 0.01 | 6.41 |
| | 20 | 1.92 | 2.20 | −12.65 | 1.85 | 2.20 | −15.96 | 0.07 | 0.00 | 3.31 |
| | 30 | 3.06 | 3.34 | −8.38 | 2.99 | 3.35 | −10.60 | 0.07 | 0.00 | 2.21 |
| 60 | 10 | 0.97 | 1.46 | −16.83 | 0.83 | 1.43 | −20.94 | 0.14 | 0.03 | 4.11 |
| | 20 | 2.53 | 3.06 | −8.72 | 2.38 | 3.05 | −10.96 | 0.15 | 0.02 | 2.23 |
| | 30 | 4.14 | 4.68 | −5.76 | 3.99 | 4.67 | −7.27 | 0.15 | 0.01 | 1.51 |
| 120 | 10 | 1.09 | 1.97 | −11.19 | 0.85 | 1.90 | −13.81 | 0.24 | 0.07 | 2.63 |
| | 20 | 3.23 | 4.23 | −5.90 | 2.94 | 4.18 | −7.44 | 0.29 | 0.05 | 1.53 |
| | 30 | 5.49 | 6.50 | −3.88 | 5.19 | 6.47 | −4.93 | 0.30 | 0.03 | 1.05 |
| Panel C: Out-Of-The-Money (Strike-to-Stock Price = 0.90) | | | | | | | | | | |
| 30 | 10 | 0.00 | 0.00 | N/A | 0.00 | 0.00 | N/A | 0.00 | 0.00 | N/A |
| | 20 | 0.05 | 0.07 | −29.07 | 0.05 | 0.07 | −30.24 | 0.00 | 0.00 | 1.17 |
| | 30 | 0.36 | 0.42 | −16.09 | 0.35 | 0.42 | −17.16 | 0.01 | 0.00 | 1.07 |
| 60 | 10 | 0.00 | 0.00 | N/A | 0.00 | 0.00 | N/A | 0.00 | 0.00 | N/A |
| | 20 | 0.22 | 0.32 | −15.96 | 0.21 | 0.32 | −17.22 | 0.01 | 0.00 | 1.26 |
| | 30 | 0.95 | 1.18 | −9.39 | 0.93 | 1.17 | −10.32 | 0.02 | 0.00 | 0.93 |
| 120 | 10 | 0.01 | 0.05 | −19.33 | 0.01 | 0.05 | −20.08 | 0.00 | 0.00 | 0.75 |
| | 20 | 0.59 | 0.94 | −9.27 | 0.55 | 0.93 | −10.20 | 0.04 | 0.01 | 0.93 |
| | 30 | 1.96 | 2.52 | −5.58 | 1.87 | 2.51 | −6.34 | 0.08 | 0.01 | 0.76 |

## Table 1.2: The Early Exercise Risk Premium, Stochastic Volatility, and Jumps

The table presents the difference in the monthly expected return between American and equivalent European put options under alternative stochastic processes employed to model the evolution of the underlying asset value. The alternative processes are Geometric Brownian motion (GBM; Panel A), a stochastic volatility (SV) process (Panel B), and a stochastic volatility-jump (SVJ) process (Panel C). Columns (1) to (3) consider in-the-money (ITM; strike-to-stock price=1.10), columns (4) to (6) at-the-money (ATM; 1.00), and columns (7) to (9) out-of-the-money (OTM; 0.90) options. Conversely, columns (1), (4), and (7) consider 30, columns (2), (5), and (8) 60, and columns (3), (6), and (9) 90 day-to-maturity options. Finally, the first, second, and third row in each panel consider options written on an underlying asset with an annualized volatility of 10%, 20%, and 30%, respectively. We calculate the options' expected payoffs and market values using Longstaff and Schwartz's (2001) least-squares method applied to simulated underlying-asset-value paths under either the physical $\mathcal{P}$ (expected option payoff) or risk-neutral $\mathcal{Q}$ (option market value) probability measure. In case of each statistic, we rely on 1,000,000 underlying-asset-value paths featuring a number of time steps equal to an option's days-to-maturity. The monthly expected option return is the expected option payoff divided by its market value minus one, scaled by the months-to-maturity. We describe the basecase parameter values in Section A.

| Vol. | ITM Options Days-to-Maturity | | | ATM Options Days-to-Maturity | | | OTM Options Days-to-Maturity | | |
|---|---|---|---|---|---|---|---|---|---|
| (%) | 30 | 60 | 120 | 30 | 60 | 120 | 30 | 60 | 120 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Panel A: Geometric Brownian Motion Model | | | | | | | | | |
| 10 | 8.09 | 8.14 | 7.69 | 6.41 | 4.11 | 2.63 | N/A | N/A | 0.75 |
| 20 | 5.65 | 4.06 | 2.50 | 3.31 | 2.23 | 1.53 | 1.17 | 1.26 | 0.93 |
| 30 | 3.40 | 2.15 | 1.35 | 2.21 | 1.51 | 1.05 | 1.07 | 0.93 | 0.76 |
| Panel B: Stochastic Volatility (SV) Model | | | | | | | | | |
| 10 | 8.14 | 8.24 | 7.83 | 6.01 | 3.96 | 2.73 | N/A | N/A | 0.36 |
| 20 | 6.08 | 4.19 | 2.64 | 3.19 | 2.19 | 1.50 | N/A | 1.03 | 0.97 |
| 30 | 3.45 | 2.22 | 1.41 | 2.24 | 1.54 | 1.03 | 1.63 | 1.32 | 0.73 |
| Panel C: Stochastic Volatility-Jump (SVJ) Model | | | | | | | | | |
| 10 | 8.13 | 8.10 | 7.23 | 1.98 | 1.54 | 1.21 | 0.10 | 0.34 | 0.48 |
| 20 | 5.40 | 3.76 | 2.30 | 2.34 | 1.75 | 1.22 | 0.48 | 0.81 | 0.71 |
| 30 | 3.68 | 2.25 | 1.34 | 2.11 | 1.51 | 1.02 | 0.65 | 1.15 | 0.70 |

## Table 1.3: Descriptive Statistics

The table presents descriptive statistics on the monthly returns of American put options (column (1)), synthetic European put options (column (2)), and spread portfolios long an American put option and short its equivalent synthetic European option (column (1)–(2)). The table further reports the moneyness (column (3)) and time-to-maturity (column (4)) of the American and European option pairs. The descriptive statistics include the mean, the standard deviation (StDev), the $t$-statistic for the mean (*Mean/StError*), the bootstrap 95% confidence interval for the $t$-statistic (*95%BS-CI*), several percentiles, and the number of observations. The observations used in columns (1) and (2) are matched along the moneyness and time-to-maturity dimension, so that each observation in column (1) corresponds to exactly one observation in column (2) with the same moneyness and time-to-maturity. We calculate moneyness as the ratio of the option strike price to the stock price. We measure time-to-maturity in terms of calendar days. With the exception of the $t$-statistic and the bootstrap 95% confidence interval for the $t$-statistic, we calculate each statistic as the time-series average of the cross-sectional statistic.

| | Monthly American Put Option Return (in %) | Monthly Synthetic European Put Option Return (in %) | Monthly Spread Portfolio Return (in %) | Moneyness Option Pair | Days to Maturity Option Pair |
|---|---|---|---|---|---|
| | (1) | (2) | (1)–(2) | (3) | (4) |
| Mean | −12.76 | −16.42 | 3.66 | 1.03 | 74 |
| StDev | 61.19 | 63.48 | 22.70 | 0.09 | 26 |
| Mean/StError | [−5.36] | [−7.06] | [8.67] | | |
| 95%BS-CI | {−2.19;1.81} | {−2.19;1.83} | {−1.85;2.11} | | |
| Percentile 1 | −91.74 | −95.88 | −45.58 | 0.83 | 48 |
| Percentile 5 | −83.38 | −87.59 | −15.47 | 0.88 | 49 |
| Quartile 1 | −55.13 | −59.07 | −2.13 | 0.98 | 50 |
| Median | −23.69 | −27.98 | 0.70 | 1.04 | 70 |
| Quartile 3 | 14.81 | 10.09 | 4.61 | 1.10 | 99 |
| Percentile 95 | 93.65 | 92.41 | 37.67 | 1.17 | 111 |
| Percentile 99 | 193.69 | 203.05 | 89.40 | 1.19 | 111 |
| Observations | 1,384 | 1,384 | 1,384 | 1,384 | 1,384 |

## Table 1.4: Portfolios Double-Sorted on Moneyness and Maturity Time

The table presents the mean returns of moneyness and time-to-maturity-sorted American put option portfolios (column (1)), synthetic European put option portfolios (column (2)), as well as spread portfolios long the American and short the European option portfolio (column (1)–(2)). At the end of each sample month $t-1$, we first sort options into portfolios according to whether their strike-to-stock price ratio ("moneyness") lies above 1.05 (Panel A), between 0.95 and 1.05 (Panel B), or below 0.95 (Panel C). Within each moneyness portfolio, we next sort them into portfolios according to whether their days-to-maturity are below 60, between 60 and 90, or above 90 days. We equally-weight the portfolios and hold them over month $t$. The observations used in columns (1) and (2) are matched, so that each observation in column (1) corresponds to exactly one observation in column (2) with the same moneyness and time-to-maturity. Plain numbers are mean monthly portfolio returns (in %), the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a lag length of twelve months, and the numbers in curly parentheses are bootstrap 95% confidence intervals for the $t$-statistic.

| Days-to-Maturity | Monthly American Put Option Return (in %) | Monthly Synthetic European Put Option Return (in %) | Monthly Spread Portfolio Return (in %) |
|---|---|---|---|
| | (1) | (2) | (1)–(2) |
| **Panel A: In-The-Money (Strike-to-Stock Price > 1.05)** | | | |
| 30-60 | −27.66 | −39.39 | 11.74 |
| | [−13.61] | [−22.98] | [12.04] |
| | {−2.12;1.87} | {−2.18;1.80} | {−2.07;1.92} |
| 60-90 | −10.13 | −13.60 | 3.47 |
| | [−5.28] | [−7.28] | [8.13] |
| | {−2.07;1.90} | {−2.09;1.87} | {−1.83;2.10} |
| 90-120 | −5.02 | −6.16 | 1.14 |
| | [−3.19] | [−3.88] | [4.49] |
| | {−2.12;1.91} | {−2.13;1.89} | {−1.84;2.11} |
| **Panel B: At-The-Money (Strike-to-Stock Price 0.95 to 1.05)** | | | |
| 30-60 | −20.43 | −24.98 | 4.55 |
| | [−7.21] | [−9.49] | [7.20] |
| | {−2.15;1.80} | {−2.27;1.77} | {−2.03;1.87} |
| 60-90 | −6.76 | −7.11 | 0.35 |
| | [−2.63] | [−2.76] | [1.31] |
| | {−2.17;1.83} | {−2.20;1.81} | {−1.73;2.24} |
| 90-120 | −3.47 | −3.16 | −0.31 |
| | [−1.72] | [−1.54] | [−2.16] |
| | {−2.12;1.84} | {−2.11;1.85} | {−1.85;2.08} |
| **Panel C: Out-Of-The-Money (Strike-to-Stock Price < 0.95)** | | | |
| 30-60 | −7.59 | −6.91 | −0.68 |
| | [−1.83] | [−1.55] | [−0.90] |
| | {−2.23;1.79} | {−2.27;1.77} | {−1.78;2.23} |
| 60-90 | −5.14 | −3.63 | −1.50 |
| | [−1.38] | [−0.84] | [−1.82] |
| | {−2.40;1.72} | {−2.58;1.65} | {−1.40;4.02} |
| 90-120 | −3.03 | −2.30 | −0.73 |
| | [−1.16] | [−0.83] | [−2.65] |
| | {−2.19;1.81} | {−2.19;1.83} | {−1.85;2.11} |

### Table 1.5: Portfolios Univariately-Sorted on Idiosyncratic Volatility

The table presents the mean returns of American put option portfolios, synthetic European put option portfolios, as well as spread portfolios long the American and short the European option portfolio sorted on idiosyncratic stock volatility. At the end of each sample month $t-1$, we sort options into portfolios according to the quintile breakpoints of their underlying stock's market-model (Panel A) or Fama-French-Carhart model (Panel B) idiosyncratic volatility over the prior 60 months. We also form a spread portfolio long the top and short the bottom quintile ("High–Low"). We equally-weight the portfolios and hold them over month $t$. The American and European option observations are matched, so that each American option observation corresponds to exactly one European option observation with the same moneyness and time-to-maturity. Plain numbers are mean monthly portfolio returns (in %), the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length, and the numbers in curly parentheses are bootstrap 95% confidence intervals for the $t$-statistic.

| Idiosyncratic Stock Volatility | | | | | |
|---|---|---|---|---|---|
| 1 (Low) | 2 | 3 | 4 | 5 (High) | High–Low |
| Panel A: Market Model Idiosyncratic Volatility | | | | | |
| *Panel A1: American Put Return* | | | | | |
| −14.31 | −13.00 | −12.34 | −13.34 | −10.48 | 3.83 |
| [−4.99] | [−5.85] | [−4.71] | [−5.55] | [−4.20] | [1.97] |
| {−2.26;1.77} | {−2.19;1.81} | {−2.14;1.83} | {−2.12;1.82} | {−2.11;1.85} | {−1.91;2.01} |
| *Panel A2: Synthetic European Put Return* | | | | | |
| −19.96 | −17.71 | −16.18 | −16.17 | −11.74 | 8.22 |
| [−6.71] | [−8.21] | [−6.47] | [−6.83] | [−4.69] | [3.65] |
| {−2.42;1.72} | {−2.16;1.83} | {−2.13;1.86} | {−2.14;1.85} | {−2.12;1.84} | {−1.87;2.10} |
| *Panel A3: Spread Portfolio Return* | | | | | |
| 5.65 | 4.71 | 3.84 | 2.83 | 1.26 | −4.39 |
| [6.60] | [10.38] | [10.01] | [8.51] | [3.56] | [−6.37] |
| {−1.62;2.42} | {−1.95;1.98} | {−1.98;1.95} | {−2.04;1.92} | {−1.87;2.10} | {−2.31;1.69} |
| Panel B: FFC Model Idiosyncratic Volatility | | | | | |
| *Panel B1: American Put Return* | | | | | |
| −14.22 | −13.09 | −12.84 | −12.63 | −10.86 | 3.36 |
| [−4.94] | [−5.87] | [−4.99] | [−5.18] | [−4.46] | [1.74] |
| {−2.22;1.75} | {−2.19;1.79} | {−2.13;1.82} | {−2.14;1.82} | {−2.10;1.85} | {−1.94;2.03} |
| *Panel B2: Synthetic European Put Return* | | | | | |
| −19.80 | −17.85 | −16.64 | −15.61 | −12.02 | 7.78 |
| [−6.61] | [−8.30] | [−6.72] | [−6.54] | [−4.92] | [3.46] |
| {−2.43;1.70} | {−2.19;1.80} | {−2.14;1.80} | {−2.13;1.84} | {−2.12;1.84} | {−1.84;2.17} |
| *Panel B3: Spread Portfolio Return* | | | | | |
| 5.59 | 4.76 | 3.80 | 2.98 | 1.16 | −4.43 |
| [6.32] | [10.89] | [9.24] | [9.51] | [3.30] | [−6.11] |
| {−1.61;2.42} | {−1.99;1.94} | {−1.98;1.92} | {−2.05;1.90} | {−1.85;2.08} | {−2.27;1.66} |

### Table 1.6: Subperiod Tests

The table presents the mean returns of moneyness and time-to-maturity-sorted spread portfolios long an American and short the equivalent European option portfolio separately calculated over the January-1996 to December-2008 (column (1)) and the January-2009 to April-2016 (column (2)) subsample periods. The table also reports the differences in mean spread portfolio returns across the subsample periods (column (2)–(1)). In Panels A, B, and C, we consider in-the-money, at-the-money, and out-of-the-money options, respectively. Within each panel, we further consider options with a short, medium, or long time-to-maturity. See the caption of Table 1.4 for details on how the double-sorted spread portfolios are created. The American and European option observations are matched, so that each American option observation corresponds to exactly one European option observation with the same moneyness and time-to-maturity. Plain numbers are mean monthly portfolio returns (in %), the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a lag length equal to twelve months, and the numbers in curly parentheses are bootstrap 95% confidence intervals for the $t$-statistic.

| | Monthly Spread Portfolio Return (in %) | | |
|---|---|---|---|
| Days-to-Maturity | Until 2008 | From 2009 | Difference |
| | (1) | (2) | (2)–(1) |
| Panel A: In-The-Money (Strike-to-Stock Price > 1.05) | | | |
| 30-60 | 13.36 | 8.88 | −4.49 |
| | [12.49] | [5.89] | [−3.26] |
| | {−2.03;1.93} | {−2.45;1.71} | {−1.69;1.69} |
| 60-90 | 4.16 | 2.24 | −1.92 |
| | [7.61] | [5.24] | [−2.58] |
| | {−1.72;2.23} | {−2.17;1.86} | {−1.68;1.72} |
| 90-120 | 1.47 | 0.56 | −0.90 |
| | [4.48] | [1.92] | [−1.94] |
| | {−1.79;2.18} | {−1.95;2.02} | {−1.71;1.68} |
| Panel B: At-The-Money (Strike-to-Stock Price 0.95 to 1.05) | | | |
| 30-60 | 5.20 | 3.39 | −1.81 |
| | [6.21] | [3.52] | [−1.32] |
| | {−2.02;1.93} | {−2.80;1.67} | {−1.63;1.62} |
| 60-90 | 0.44 | 0.18 | −0.26 |
| | [1.22] | [0.49] | [−0.42] |
| | {−1.61;2.51} | {−2.20;1.78} | {−1.65;1.61} |
| 90-120 | −0.08 | −0.71 | −0.63 |
| | [−0.56] | [−3.03] | [−1.63] |
| | {−1.84;2.13} | {−1.77;2.32} | {−1.71;1.76} |
| Panel C: Out-Of-The-Money (Strike-to-Stock Price < 0.95) | | | |
| 30-60 | 0.94 | −3.55 | −4.50 |
| | [1.23] | [−3.89] | [−3.26] |
| | {−1.68;2.37} | {−1.78;2.30} | {−1.68;1.68} |
| 60-90 | −2.00 | −0.64 | 1.36 |
| | [−1.61] | [−1.19] | [0.74] |
| | {−1.34;5.37} | {−1.83;2.20} | {−1.49;1.66} |
| 90-120 | −0.60 | −0.96 | −0.36 |
| | [−1.91] | [−1.87] | [−0.55] |
| | {−1.84;2.11} | {−1.84;2.19} | {−1.71;1.67} |

## Table 1.7: Fama-MacBeth (1973) Regressions

The table presents the results of Fama-MacBeth (1973) regressions of the return over month $t$ of a spread portfolio long an American put option and short its equivalent synthetic European option (Panel A), an American put option (Panel B), or a synthetic European put option (Panel C) on subsets of stock and option characteristics plus a constant. The characteristics include the strike-to-stock price ratio ("moneyness"), time-to-maturity (as fraction of a year), and idiosyncratic underlying-stock volatility, all measured at the start of month $t$. We calculate idiosyncratic stock volatility from the Fama-French-Carhart model estimated over the prior 60 months. The American and European option observations are matched, so that each American option observation corresponds to exactly one European option observation with the same moneyness and time-to-maturity. The plain numbers are premium estimates, the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length, and the numbers in curly parentheses are the bootstrap 95% confidence intervals for the $t$-statistic.

| | Regression Model: | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: Spread Portfolio Return** | | | | |
| Constant | 0.04 | −0.24 | 0.07 | −0.22 |
| | [9.44] | [−13.30] | [8.75] | [−11.44] |
| | {−1.90;1.90} | {−2.08;1.90} | {−2.00;2.05} | {−1.84;1.95} |
| Moneyness | | 0.35 | | 0.35 |
| | | [19.27] | | [19.17] |
| | | {−2.00;1.89} | | {−1.99;2.09} |
| Time-to-Maturity | | −0.41 | | −0.40 |
| | | [−12.22] | | [−12.57] |
| | | {−1.79;2.06} | | {−2.00;1.93} |
| Volatility | | | −0.07 | −0.05 |
| | | | [−6.21] | [−5.50] |
| | | | {−1.88;1.91} | {−1.93;2.06} |
| **Panel B: American Put Option Return** | | | | |
| Constant | −0.13 | 0.05 | −0.15 | 0.04 |
| | [−6.24] | [0.42] | [−6.07] | [0.31] |
| | {−1.97;2.14} | {−2.04;1.90} | {−2.03;1.87} | {−1.97;1.94} |
| Moneyness | | −0.39 | | −0.38 |
| | | [−4.41] | | [−4.30] |
| | | {−1.94;1.89} | | {−1.90;1.96} |
| Time-to-Maturity | | 1.08 | | 1.08 |
| | | [23.27] | | [24.39] |
| | | {−1.93;1.97} | | {−1.80;2.09} |
| Volatility | | | 0.04 | 0.00 |
| | | | [1.51] | [0.02] |
| | | | {−1.86;1.96} | {−1.92;1.99} |
| **Panel C: Synthetic European Put Option Return** | | | | |
| Constant | −0.16 | 0.29 | −0.21 | 0.26 |
| | [−7.74] | [2.34] | [−8.15] | [2.06] |
| | {−1.99;1.81} | {−2.03;1.81} | {−2.01;1.87} | {−1.96;1.89} |
| Moneyness | | −0.74 | | −0.73 |
| | | [−7.64] | | [−7.54] |
| | | {−1.88;1.98} | | {−1.90;2.07} |
| Time-to-Maturity | | 1.50 | | 1.48 |
| | | [30.49] | | [32.88] |
| | | {−2.04;1.93} | | {−2.03;1.84} |
| Volatility | | | 0.11 | 0.05 |
| | | | [3.53] | [1.84] |
| | | | {−2.03;1.94} | {−1.90;1.90} |

## Table 1.8: Controlling for Stock Short-Selling Constraints

The table presents the mean returns of moneyness and time-to-maturity-sorted spread portfolios long an American and short the equivalent European option portfolio formed from only options written on stocks with a non-missing DCBS value or one equal to or below eight, seven, and five at the start of the option return period (columns (1) to (4), respectively). In Panels A, B, and C, we consider in-the-money, at-the-money, and out-of-the-money options, respectively. Within each panel, we further consider options with a short, medium, or long time-to-maturity. See the caption of Table 1.4 for details on how the double-sorted spread portfolios are created. The American and European option observations are matched, so that each American option observation corresponds to exactly one European option observation with the same moneyness and time-to-maturity. Plain numbers are mean monthly portfolio returns (in %), while the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length. An asterisk (*) indicates that the $t$-statistic lies outside of its bootstrap 95% confidence interval.

| | Monthly Spread Portfolio Return (in %) | | | |
|---|---|---|---|---|
| Days-to-Maturity | All Stocks with Available DCBS Value | Stocks with DCBS Value $\leq 8$ | Stocks with DCBS Value $\leq 7$ | Stocks with DCBS Value $\leq 5$ |
| | (1) | (2) | (3) | (4) |
| Panel A: In-the-Money (Strike-to-Stock Price > 1.05) | | | | |
| 30-60 | 12.12* | 12.34* | 12.43* | 12.68* |
| | [8.13] | [8.38] | [8.46] | [8.74] |
| 60-90 | 3.58* | 3.68* | 3.78* | 3.89* |
| | [6.51] | [6.76] | [7.09] | [7.39] |
| 90-120 | 1.18* | 1.26* | 1.33* | 1.43* |
| | [2.94] | [3.24] | [3.61] | [4.12] |
| Panel B: At-The-Money (Strike-to-Stock Price 0.95 to 1.05) | | | | |
| 30-60 | 4.16* | 4.41* | 4.52* | 4.71* |
| | [5.96] | [6.19] | [6.42] | [6.83] |
| 60-90 | 0.64* | 0.83* | 1.01* | 1.25* |
| | [2.68] | [3.22] | [3.73] | [4.51] |
| 90-120 | −0.09 | 0.14 | 0.25 | 0.44* |
| | [−0.42] | [0.68] | [1.32] | [2.35] |
| Panel C: Out-Of-The-Money (Strike-to-Stock Price < 0.95) | | | | |
| 30-60 | −1.23 | −0.79 | −0.51 | −0.14 |
| | [−1.07] | [−0.71] | [−0.48] | [−0.14] |
| 60-90 | −0.74 | −0.28 | −0.09 | 0.43 |
| | [−1.59] | [−0.69] | [−0.23] | [1.10] |
| 90-120 | −0.57 | −0.15 | 0.10 | 0.43 |
| | [−1.64] | [−0.42] | [0.29] | [1.18] |

## Table 1.9: Controlling for Stock and Option Illiquidity

The table presents the mean returns of illiquidity-sorted spread portfolios long an American and short the equivalent European option portfolio formed using only in-the-money (Panel A), at-the-money (Panel B), or out-of-the-money (Panel C) options with a short, medium, or long time-to-maturity and on stocks with a DCBS value equal to or below five. See the captions of Tables 1.4 and 1.8 for details on the moneyness, time-to-maturity, and DCBS classifications. At the end of each sample month $t-1$, we first sort the option pairs into portfolios based on the median of the illiquidity-proxy value for the American call option, then based on the median of that for the American put option, and finally based on the median of that for the underlying stock. We use either the inverse of scaled open interest or the scaled bid-ask spread to proxy for option illiquidity and Amihud's (2002) measure to proxy for stock illiquidity. The intersection of the three portfolio sets yields the independently triple-sorted illiquidity portfolios. We equally-weight the triple-sorted portfolios and hold them over month $t$. The American and European option observations are matched, so that each American option observation corresponds to exactly one European option observation with the same moneyness and time-to-maturity. Plain numbers are mean monthly portfolio returns (in %), while the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length. An asterisk (*) indicates that the $t$-statistic lies outside of its bootstrap 95% confidence interval. To conserve space, the table only reports the mean returns of those triple-sorted portfolios for which either all stocks and options are in the high (H-H-H) or in the low (L-L-L) univariate illiquidity portfolios.

| | Monthly Spread Portfolio Return (in %) Option Illiquidity Proxy | | | |
|---|---|---|---|---|
| | Bid-Ask Spread | | Inverse Open Interests | |
| Days-to-Maturity | H-H-H | L-L-L | H-H-H | L-L-L |
| | (1) | (2) | (3) | (4) |
| Panel A: In-The-Money (Strike-to-Stock Price > 1.05) | | | | |
| 30-60 | 15.01* | 9.10* | 10.72* | 10.93* |
| | [8.98] | [6.15] | [7.27] | [6.54] |
| 60-90 | 4.91* | 2.00* | 3.42* | 2.67* |
| | [7.62] | [3.24] | [4.76] | [3.76] |
| 90-120 | 1.55* | 0.50 | 1.65* | 1.27* |
| | [3.23] | [1.26] | [3.18] | [3.70] |
| Panel B: At-The-Money (Strike-to-Stock Price 0.95 to 1.05) | | | | |
| 30-60 | 5.27* | 4.01* | 5.19* | 3.51* |
| | [5.86] | [5.11] | [4.31] | [6.13] |
| 60-90 | 2.49* | 0.64 | 1.93* | 0.31 |
| | [4.20] | [1.61] | [3.27] | [0.94] |
| 90-120 | 0.05 | 0.20 | −0.13 | 1.29* |
| | [0.07] | [0.28] | [−0.27] | [2.90] |
| Panel C: Out-Of-The-Money (Strike-to-Stock Price < 0.95) | | | | |
| 30-60 | 1.52 | −3.12 | 1.54 | −1.24 |
| | [0.80] | [−1.42] | [1.12] | [−1.02] |
| 60-90 | 3.12 | −0.08 | 2.93 | −1.05 |
| | [1.85] | [−0.16] | [1.87] | [−1.49] |
| 90-120 | 1.04 | 0.17 | −0.12 | 0.22 |
| | [1.02] | [0.44] | [−0.11] | [0.38] |

## Table 1.10: Stocks That Never Paid Dividends

The table presents the mean returns of moneyness and time-to-maturity-sorted American put option portfolios (column (1)), synthetic European put option portfolios (column (2)), as well as spread portfolios long the American and short the European option portfolio (column (1)–(2)) formed using only options written on stocks that never paid out a dividend until the start of the option return period. See the caption of Table 1.4 for details on how the double-sorted American put, European put, and spread portfolios are created. The American and European option observations are matched, so that each American option observation corresponds to exactly one European option observation with the same moneyness and time-to-maturity. The plain numbers are mean monthly portfolio returns (in %), while the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a lag length equal to twelve months. An asterisk (*) indicates that the $t$-statistic lies outside of its bootstrap 95% confidence interval.

| Days-to-Maturity | American Put Option Return (in %) | Syn. European Put Option Return (in %) | Spread Portfolio Return (in %) |
|---|---|---|---|
| | (1) | (2) | (1)–(2) |
| Panel A: In-The-Money (Strike-to-Stock Price > 1.05) | | | |
| 30-60 | −27.99* | −38.06* | 10.07* |
| | [−14.54] | [−21.93] | [12.05] |
| 60-90 | −9.79* | −12.67* | 2.88* |
| | [−5.12] | [−6.77] | [8.31] |
| 90-120 | −5.11* | −6.20* | 1.08* |
| | [−3.39] | [−4.12] | [4.36] |
| Panel B: At-The-Money (Strike-to-Stock Price 0.95 to 1.05) | | | |
| 30-60 | −20.64* | −24.33* | 3.69* |
| | [−8.25] | [−10.59] | [9.00] |
| 60-90 | −6.72* | −7.04* | 0.32 |
| | [−2.67] | [−2.79] | [1.47] |
| 90-120 | −4.09* | −3.82 | −0.26 |
| | [−2.14] | [−1.95] | [−1.63] |
| Panel C: Out-Of-The-Money (Strike-to-Stock Price < 0.95) | | | |
| 30-60 | −8.31* | −8.01 | −0.30 |
| | [−2.10] | [−1.90] | [−0.42] |
| 60-90 | −5.19 | −4.12 | −1.07* |
| | [−1.56] | [−1.15] | [−2.48] |
| 90-120 | −3.58 | −3.03 | −0.55 |
| | [−1.41] | [−1.12] | [−1.73] |

**Table 1.11: Accounting for Bid-Ask Transaction Costs**

The table presents the mean returns of moneyness and time-to-maturity-sorted spread portfolios long an American and short the equivalent European option portfolio under the assumption that investors always buy (sell) at the midpoint plus (minus) $S$ times the bid-ask spread. We set $S$ equal to 0.00, 0.10, 0.25, 0.50 in columns (1) to (4), respectively. In Panels A and B, we use either all option pairs or only those based on low illiquidity assets (those in the L-L-L portfolio in Table 1.9) and on stocks with a DCBS value equal to or below five, respectively. In subpanels 1, 2, and 3, we then consider in-the-money, at-the-money, and out-of-the-money options, respectively. Within each subpanel, we further consider options with a short, medium, or long time-to-maturity. See the caption of Table 1.4 for details on how the double-sorted spread portfolios are created. The American and European option observations are matched, so that each American option observation corresponds to exactly one European option observation with the same moneyness and time-to-maturity. Plain numbers are mean monthly portfolio returns (in %), while the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length. An asterisk (*) indicates that the $t$-statistic lies outside of its bootstrap 95% confidence interval.

| | Monthly Spread Portfolio Return (in %) Bid-Ask Spread Fraction $S$ Equal to: | | | |
|---|---|---|---|---|
| Days-to-Maturity | 0.00 | 0.10 | 0.25 | 0.50 |
| | (1) | (2) | (3) | (4) |
| Panel A: Full Sample | | | | |
| *Panel A1: In-The-Money (Strike-to-Stock Price > 1.05)* | | | | |
| 30-60 | 11.74* | 8.10* | 4.45* | −0.64 |
| | [12.04] | [8.96] | [4.66] | [−0.51] |
| 60-90 | 3.47* | 0.08 | −3.89* | −9.26* |
| | [8.13] | [0.20] | [−6.44] | [−8.56] |
| 90-120 | 1.14* | −2.13* | −6.30* | −11.95* |
| | [4.49] | [−6.75] | [−12.08] | [−11.54] |
| *Panel A2: At-The-Money (Strike-to-Stock Price 0.95 to 1.05)* | | | | |
| 30-60 | 4.55* | −0.99 | −7.71* | −17.18* |
| | [7.20] | [−1.62] | [−7.27] | [−8.52] |
| 60-90 | 0.35 | −5.00* | −12.24* | −22.76* |
| | [1.31] | [−10.73] | [−13.04] | [−12.05] |
| 90-120 | −0.31* | −5.63* | −12.93* | −23.68* |
| | [−2.16] | [−15.23] | [−16.71] | [−15.61] |
| *Panel A3: Out-Of-The-Money (Strike-to-Stock Price < 0.95)* | | | | |
| 30-60 | −0.68 | −9.73* | −20.90* | −35.85* |
| | [−0.90] | [−7.34] | [−9.74] | [−10.71] |
| 60-90 | −1.50* | −10.48* | −22.82* | −38.42* |
| | [−1.82] | [−9.59] | [−9.89] | [−11.55] |
| 90-120 | −0.73* | −9.65* | −21.98* | −40.05* |
| | [−2.65] | [−11.54] | [−12.63] | [−13.17] |
| Panel B: Sample with Low Option and Stock Illiquidity and DCBS Value ≤ Five | | | | |
| *Panel B1: In-The-Money (Strike-to-Stock Price > 1.05)* | | | | |
| 30-60 | 9.10* | 8.27* | 7.17* | 5.14* |
| | [6.15] | [5.83] | [5.30] | [4.01] |
| 60-90 | 2.00* | 1.74* | −0.13 | −2.89* |
| | [3.24] | [2.45] | [−0.17] | [−3.45] |
| 90-120 | 0.50 | −0.64 | −2.42* | −5.44* |
| | [1.26] | [−1.36] | [−4.34] | [−6.47] |

*(continued on next page)*

**Table 1.11: Accounting for Bid-Ask Transaction Costs (Cont.)**

| | Monthly Spread Portfolio Return (in %) Bid-Ask Spread Fraction $S$ Equal to: | | | |
|---|---|---|---|---|
| Days-to-Maturity | 0.00 | 0.10 | 0.25 | 0.50 |
| | (1) | (2) | (3) | (4) |
| *Panel B2: At-The-Money (Strike-to-Stock Price 0.95 to 1.05)* | | | | |
| 30-60 | 4.01* | 2.70* | −0.02 | −4.70* |
| | [5.11] | [3.12] | [−0.02] | [−3.51] |
| 60-90 | 0.64 | −1.18* | −3.70* | −8.61* |
| | [1.61] | [−2.88] | [−6.05] | [−10.06] |
| 90-120 | 0.20 | −1.76* | −5.21* | −9.94* |
| | [0.28] | [−2.67] | [−7.36] | [−9.68] |
| *Panel B3: Out-Of-The-Money (Strike-to-Stock Price < 0.95)* | | | | |
| 30-60 | −3.12 | −8.99* | −17.91* | −33.56* |
| | [−1.42] | [−3.30] | [−4.12] | [−3.43] |
| 60-90 | −0.08 | −4.20* | −10.66* | −21.30* |
| | [−0.16] | [−5.42] | [−6.27] | [−5.43] |
| 90-120 | 0.17 | −3.83* | −9.34* | −17.99* |
| | [0.44] | [−7.58] | [−7.53] | [−8.50] |

## Table 1.12: Time-Series Asset Pricing Tests

The table shows the results from time-series regressions of the return over month $t$ of a spread portfolio long an American and short the equivalent European put option on several sets of benchmark factors measured over the same month and a constant. The factors include the market return minus the risk-free rate of return (MKT), the return of a spread portfolio long small and short large stocks (SMB), the return of a spread portfolio long value and short growth stocks (HML), the return of a spread portfolio long winner and short loser stocks (MOM), the return of a spread portfolio long profitable and short unprofitable stocks (PRF), the return of a spread portfolio long non-investing and short investing stocks (INV), the change in the VIX option-implied volatility index (VIX), the three-month LIBOR rate minus the treasury bill rate (TED), and the return of a spread portfolio long high-liquidity and short low-liquidity stocks (LIQ). Plain numbers are estimates, while the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length. An asterisk (*) indicates that the $t$-statistic lies outside of its bootstrap 95% confidence interval.

| | Time-Series Regression Model: | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| MKT | 0.22* | 0.23* | 0.23* | 0.28* | 0.13 |
| | [2.67] | [2.64] | [2.48] | [2.59] | [0.96] |
| SMB | | −0.10 | −0.10 | 0.03 | 0.06 |
| | | [−0.82] | [−0.82] | [0.24] | [0.40] |
| HML | | −0.09 | −0.09 | −0.13 | −0.09 |
| | | [−0.76] | [−0.73] | [−0.69] | [−0.46] |
| MOM | | | 0.00 | −0.01 | −0.00 |
| | | | [0.01] | [−0.11] | [−0.04] |
| PRF | | | | 0.33 | 0.30 |
| | | | | [1.74] | [1.58] |
| INV | | | | −0.19 | −0.28 |
| | | | | [−0.78] | [−1.13] |
| VIX | | | | | −0.05* |
| | | | | | [−2.15] |
| TED | | | | | 1.24 |
| | | | | | [0.72] |
| LIQ | | | | | −0.20 |
| | | | | | [−1.89] |
| Constant | 0.04* | 0.04* | 0.04* | 0.04* | 0.04* |
| | [9.16] | [9.20] | [9.09] | [8.52] | [9.01] |

# Chapter 2

# Taking Money Off the Table: Suboptimal Early Exercises, Risky Arbitrage, and American Put Returns

***Keywords***: Empirical asset pricing; cross-sectional option pricing; put options; early exercise.

## 2.1 Introduction

Many studies suggest that American option investors do not always follow optimal early exercise policies, with them frequently exercising their positions too late. Pool et al. (2008), for example, estimate that the total profits lost from not optimally exercising single-stock American calls on ex-dividend dates are about $500 million over a ten-year period, while Barraclough and Whaley (2012) estimate that those from not optimally exercising American puts are about $1.9 billion over a twelve-year period. Yet, when American puts can be exercised too late, there is an arbitrage opportunity in perfect markets, exploitable by longing a dynamic underlying-asset-and-riskfree-asset portfolio replicating the put and shorting the traded put (Shreve (2004)). Intuitively, after the put should have but has not been exercised, the portfolio consists of one short underlying asset unit and an investment of the strike price into the money market, covering any obligations arising from the put but also earning interest. The earned interest represents the arbitrage profit.

In our paper, we study the profitability and total and systematic risk of the arbitrage strategy in real markets in which options can only be imperfectly replicated. While real investors are thus unable to earn a proper arbitrage profit, they may still be able to earn some profit with an only low risk ("risky arbitrage profit"). In accordance, we find that longing daily-rebalanced replication portfolios of single-stock American puts and shorting those puts earns us a highly significant mean return of 4.11% per month before transaction costs. Suggesting daily rebalancing creates efficient replication portfolios, the returns on the two legs of that strategy share a mean cross-sectional correlation of –0.85; the variance of the strategy return is only 30-35% of those on the returns of the legs; and the strategy only weakly loads on risk factors. In line with theory, the mean strategy return increases with the strike price and the interest rate. Also, it is higher for short time-to-maturity puts on low volatility stocks, aligning with further evidence that investors do not understand how those variables condition the optimal early exercise decision. Crucially, since the strategy requires frequent rebalancing and short-selling, we finally show that it survives accounting for trading and shorting costs, at least when executed on liquid assets.

We use options market data from Optionmetrics and early exercise data from Bob Whaley to study our arbitrage strategy in real markets. To derive the return on the long leg of the strategy, we invest the market price of the put into a portfolio consisting of the underlying stock and the riskfree asset at the start of the strategy return period, with the number of stocks equal to the (negative) delta of the put. At the end of each day, we then rebalance the stock holdings in the portfolio to its new delta. To derive the return on the short leg, we calculate the return of the traded put over the same period as the compounded early exercise payoff (if there is an early exercise) or the market price (if there is none) of the put at the end of that period to its market price at the start. To find out if and when the put is early exercised, we recognize that the Option Clearing Corporation randomly assigns exercise obligations to outstanding short positions (Pool et al. (2008)). We thus assume that the short leg is terminated over a day if a draw from the univariate distribution lies below the ratio of early exercised contracts of the put over that day to its open interest at the end of the prior day ("daily early exercise probability").[1]

We finally calculate the return of the risky arbitrage strategy as the spread in returns between the replication portfolio and the traded put over the strategy return period. Importantly, however, the strategy return period is not fixed, ranging from the start of a month to the earlier of the day over which the traded put is exercised and the end of the month. In other words, we always liquidate the long and the short leg of the risky arbitrage strategy on the same date.

Our evidence suggests that the risky arbitrage strategy is highly profitable with an only low total or systematic risk. Longing an equally-weighted portfolio of daily-rebalanced replication portfolios of all outstanding American puts and shorting an equally-weighted portfolio of those same puts yields a mean monthly return of 4.11% ($t$-statistic: 5.21). While the monthly variances of the returns on the long and the short leg of the strategy are, respectively, 0.32 and 0.26, the monthly variance of the strategy return is a much lower 0.09. Regressing the strategy's return on popular risk factors, such as the Fama-French (2015) five-factor model factors, we obtain alphas close to identical to the strategy's mean return but with higher $t$-statistics. In addition, we also obtain factor loadings markedly attenuated compared to those obtained from the two legs of the strategy. Calculating key performance evaluation statistic (e.g., the Sharpe (1966) ratio), we find yet more evidence that the strategy bodes satisfactory risk-adjusted performance.

[1]Given that the put return depends on draws from a random distribution, it varies each time that we recalculate it, leading to concern that our empirical results cannot be replicated. Fortunately, however, those variations almost cancel out in the aggregate, likely due to a law of large numbers. Recalculating the mean monthly pooled-sample return of the risky arbitrage strategy three times, we, for example, obtain values of 4.11455%, 4.11022%, and 4.11812%. We plan to include a histogram of 1,000 mean strategy returns in future versions of our paper.

We next condition the risky arbitrage strategy on put, stock, and macroeconomic characteristics including the strike price, the interest rate, moneyness, time-to-maturity, and underlying stock volatility. While theory suggests that, ceteris paribus, the strategy is more profitable for high strike-price puts in high interest-rate regimes, it is hard to generate further predictions since the strategy's profitability for some class of puts ultimately hinges on the extent to which investors early exercise puts too late within that class. Our evidence shows that the mean strategy return significantly increases with the strike price and the interest rate, but significantly decreases with moneyness, time-to-maturity, and stock volatility. Corroborating that the relations with time-to-maturity and stock volatility, but not that with moneyness, originate from variations in investors' tendency to exercise puts too late, we offer further evidence that investors do not seem to understand the negative effects of those characteristics on the optimal early exercise decision.

We finally turn to the transaction costs incurred by the risky arbitrage strategy. Assuming that an asset's trading costs are proportional to its bid-ask spread and that a stock can be borrowed at Markit's indicative rate, we show that trading and borrowing costs greatly eat into the profitability of that strategy. Considering in-the-money (ITM) puts with 30-60 days-to-maturity, the mean strategy return, for example, drops from 4.83% per month ($t$-statistic: 5.87) in the no-transaction-cost case to –2.17% ($t$-statistic: –2.05) in the 25% bid-ask-spread transaction-cost case. Importantly however, the strategy remains profitable even net of transaction costs when we restrict our attention to liquid puts written on liquid underlying assets. Excluding from the above ITM puts those with a bid-ask spread above the median and/or written on stocks with an Amihud (2002) illiquidity value above the median, the mean strategy return is, for example, 3.92% ($t$-statistic: 3.20) even in the 25% bid-ask-spread transaction-cost case. We also establish that decreasing the rebalancing frequency of the replication portfolio can further help to raise the strategy's profitability net of transaction costs, without it greatly boosting the volatility of the strategy return.

Our work builds up on empirical studies suggesting that American option investors often do not follow optimal early exercise policies. While Overdahl and Martin (1994) show that most early exercises of single-stock American calls and puts fall within optimal boundaries on the underlying stock's price, Brennan and Schwartz (1977) document that the early exercises of such puts are typically inconsistent with the Black-Scholes (1973) framework. Finucane (1997) finds that investors often early exercise American calls on non-dividend stocks, conflicting with Merton's (1973) insight that it is never optimal to early exercise such calls. Digging deeper into Finucane's (1997) results, Poteshman and Serbin (2003) show that only individual but not institutional investors sometimes early exercise the former calls.[2] As already said, Pool et al. (2008) and Barraclough and Whaley (2012) find that the foregone profits from failing to optimally early exercise single-stock American calls and puts are economically large. More generally, Bauer et al. (2009) report that retail investors do not perform well in their option investments. While we offer further evidence that investors' early exercise strategies are often suboptimal, our main contribution to the above literature is to show how to make money from that suboptimality using a simple trading strategy.

We further add to empirical studies examining the performance of trading strategies. While stock strategies are dominant among those studies (see, e.g., Fama and French (1992), Lakonishok et al. (1994), and Carhart (1997)), recent studies have also evaluated option strategies. Coval and Shumway (2001), for example, report that writing zero-beta index straddles is profitable even after considering transaction costs. Vasquez (2017) documents that longing straddles with a high slope of their implied volatility term structure and shorting those with a low slope yields positive mean cost-adjusted returns. Conversely, Goyal and Saretto (2009) find that longing options with a large difference between realized and implied volatility and shorting those with a low difference yields positive mean "delta-hedged returns" (i.e., the options' returns neutralized with respect to their underlying stocks' returns). Cao et al. (2020) show that sorting options

[2]The recent studies of Jensen and Pedersen (2016), Battalio et al. (2020), and Figlewski (2019) highlight that trading costs can make it optimal to early exercise an American call on a non-dividend asset, implying that some of the early exercises classified as suboptimal by Finucane (1997) could be optimal. Notwithstanding, Pool et al.'s (2008) evidence that the vast majority of such calls are early exercised by retail (but not institutional) investors leads one to suspect that Finucane (1997) correctly classifies most of his early exercises.

based on characteristics used to predict stock returns also often yields significant mean delta-hedged returns, which can, however, differ in sign from the corresponding mean stock returns. We add to this literature by proposing a new risky arbitrage trading strategy involving single-stock American puts and their delta-replication portfolio firmly grounded in mathematical finance theory.

We finally add to studies looking into how option prices deviate from the values of their replication portfolios. While in a Black-Scholes (1973) perfect capital market the value of any option can be perfectly hedged/replicated using a dynamic asset-and-riskfree-asset portfolio, Leland (1985) proves that transaction costs can drive a wedge between the option's value and the replication portfolio value. Adding underlying-asset dividends and stochastic volatility, Perrakis and Lefoll (2000) and Gondzio et al. (2003) come to the same conclusion. Relying on existing mathematical finance theory, we contribute to this literature by showing that suboptimally late early exercises of American puts can also make the values of those puts deviate from the values of their (standard/non-consuming) replication portfolios, not only in theory but also in practice.

We proceed as follows. In Section 2.2, we review the underlying theory. In Section 2.3, we discuss our methodology and data. In Sections 2.4, 2.5, and 2.6, we present the historical performance of the risky arbitrage strategy, condition the strategy on put, stock, and macroeconomic characteristics, and adjust it for trading and borrowing transaction costs, respectively. Section 2.7 concludes.

## 2.2 Theory

In this section, we briefly review the mathematical finance theory suggesting that an ex-ante non-zero probability that an American put is exercised too late creates an arbitrage opportunity in perfect capital markets. We keep our review as intuitive as possible, referring to other papers for more technical details. Consider an American put giving its owner the option to sell an underlying asset worth $S(t)$ at time $t$ for a constant price of $K$ ("strike price") in each instant within the time period $t \in [0, T]$ ("maturity time"). To keep matters simple, the underlying asset does not pay out cash over the maturity time, and its value evolves according to the following geometric Brownian motion (GBM) under the risk-neutral ("martingale") measure $\mathbb{Q}$:

$$dS(t) = rS(t)dt + \sigma S(t)d\widetilde{W}(t), \tag{2.1}$$

where $r$ is the annualized risk-free rate of return, $\sigma$ is the annualized volatility of the underlying asset's return, and $d\widetilde{W}(t)$ is the differential of a Brownian motion.

Using risk-neutral pricing techniques, the value of the American put, $p(t, S(t))$, is:

$$p(t, S(t)) = \max_{\tau \in \Omega(t,T)} \tilde{E}[e^{-r(\tau - t)}(K - S(\tau))|S(t)], \tag{2.2}$$

where max is the maximum operator, $\tilde{E}$ the expectation under the $\mathbb{Q}$ measure, $\tau$ the (random) early exercise time associated with some *feasible* early exercise strategy (i.e., a strategy based only on information available at the current time), and $\Omega(t, T)$ the set of early exercise times associated with all feasible strategies. Karatzas (1988) and Jacka (1991) prove that the optimal feasible strategy is to early exercise the put as soon as the underlying asset value $S(t)$ drops below some time-variant boundary $L(t)$, which is bounded and increases convexly with time $t$. As long as $S(t) > L(t)$, we can thus replicate the put by investing $p(t, S(t))$ into a dynamically rebalanced portfolio containing the underlying asset and the riskfree asset and ensuring that the underlying asset investment is equal to $\Delta(t, S(t))S(t)$ in each instant, where $\Delta(t, S(t)) \equiv \partial p(t, S(t))/\partial S(t)$ is the (negative) put delta. It follows from the replication strategy that:

$$\frac{1}{2}\sigma^2 S(t)^2 p_{SS}(t, S(t)) + rS(t)p_S(t, S(t)) + p_t(t, S(t)) - rp(t, S(t)) = 0, \tag{2.3}$$

where $p_S \equiv \partial p(t, S(t))/\partial S(t)$, $p_{SS} \equiv \partial^2 p(t, S(t))/\partial S(t)^2$, and $p_t \equiv \partial p(t, S(t))/\partial t$. Conversely, as soon as $S(t) \leq L(t)$, $p(t, S(t)) = K - S(t)$. A direct computation then yields as following:

$$\frac{1}{2}\sigma^2 S(t)^2 p_{SS}(t,S(t)) + rS(t)p_S(t,S(t)) + p_t(t,S(t)) - rp(t,S(t)) = rK. \quad (2.4)$$

We are now in a good position to repeat the insights in Shreve's (2004) Corollary 8.4.3:

COROLLARY 8.4.3 *Consider an agent with initial capital $X(0) = p(0, S(0))$. Suppose that, in each instant, this agent holds a portfolio consisting of $\Delta(t, S(t))$ units of the underlying asset and the residual value invested into the risk-free asset. Further, assume the agent consumes cash, $C(t)$, from that portfolio at rate $rK$ per time unit if $S(t) \leq L(t)$ and else at rate zero. Then $X(t) = p(t, S(t))$ for all times $t \in [0, T]$. In particular, $X(t) \geq max(K - S(t), 0)$ for all times $t$ until $T$, so the agent can pay off a short position regardless of when the option is expired.*

To prove the corollary, Shreve (2004) starts with applying Itô's lemma to the differential of the discounted value of the American put, $d(e^{-rt}p(t, S(t))$:

$$\begin{aligned} d(e^{-rt}p(t,S(t))) &= e^{-rt}\Big[ - rp(t,S(t))dt + p_t(t,S(t))dt + p_S(t,S(t))dS(t) \\ &\quad + \frac{1}{2}p_{SS}(t,S(t))dS(t)dS(t)\Big] \qquad (2.5) \\ &= e^{-rt}\Big[ - rp(t,S(t)) + p_t(t,S(t)) + rS(t)p_S(t,S(t)) \\ &\quad + \frac{1}{2}\sigma^2 S(t)^2 p_{SS}(t,S(t))\Big] dt \\ &\quad + e^{-rt}\sigma S(t) p_S(t,S(t)) d\tilde{W}(t), \qquad (2.6) \end{aligned}$$

noting that Equations (2.3) and (2.4) imply that the term in square parentheses in Equation (2.6) is zero if $S(t) > L(t)$ and else $-rK$. He then stresses that the differential of the value of the portfolio containing the underlying asset and the riskfree asset is equal to:

$$dX(t) = \Delta(t)dS(t) + r(X(t) - \Delta(t)S(t))dt - C(t)dt, \quad (2.7)$$

so that the differential of the discounted value of that portfolio is equal to:

$$d(e^{-rt}X(t)) = e^{-rt}(-rX(t)dt + dX(t)) \tag{2.8}$$

$$= e^{-rt}(\Delta(t)dS(t) - r\Delta(t)S(t)dt - C(t)dt) \tag{2.9}$$

$$= e^{-rt}\sigma S(t)p_S(t, S(t))d\tilde{W}(t) - e^{-rt}rK\mathbb{I}_{S(t)\leq L(t)}dt, \tag{2.10}$$

where $\mathbb{I}_{S(t)\leq L(t)}$ is an indicator equal to one if $S(t) \leq L(t)$ and else zero. Comparing Equations (2.6) and (2.10), it is obvious that the values of the put and the portfolio are not only the same at the initial time $t = 0$, but also at any other time $t$ over the time-to-maturity.

Corollary 8.4.3 implies that the arbitrage profit in dollars, obtained from longing the asset-and-riskfree-asset portfolio and shorting the put over the period until the put expires is $e^{r\times\max(t^E-\tau,0)}\max(t^E - \tau, 0)rK$ in perfect markets, where $t^E$ is the actual expiration date of the put. As a result, the dollar arbitrage profit increases with the length of the period over which the put should but has not been early exercised ($\max(t^E - \tau, 0)$), the risk-free rate of return ($r$), as well as the strike price ($K$). Intuitively, as the underlying asset value $S(t)$ reaches the early exercise boundary $L(t)$ from above, the replication portfolio consists of one short unit of the underlying asset and an investment of the strike price $K$ into a money market account. If the put owner optimally early exercises at that point, we transfer the money market account to him/her in return for one long unit of the underlying asset, which we use to extinguish our short position in that asset. Conversely, if he/she does not early exercise, we hold onto the money market investment and the one short underlying asset unit, enabling us to earn interest on the money market investment equal to $rKdt$ in each instant.

Three remarks about the above analysis are in order. First, while we rely on the simplest possible assumptions in that analysis, Shreve (2004) highlights that the arbitrage opportunity generally exists in complete markets as long as investors' suboptimal exercise policies render the discounted American put value a *supermartingale* under the $\mathbb{Q}$ measure. The arbitrage opportunity thus also exists in, for example, stochastic volatility and mixed jump-diffusion models. Second, while there may be concern that the arbitrage payoff is negligible due to its dependence on the interest rate, that payoff can be large even in a low interest rate regime. To see that, consider an American put with $S(0) = K = 30$, $\sigma = 0.30$, $T = 0.25$, and $r = 0.01$, which ends up being early exercised six weeks too late (i.e., $\max(t^E - \tau, 0) = 0.125$). Scaling the compounded arbitrage payoff, $e^{0.01\times0.125} \times 0.01 \times 30 \times 0.125 = 0.0376$, by the initial put value, 1.7659, we obtain a three-month arbitrage return of 2.1%, translating into a monthly return of 70.80 basis points.[3,4] Hence, even when $r = 0.01$, the arbitrage return compares well to the mean returns of popular risky stock strategies, such as the SMB, HML, and MOM strategies. Notwithstanding, the arbitrage return is obviously much larger in a higher interest rate regime, with it, for example, being 3.8% *per month* when $r = 0.05$. Third, while the arbitrage profit should theoretically depend on the strike price, the arbitrage return should not. This is because as the strike price increases, both the arbitrage profit and option value required to calculate the arbitrage return increase proportionately, making the arbitrage return insensitive to changes in the strike price in theory.

In the remainder of our study, we evaluate the mean returns and the total and systematic risk of the arbitrage strategy outlined in this section in real markets. While many empirical studies suggest that real American put investors often early exercise their positions too late (a necessary condition for the strategy to be profitable), market imperfections such as transaction costs and discontinuous trading imply that we cannot perfectly replicate puts using their underlying assets and the riskfree asset in real markets. Given that, the arbitrage strategy in perfect markets, with a zero probability of a loss but a positive probability of a gain, becomes, at best, a risky arbitrage strategy in real markets, with a low probability of a loss but a high probability of a gain.

[3]In these calculations, we follow the standard practice in the stock literature to compute the return of a long-short strategy as the difference in returns between the long and short leg. In our case, the long and short leg share the same initial value, so that the return of the strategy becomes the difference in payoffs between the two legs (which is the arbitrage payoff $\max(t^E - \tau, 0)rK$) scaled by the initial value of the American put.

[4]We use a binomial tree with 1,000 time steps to calculate the American put value.

## 2.3 Methodology and Data

In this section, we review our methodology and data. We first explain how we calculate the returns of American puts and their replication portfolios with or without transaction costs, also elaborating on how we decide whether a short put is expired early. We next discuss how we compute optimal early exercise probabilities conditional on the Black-Scholes (1973) framework using the Longstaff-Schwartz (2001) approach. We finally outline our data and data sources.

### 2.3.1 Calculating the American Put Return

We compute the holding-period gross return of an American put, $R^p(t_0, t_H)$, as its compounded-up early exercise payoff (if there is an early exercise over that period) or its market price at the end of that period (if there is none) to its market price at the start of that period:

$$R^p(t_0, t_H) = v(t_H)/p(t_0), \tag{2.11}$$

where $t_0$ and $t_H$ are, respectively, the start and end day of the holding period, $v(t_x)$ is the put's value at the end of day $t_x$, and $p(t_x)$ is its market price at the end of the same day. If the put is early exercised on day $t_E < t_H$, its value at the end of the holding period, $v(t_H)$, is: $e^{r_f(t_E,t_H)} \max(K - S(t_E), 0)$, where $r_f(t_x, t_y)$ is the net riskfree rate of return from end of day $x$ to end of day $y$, $K$ is the strike price, and $S(t_x)$ is the underlying asset's value at the end of day $x$. Else, the value of the put on that date is its market price, $p(t_H)$, on the same date.[5]

Keeping in mind that the risky arbitrage strategy requires us to be short in the put, it is the put owners, not us, who determine if and when they early exercise their positions. When a put owner decides to early exercise a put, the Options Clearing Corporation randomly assigns the exercise to an outstanding short position.

[5] Consistent with convention, Equation (2.11) gives the return on one long unit of the American put. Given that we are, however, short that put in our risky arbitrage strategy, what we mostly care about is $-R^p(t_0, t_H)$, and not $R^p(t_0, t_H)$. Recognizing that is especially important when we adjust returns for transaction costs since else one could erroneously gain the impression that our adjustments increase, and not decrease, returns.

To mimic that procedure, we do the following. For each day within the holding period, we first calculate the daily early exercise probability for the entire put issue, defined as the number of early exercises of that issue over the day scaled by its open interest at the end of the prior day. Starting with the first day in that period, we draw a number from the univariate distribution with bounds zero and one, assuming that our short put position is expired if the drawn number lies below the daily early exercise probability. Unless the short position is expired, we move to the second day, again drawing a number from the univariate distribution and assuming the short position is expired if the number lies below the daily early exercise probability. We continue in that way until the end of the holding period.

Our return calculations above ignore transaction costs arising from trading the put. Following Goyal and Saretto (2009) and Cao and Han (2013), we assume that these costs are proportional to the put's bid-ask spread, $\varphi BAS^p(t_x)$, where $\varphi$ is a constant and $BAS^p(t_x)$ that spread at the end of day $t_x$. Again keeping in mind that we are short the put, we compute the transaction-cost-adjusted gross put return, $R^{p,tca}(t_0, t_H)$, in the no-early-exercise case by *adding* $\varphi BAS^p(t_H)$ to the numerator of Equation (2.11) and *subtracting* $\varphi BAS^p(t_0)$ from its denominator. Conversely, in the early-exercise case, we add $(1 - abs(\Delta(t_E)))\varphi BAS^s(t_E)$ to the early exercise payoff in the numerator, where $\varphi BAS^s(t_x)$ and $\Delta(t_x)$ are, respectively, the underlying asset's bid-ask spread and the put's delta at the end of day $t_x$ and abs is the absolute-value operator, while we again subtract $\varphi BAS^p(t_0)$ from the denominator. Thus, $R^{p,tca}(t_0, t_H)$ equals:

$$R^{p,tca}(t_0, t_H) = \begin{cases} \frac{e^{r_f(t_E,t_H)}\left(\max(K-S(t_E),0)+(1-abs(\Delta(t_E)))\varphi BAS^s(t_E)\right)}{p(t_0)-\varphi BAS^p(t_0)}; & t_E \leq t_H, \\ \frac{p(t_H)+\varphi BAS^p(t_H)}{p(t_0)-\varphi BAS^p(t_0)}; & t_E > t_H. \end{cases} \tag{2.12}$$

The $(1 - abs(\Delta(t_E)))\varphi BAS^s(t_E)$ adjustment in the early-exercise case arises since the long leg of our risky arbitrage strategy (the put replication portfolio to be discussed in the next subsection) is generally short delta underlying asset units on the early exercise date. As a result, when the put is early exercised against us, we use the underlying asset obtained from the put owner to extinguish our short delta position in that asset, implying that we only need to sell ($1-$ $abs(\Delta(t_E))$) of that asset in the market at a unit transaction cost of $\varphi BAS^s(t_E)$.[6]

Using the daily early exercise probabilities above, we also calculate the real-world early exercise probability of a put within an issue over some other period. To do so, note that one minus a daily early exercise probability yields the corresponding daily no-exercise ("survival") probability. Computing one minus the product of the daily survival probabilities over the period, we obtain the real-world early exercise probability over that same period.

### 2.3.2 Calculating the Replication Portfolio Return

We next consider a dynamic portfolio noisily replicating an American put by shorting a number of underlying assets close to the put's delta and investing the remaining portfolio value into the riskfree asset. Assuming we invest the market price of the put into that portfolio at the end of day $t_0$ and rebalance at the end of each day in $\mathbb{B} \in [t_1, t_2, \ldots t_N]$, where $t_0 < t_1 < \ldots < t_N$ and $t_N$ is the last date before the earlier of the put expiration date ($t_E$) and the holding period end date ($t_H$), we calculate the portfolio's return, $R^X(t_0, t_H)$, over the holding period as follows. We first calculate the value of the portfolio at the end of the first rebalancing day, $X(t_1)$, as:

$$X(t_1) = \Delta(t_0)S(t_0)R(t_0, t_1) + (p(t_0) - \Delta(t_0)S(t_0))R_f(t_0, t_1), \qquad (2.13)$$

where $R(t_x, t_{x+1})$ and $R_f(t_x, t_{x+1})$ are the gross underlying asset return and the riskfree rate of return from end of day $t_x$ to end of day $t_{x+1}$, respectively. The

[6]In case the put owner exercises his/her position on the optimal date or later, the put replication portfolio contains one unit of the underlying asset, implying that: $(1 - abs(\Delta(t_E)))\varphi BAS^s(t_E)) = 0$. The adjustment thus deals with the unlikely case in which the put owner exercises his/her position too early.

portfolio's value at the end of any other rebalancing date $t_k$ is then:

$$X(t_k) = \Delta(t_{k-1})S(t_{k-1})R(t_{k-1}, t_k) + (X(t_{k-1}) - \Delta(t_{k-1})S(t_{k-1}))R_f(t_{k-1}, t_k), \tag{2.14}$$

while its value at the end of the holding period, $t_H$, can be written as:

$$X(t_H) = e^{r_f(t_E, t_H)} \left(\Delta(t_N)S(t_N)R(t_N, t_E) + (X(t_N) - \Delta(t_N)S(t_N))R_f(t_N, t_E)\right) \tag{2.15}$$

if the put is expired before the end of the holding period and as:

$$X(t_H) = \Delta(t_N)S(t_N)R(t_N, t_H) + (X(t_N) - \Delta(t_N)S(t_N))R_f(t_N, t_H) \tag{2.16}$$

if it is not. Notice that the compounding in Equation (2.15) ensures that the portfolio's value is measured at time $t_H$ even when the put is expired earlier. We finally calculate the portfolio's gross return over the holding period by scaling by the initial investment:

$$R^X(t_0, t_H) = X(t_H)/X(t_0) = X(t_H)/p(t_0). \tag{2.17}$$

While we again abstract from transaction costs in our initial return calculations, such costs are likely to be even more important for the replication portfolio than the put due to the potentially frequent underlying-asset buys and sales necessary to ensure that the replication portfolio and put have similar deltas and due to the portfolio being short in the underlying asset. Again assuming that an asset's trading costs are proportional to its bid-ask spread, the total trading costs of the replication portfolio at the end of the holding period, $C^{TC}(t_0, t_H)$, are:

$$\begin{aligned} C^{TC}(t_0, t_H) \quad = \quad & \varphi\Bigg( e^{r_f(t_0, t_H)} \text{abs}(\Delta(t_0)) BAS^s(t_0) \\ + \quad & \sum_{i=1}^{N} e^{r_f(t_i, t_H)} \text{abs}(\Delta(t_i) - \Delta(t_{i-1})) BAS^s(t_i) && (2.18) \\ + \quad & \mathbb{I}_{\{t_E > t_H\}} \text{abs}(\Delta(t_H)) BAS^s(t_H) \Bigg), && (2.19) \end{aligned}$$

where $\mathbb{I}_{\{t_E>t_H\}}$ is a dummy variable equal to one of $t_E > t_H$ and else zero. Intuitively, we only need to buy back the shorted underlying asset if the put is not exercised against us in the holding period. If it is, we obtain the underlying asset from the put owner, saving us $\varphi \text{abs}(\Delta(t_H))\ BAS^s(t_H)$ in bid-ask trading costs. Conversely, the total costs originating from short-selling the underlying asset at the end of the holding period, $C^{BC}(t_0, t_H)$, are equal to:

$$C^{BC}(t_0, t_H) = \sum_{i=0}^{D} e^{r_f(i+1,D)} r^{bc}(i+1) \text{abs}(\Delta(i)) S(i), \tag{2.20}$$

where the sum is taken over all days within the holding period, $D$ is the number of days in that period, and $r^{bc}(i+1)$ is the daily stock-borrowing rate over day $i+1$. Subtracting the trading and stock-borrowing costs measured at the end of the holding period from the numerator of the unadjusted portfolio return, the adjusted portfolio return, $R^{X,tca}(t_0, t_H)$, is equal to:

$$R^{X,tca}(t_0, t_H) = \frac{X(t_H) - C^{TC}(t_0, t_H) - C^{BC}(t_0, t_H)}{p(t_0)}. \tag{2.21}$$

### 2.3.3 Calculating Theoretical Early Exercise Probabilities

To see whether real investors follow differentially suboptimal early exercise policies across different types of puts, we also contrast real-world early exercise probabilities with theoretical probabilities deduced from the optimal policies implied by the Black-Scholes (1973) model. To calculate the latter probabilities, we use Longstaff and Schwartz's (2001) least-squares approach. To be specific, we use a GBM to simulate $q$ underlying-asset-value paths under the $\mathbb{Q}$ measure (see Equation (2.1)), sampling the asset's value at times $t_0 < t_1 < \ldots < t_k = T$. We next move backward through the paths, starting with calculating the path-specific maturity payoff of the put, $\max(K - S(t_k), 0)$. Moving to time $t_{k-1}$, we compare each path's early exercise payoff at that time, $K - S(t_{k-1})$, with the put's continuation value, which we define as the fitted value from a regression of the put's maturity payoffs discounted to time $t_{k-1}$ on a function of the underlying asset value at time $t_{k-1}$.[7] If the early exercise payoff exceeds the continuation value, we assume that an early exercise occurs for that path and at that time, replacing the put's value with the early exercise payoff.

Moving back to time $t_{k-2}$, we again regress the put's future payoff discounted to time $t_{k-2}$ on the function of the underlying asset's value at time $t_{k-2}$. This time, however, the future payoff is either the earliest early exercise payoff (if the put is early exercised in the future) or the maturity payoff (if it is not). As before, we assume that an early exercise occurs for a path at that time if the early exercise payoff exceeds the continuation value. We continue in that way until we reach time $t_0$. We finally compute the *unconditional* theoretical early exercise probability over the put's time-to-maturity and imputed from the Black-Scholes (1973) model as the ratio of the number of paths over which the put is early exercised to the total number of paths $q$. Similarly, we compute the corresponding *conditional* theoretical probability over some window within the time-to-maturity as the ratio of the number of paths over which the put is first early exercised in that window to the total number of paths without an early exercise at the start of the window.

To implement the above methodology, we use 100,000 simulated underlying-asset-value paths for each put-month observation, with a number of time steps equal to the days-to-maturity of the put. We use a third-order polynomial to estimate the put's continuation value. While we directly observe the stock and strike price, the days-to-maturity, and the riskfree rate for each put-month observation, we estimate the underlying stock's volatility using monthly returns over the prior 60 months, allowing us to compute a forward-looking theoretical early exercise probability.

### 2.3.4 Data and Data Sources

We obtain daily data on American puts written on dividend-paying and non-dividend-paying single stocks and on those stocks from Optionmetrics. We retrieve riskfree rates of return from the zero-coupon yield curves also provided by Option-

[7] To avoid bias, we run the regression on only observations for which the put is in-the-money at time $t_{k-1}$.

metrics, and we always use that riskfree rate in our empirical work whose maturity date is closest to the date to which we want to compound or discount a cash flow. We obtain unique early exercise data manually collected from the Option Clearing Corporation's archives and containing the daily number of contracts exercised by put issue and owner (customers, market makers, and firms) from Bob Whaley.[8] Due to the availability of the early exercise data, our sample period is July 2001 to June 2014. We exclude the months November 2001, January 2002, July 2002, and January 2006 from that sample period because the early exercise data are consistently missing in these months. We further omit puts with a strike price-to-stock price ratio (moneyness) below 0.975 from our sample since such puts are hardly ever early exercised in real markets. We extract stock short-selling fees from Markit. We finally obtain the Fama-French benchmark factors, the VIX index, the TED spread, and a stock liquidity factor from Kenneth French's website, the CBOE website, the Fred Database, and Lubos Pastor's website, respectively.[9]

We apply standard filters to our data (see Goyal and Saretto (2009) and Cao and Han (2013)). To be specific, we exclude put-day observations for which the put price violates standard arbitrage bounds (as, e.g., that the put's price must lie below its strike price). We further omit observations (i) for which the put price is below \$1 or one-half the bid-ask spread; (ii) for which the bid-ask spread is negative; or (iii) for which the underlying stock's price is missing.

[8] We are grateful to Bob Whaley for sharing these data with us. The shared data are an updated version of the data also used in Pool et al. (2008), Barraclough and Whaley (2012), and Jensen and Pedersen (2016).

[9] The benchmark factors are obtainable from: <https://mba.tuck.dartmouth.edu/pages/faculty /ken.french/> and the stock liquidity factor from: <https://faculty.chicagobooth.edu/lubos.pastor /research/>. The URL for the VIX data are: <http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data)>, and for the TED spread: <https://fred.stlouisfed.org/series/TEDRATE>.

## 2.4 The Performance of the Risky Arbitrage Strategy

In this section, we examine the historical performance of the risky arbitrage strategy in the absence of transaction costs using the raw returns in Equations (2.11) and (2.17). We first look at the mean returns and return volatilities of the strategy and its two legs. We next adjust the mean returns for popular systematic factors, including firm-characteristic and macroeconomic factors.

### 2.4.1 Mean Return and Volatility of the Strategy

In Table 2.1, we offer descriptive statistics on the returns of our sample American put replication portfolios and traded American puts (columns (1) to (2), respectively), the spread across their returns (column (1)–(2)), and the moneyness and days-to-maturity of the puts (columns (3) to (4), respectively). We do not adjust returns for transaction costs, rebalance the replication portfolios daily, and define the strategy period to be one calendar month long. Each replication portfolio observation in column (1) corresponds to exactly one put observation in column (2). The descriptive statistics include the mean, the standard deviation (StDev), the Sharpe (1966) ratio, the mean's $t$-statistic (Mean/StError), several percentiles, and the number of observations. With the exception of the $t$-statistic and the Sharpe ratio, we calculate the statistics by sample month and then average over time. Given that, we can interpret the means in columns (1) and (2) as the mean returns of equally-weighted portfolios of, respectively, the replication portfolios and of the puts. We calculate the $t$-statistic as the mean divided by the product of the standard deviation and the square root of the number of observations and the Sharpe (1966) ratio as the mean excess return (i.e., the mean return minus the riskfree rate of return) divided by the standard deviation. We finally compute moneyness as the strike-to-stock price ratio and measure both moneyness as well as time-to-maturity at the start of the strategy period.

TABLE 2.1 ABOUT HERE

Table 2.1 suggests that the risky arbitrage strategy is highly profitable. While both legs of the strategy yield significantly negative mean monthly returns in columns (1) and (2), the mean return of the portfolio of replication portfolios (the long leg) is a less negative –9.08% ($t$-statistic: –4.09) compared to the –13.20% ($t$-statistic: –4.88) mean return of the put portfolio (the short leg). In turn, the mean spread return across them is 4.11% ($t$-statistic: 5.21) in column (1)–(2). Crucially, the spread return is much less volatile than the returns of the legs, as can be seen from the standard deviations and percentiles. While the annualized standard deviation of the spread portfolio is, for example, only 30.18%, those of the portfolio of replication portfolios and of the puts are 56.70% and 51.16%, respectively (compare columns (1), (2), and (1)–(2)). Using the three standard deviations, we can easily calculate the correlation between the two legs to be –0.85.[10] Indeed, Figure 2.1 graphically shows that the mean monthly holding period returns for the legs are heavily correlated over time. Nevertheless, despite being highly correlated, the replicating portfolio generates larger upward drift than the put option portfolio, as indicated in the cumulative profit graph in Figure 2.2, enough to make the long-short spread portfolio of our trading strategy a profitable one. The low volatility of the spread portfolio implies that it has a higher $t$-statistic than the two legs in absolute terms, despite it having a far less extreme mean return. It further implies that the spread portfolio has an impressive annualized Sharpe ratio of 1.64. The moneyness and days-to-maturity statistics in columns (3) and (4) suggest that the average put in our risky arbitrage strategy is in-the-money (moneyness: 1.07) and has slightly more than two months to maturity.

[10]To wit, $0.5670^2 + 0.5116^2 - 2 \times 0.8482 \times 0.5670 \times 0.5116 = 0.3018^2$.

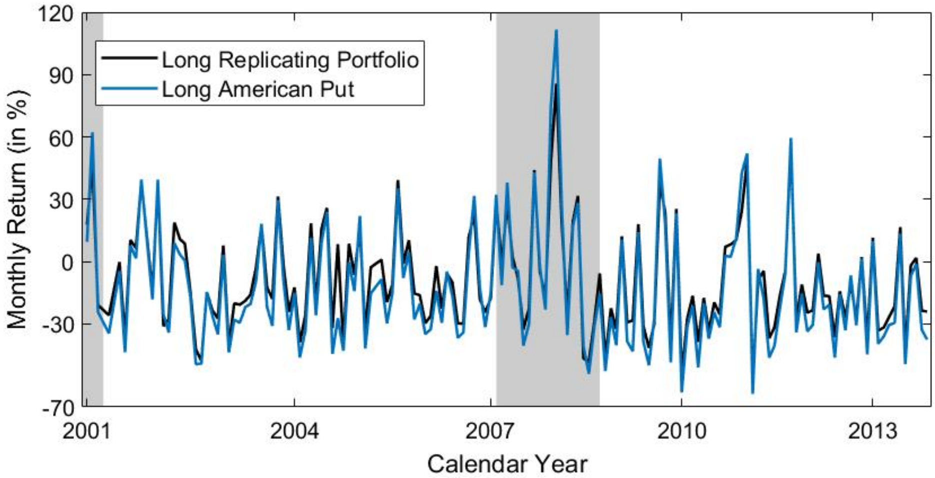


**Figure 2.1: Monthly Returns From Replicating Portfolio and Traded Put Positions**

The figure plots the mean monthly return from longing the equally-weighted replicating portfolio or traded American put portfolio. Option positions are taken at the start of each sample month and are held over the month. The grey areas are NBER recession periods.

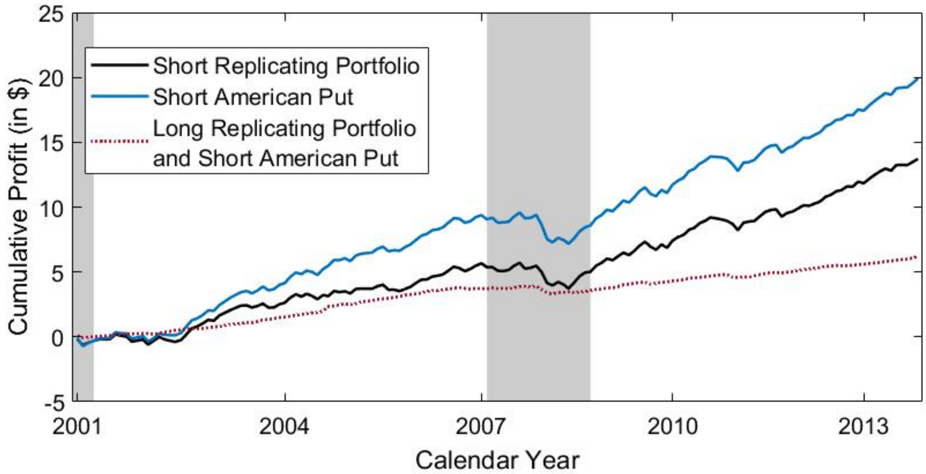


**Figure 2.2: Cumulative Profits From Replicating Portfolio and Traded Put Positions**

The figure plots the cumulative profits from shorting $1 of the equally-weighted replicating portfolio or traded American put portfolio or from longing $1 of the replicating and shorting the same amount of the traded put portfolio. Option positions are taken at the start of each sample month and are held over the month. The grey areas are NBER recession periods.

### 2.4.2 The Factor Model Alphas of the Strategy

In Table 2.2, we study the performance of the risky arbitrage strategy after taking its systematic risk into account. To do so, we use Black et al.'s (1972) time-series methodology and regress the monthly return on the spread portfolio long the equally-weighted portfolio of replication portfolios and short the equally-weighted put portfolio on risk factors, interpreting the constant ("alpha") from that regression as the risk-adjusted mean return of the strategy. In column (1), we use only the excess market return and a constant as exogenous variables. Column (2) adds Fama and French's (1993) benchmark factors SMB and HML, while column (3) also adds Carhart's (1997) MOM factor. Column (4) adds Fama and French's (2015) additional benchmark factors PRF and INV. Column (5) finally adds the change in the VIX index, the TED funding spread, and Pastor and Stambaugh's (2003) liquidity factor.[11] Plain numbers in the table are monthly premium estimates, while numbers in parentheses are $t$-statistics derived from Newey and West's (1987) formula with a twelve-month lag length.

Table 2.2 About Here

Table 2.2 suggests that, independent of the risk factors used in the regressions, the alpha of the risky arbitrage strategy is always significantly positive, with it, in fact, hardly differing from the mean strategy return. Considering the most comprehensive model in column (5), the alpha is, for example, 4.05% per month ($t$-statistic: 8.70) — virtually identical to the mean strategy return of 4.11% ($t$-statistic: 5.21) in Table 2.1. Notwithstanding, the strategy does load significantly on several risk factors due to us being unable to perfectly replicate puts in real markets. Columns (1) to (4), for example, suggest that the strategy loads positively and significantly on MKT, SMB, and MOM. Interestingly, however, the MKT, SMB, and MOM loadings seem mostly attributable to volatility and liquidity risk, as can be seen from column (5). Not

[11] The SMB factor is the return of a portfolio long small and short big stocks controlling for book-to-market, while the HML factor is the return of a portfolio long high book-to-market ("value") and short low book-to-market ("growth") stocks controlling for size. The MOM factor is the return of a portfolio long stocks with high returns over the recent past and short stocks with low returns over that period. The PRF factor is the return of a portfolio long more profitable and short less profitable stocks, while INV is the return of a portfolio long low-investment and short high-investment stocks, with both factors controlling for size. See Kenneth French's website for more details. The VIX index is a portfolio of options mimicking option-implied volatility, the TED spread is the difference between the interest rate on short-term U.S. government debt and on interbank loans, and the systematic liquidity factor is the return of a portfolio long stocks with a high liquidity exposure and short stocks with a low exposure. See Lubos Pastor's website for more details.

only do VIX and LIQ command significant loadings of, respectively, –0.16 ($t$-statistic: –4.67) and 0.44 ($t$-statistic: 3.65) in that column, but they also drive out the MKT, SMB, and MOM loadings. Although the risky arbitrage strategy thus loads on some risks, its loadings are markedly attenuated compared to those of its two legs. While its univariate MKT loading is, for example, 0.76 in column (1), the portfolio of replication portfolios (the long leg) and the put portfolio (the short leg) attract corresponding MKT loadings of –5.52 and –6.27, respectively (unreported to conserve space).

Taken together, the empirical results in this section suggest that the return of the risky arbitrage strategy is not spanned by those of other well-known trading strategies and that the strategy thus represents a novel trading opportunity for investors.

## 2.5 Conditioning the Strategy Return

In this section, we condition the performance of the risky arbitrage strategy on put, stock, and macroeconomic factors. We start with those factors for which we can deduce the sign of their conditioning effect from the theory in Section 2.2, which are the strike price and the interest rate. We then turn to factors which may condition the strategy's performance through them capturing variations in investors' tendency to early exercise puts too late, namely, moneyness, time-to-maturity, and stock volatility. We finally offer evidence suggesting that the conditioning ability of the latter factors indeed comes through suboptimally late early exercises, with us showing that investors do not seem to correctly condition their early exercise decisions on time-to-maturity and stock volatility. As in Section 2.4, we continue to ignore transaction costs.

### 2.5.1 Conditioning on the Strike Price and the Interest Rate

Assuming that American put investors can sometimes exercise their positions too late, our theory in Section 2.2 suggests that the performance of the risky arbitrage strategy is higher for high strike-price puts in high interest-rate regimes. Additionally, the theory in Section 2.2 shows that the returns on the strategy should not be conditional upon changes in the strike price. To test these predictions, we start with sorting our sample replication portfolios, their associated puts, and the spread portfolios long a replication portfolio and short its associated put into portfolios according to the quintile breakpoints of the strike price in month $t-1$. We label these portfolios our "unconditional strike price portfolios." Since the strike price may, however, relate to moneyness (defined as the strike-to-stock price ratio), possibly leading moneyness to confound the unconditional portfolio results, we also construct "strike-price portfolios controlling for moneyness" in the spirit of An et al. (2014). To do so, we further split each unconditional portfolio into portfolios based on the quintile breakpoints of moneyness in month $t-1$ and then form equally-weighted portfolios of those portfolios within the same strike-price classification. For both sets of portfolios, we equally-weight the portfolios, set up a spread portfolio long the top and short the bottom quintile, and hold the portfolios over month $t$.

Table 2.3 presents the results from the univariate portfolio exercise, with Panel A focusing on the unconditional portfolios and Panel B on those controlling for moneyness. Plain numbers are mean monthly returns, while those in square parentheses are $t$-statistics calculated from Newey and West's (1987) formula with a twelve-month lag length. Whilst the theory predicts no relationship between the strike price and the arbitrage strategy return, the results in Table 2.3 shows that the arbitrage return increases as the strike price increases, a result that is not consistent with the theory. Looking at the unconditional portfolios in Panel A, for example, the mean return of the strategy increases from 2.34% ($t$-statistic: 3.59) in the lowest strike-price quintile to 9.02% ($t$-statistic: 5.22) in the highest quintile. This increase in return is equal to a statistically significant 6.68% ($t$-statistic: 4.57), suggesting there is an anomaly in the relationship between the strike price and the return on the risky arbitrage strategy. This anomaly is also present in Panel B where we control for moneyness in the strike-price sorts. A preliminary investigation of the anomaly reveals a reversed J-shaped pattern in the relationship between the returns on the replicating portfolio (the long leg) and changes in the strike price. This pattern, however, is not present in the relationship between associated put returns (the short leg) and changes in the strike price. While the American put portfolio shows a steady decline in option returns as the strike price increases, the returns on the replicating portfolio abruptly increase, especially for the highest strike-price-sorted quintile. This creates a significantly higher arbitrage return for this quintile which is reflected in the very significant difference in returns on the "High" and "Low" portfolios. To explore whether this anomaly is due to the pattern we observe in the replicating portfolio returns, we calculate the difference in returns from a strategy long in the fourth highest (but not the highest) strike-price-sorted quintile and short in the lowest strike-price-sorted quintile in Panel A. The return on the "High–Low" portfolio from this strategy is a much lower and statistically insignificant 0.91% ($t$-statistic: 1.79) compared to the 6.68% return on the strategy when we use the highest quintile portfolio. This suggests that the anomaly mainly arises from the replicating portfolio return on the highest strike-price-sorted quintile. We plan to explore this anomaly in greater detail in future research.

TABLE 2.3 ABOUT HERE

We next condition on the interest rate on top of the strike price. To do so, we first remember that our theory in Section 2.2 suggests that the arbitrage profit is proportional to the interest rate times the strike price ($rK$) in a perfect market. We then recognize that the mean interest rate drops significantly from 2.83% per annum over the first half of our sample period (2001-2007) to only 0.27% over the second half (2008-2014). Combining these two observations, it becomes obvious that we can use a difference-in-difference (DID) approach to investigate the effect of the interest rate on the profitability of the risky arbitrage strategy, relying on the drop in the interest rate as shock variable and the strike price as treatment variable. To facilitate that approach, we first calculate the mean returns of the unconditional strike-price spread portfolios from Panel A of Table 2.3 separately for the 2001-2007 and 2008-2014 subsample periods. We then calculate the difference in mean returns for each portfolio over the two periods and finally the difference in those differences across the highest and lowest strike-price quintile ("DID estimate").

Table 2.4 presents the results from our DID tests studying how the interest rate conditions the profitability of the risky arbitrage strategy. As before, the plain numbers in the table are mean monthly returns, while those in square parentheses are $t$-statistics calculated from Newey and West's (1987) formula with a twelve-month lag length. In accordance with the theory, the table suggests that a higher interest rate indeed positively conditions the mean strategy return, with the positive effect, however, being amplified by the strike price. While the mean monthly return of the strategy is only an insignificant 0.75 percentage points ($t$-statistic: 0.79) higher in the high relative to the low interest rate subsample period in the lowest strike-price quintile, it is a significant 7.94 points ($t$-statistic: 4.87) higher in the highest quintile. The difference in these two percentage-point increases is a highly significant 7.18% ($t$-statistic: 5.38). It is important to note, however, that a part of this difference might be attributable to the strike-price and arbitrage return anomaly reported in Table 2.3. This is something we plan to investigate in future research.

TABLE 2.4 ABOUT HERE

### 2.5.2 Conditioning on Moneyness, Maturity Time, and Volatility

While the former subsection supports our theory by establishing that the risky arbitrage strategy is more profitable on higher strike-price puts in higher interest-rate regimes, it is conceivable that other factors condition the strategy's profitability if they capture variations in investor's tendency to exercise puts too late (i.e., if they predict $\max(t^E - \tau, 0)$). To study that possibility, we now condition the strategy's return on a parsimonious set of put and stock characteristics potentially capturing such variations, including moneyness, time-to-maturity, and stock volatility. We start with moneyness and time-to-maturity. At the end of each month $t-1$, we first sort the replication portfolios, the puts, and the spread portfolios into a deep in-the-money (DITM; strike-to-stock price above 1.10), in-the-money (ITM; 1.025-1.10), and at-the-money (ATM; 0.975-1.025) portfolio according to the relevant put's moneyness. We next independently sort them into a short (below 60 days), medium (60-90 days), and long (above 90 days) time-to-maturity portfolio according to the relevant put's time-to-maturity. The intersection of the two univariate portfolio sorts then yields $3 \times 3$ portfolios double-sorted on moneyness and time-to-maturity. We equally-weight the constituents of those portfolios. We finally hold the portfolios over month $t$.

Table 2.5 presents the results from the double-sorted portfolio exercise, with columns (1), (2), and (1)–(2) focusing on the put replication portfolios, the puts, and the spread portfolios long a replication portfolio and short the associated put, respectively. In turn, Panels A, B, and C look into DITM, ITM, and ATM put strategies, respectively. As before, plain numbers are mean monthly portfolio returns, whereas the numbers in parentheses are $t$-statistics calculated from Newey and West's (1987) formula with a twelve-month lag length. Column (1)–(2) in the table offers strong evidence that the profitability of the risky arbitrage strategy decreases in both moneyness and time-to-maturity. Looking at strategies based on 30-60 day puts, the mean strategy return, for example, decreases from 9.66% ($t$-statistic: 6.53) in the ATM-put portfolio in Panel C to 2.49% ($t$-statistic: 5.28) in the DITM-put portfolio in Panel A. Conversely, looking at strategies based on ITM puts in Panel B, that same return decreases from 4.94% ($t$-statistic: 6.03) in the 30-60 day put portfolio to 2.58% ($t$-statistics: 4.16) in the 90-120 day put portfolio. Turning to the underlying replication portfolios and puts in columns (1) and (2), their mean returns drop with moneyness but rise with time-to-maturity, in line with the put results in Aretz and Gazi (2020).

<u>Table 2.5 About Here</u>

As a next step, we condition the performance of the risky arbitrage strategy on the idiosyncratic volatility of the underlying stocks. At the end of each month $t-1$, we thus split the replication portfolios, puts, and the spread portfolios long a replication portfolio and short the associated put into portfolios according to the quintile breakpoints of that volatility estimated using the market model or the Fama-French-Carhart (FFC; 1997) model. We can write the market model as:

$$R_{i,\tau} = \alpha_i + \beta_i^{mkt}(R_\tau^{mkt} - R_{f\tau}) + \epsilon_{i,\tau}, \tag{2.22}$$

where $R_{i,\tau}$ is stock $i$'s return over month $\tau$, $R_\tau^{mkt} - R_{f\tau}$ is the excess market return, $\alpha_i$ and $\beta_i^{mkt}$ are parameters, and $\epsilon_{i,\tau}$ is the residual. We can write the FFC model as:

$$R_{i,\tau} = \alpha_i + \beta_i^{mkt}(R_\tau^{mkt} - R_{f\tau}) + \beta_i^{smb} R_\tau^{smb} + \beta_i^{hml} R_\tau^{hml} + \beta_i^{mom} R_\tau^{mom} + \epsilon_{i,\tau}, \tag{2.23}$$

where $R_\tau^{smb}$, $R_\tau^{hml}$, and $R_\tau^{mom}$ are the returns of spread portfolios on size, the book-to-market ratio, and the eleven-month (momentum) past return, respectively, and $\beta_i^{smb}$, $\beta_i^{hml}$, and $\beta_i^{mom}$ are additional parameters. We estimate both models over the prior 60 months of monthly data, calculating idiosyncratic volatility as the standard deviation of the residual, $\epsilon_{i,\tau}$. We equally-weight the constituents of the quintile portfolios and hold them over month $t$.

Table 2.6 presents the results from the univariate portfolio exercise, with Panels A and B using market- and FFC-model estimates to proxy for idiosyncratic stock volatility, respectively. As before, plain numbers are mean monthly returns, whereas numbers in parentheses are Newey-West (1987) $t$-statistics. The table offers strong evidence that the profitability of the risky arbitrage strategy deteriorates with idiosyncratic stock volatility. Looking at the market-model sorts in Panel A, the mean strategy return in the third row, for example, drops from 6.07% ($t$-statistic: 4.70) in the low-volatility portfolio to 2.81% ($t$-statistic: 5.36) in the high-volatility portfolio. The difference in those two numbers is a highly significant –3.26% ($t$-statistic: –3.19). Turning to the replication portfolios and puts in the first and second rows of each panel, their mean returns increase with idiosyncratic stock volatility, aligning with Aretz and Gazi's (2020) put results.

TABLE 2.6 ABOUT HERE

### 2.5.3 Jointly Conditioning on All Factors

We next investigate how the strike price, moneyness, time-to-maturity, and stock volatility jointly condition the performance of the risky arbitrage strategy and verify that our portfolio results are robust to variations in methodology. To do so, we run Fama-MacBeth (FM; 1973) regressions of the returns on the replication portfolios, the puts, and the spread portfolios long a replication portfolio and short its associated put over month $t$ on combinations of those conditioning factors calculated using only data until the end of month $t-1$. To mitigate outlier effects, we use the log strike price (instead of the strike price) and time-to-maturity stated as fraction of a year in the regressions. As before, moneyness is the strike-to-stock price ratio, while stock volatility is the annualized FFC idiosyncratic volatility estimate, introduced in Section 2.5.2.

Table 2.7 gives the regression results, with Panels A, B, and C using the spread portfolio return, replication portfolio return, and put return as dependent variable, respectively. Plain numbers are monthly premium estimates, whereas the numbers in parentheses are Newey and West (1987) $t$-statistics with a twelve-month lag length. The table shows that the regressions yield results almost exactly identical with those from the portfolio exercises. To be specific, column (1) in Panel A reveals that the unconditional mean spread return is 4.11% ($t$-statistic: 7.33) per month, aligning with column (1)–(2) in Table 2.1. Moreover, columns (2)–(5) in that panel confirm that the mean spread return significantly rises in the strike price but significantly drops in moneyness, time-to-maturity, and stock volatility, at least when strike price and stock volatility are not jointly included as independent variables. Interestingly, when we jointly include the strike price and stock volatility, as we do in column (6) of Panel A, the strike-price premium hardly changes, while the stock volatility premium switches from being significantly negative to significantly positive. The reason for this unexpected outcome is that the strike price and stock volatility are negatively correlated, with higher volatility stocks often having many low strike-price puts written on them.

TABLE 2.7 ABOUT HERE

Turning to the replication portfolio and put results in Panels B and C of Table 2.7, we notice that their monthly premiums on the strike price, moneyness, time-to-maturity, and stock volatility are of exactly the same sign as those obtained in the portfolio exercises.

### 2.5.4 Why Do Moneyness, Time-to-Maturity, and Stock Volatility Condition the Success of the Risky Arbitrage Strategy?

While we argue that the ability of moneyness, time-to-maturity, and stock volatility to condition the risky arbitrage strategy in Sections 2.5.2 and 2.5.3 is due to those factors capturing variations in investors' tendency to exercise puts too late, there may be other reasons for that ability. To offer more support for the hypothesis that suboptimally late exercises do indeed lie behind that ability, we next take a closer look at real investors' early exercise behavior. We start with benchmarking the real-world early exercise probabilities of our sample puts against their corresponding optimal probabilities deduced from the Black and Scholes (1973) model. We calculate the probabilities as described in Sections 2.3.1 and 2.3.3, respectively. While we acknowledge that the shortcomings of the Black and Scholes (1973) model imply that the optimal probabilities deduced from it differ from the true optimal probabilities, we nonetheless hope to learn some broader lessons about the optimality of real investor's early exercise strategies from the comparisons.[12]

[12]It is not obvious to us how to calculate more accurate theoretical early exercise probabilities. Switching to a more sophisticated option pricing model, as, for example, the Heston (1993) model, would require us to estimate additional parameters governing, for example, the mean re-

To take a first look at how the real-world and Black-Scholes (1973) early exercise probabilities of our sample puts over their entire times-to-maturity are related, Table 2.8 reports their mean values over portfolios formed according to the decile breakpoints of the theoretical probabilities at the start of the strategy return period. We calculate the mean values first by cross-section and then average over time. The table reveals that the real-world and the Black-Scholes (1973) probabilities are strongly positively correlated. While the monotonic increase in the mean Black-Scholes (1973) probabilities over the portfolios from 16.35% to 76.74% is by construction, the mean real-world probabilities, remarkably, also monotonically increase over them, from 6.27% to 27.05%, sharing an average cross-sectional correlation of 0.23 between the two probabilities. Notwithstanding, the mean real-world probabilities are consistently only around one-third of the mean Black-Scholes (1973) probabilities, suggesting that real-world put investors often wait too long with exercising their positions and that the necessary condition for a risky arbitrage profit to exist is fulfilled. While relaxing certain Black-Scholes (1973) assumptions (as, e.g., the constant volatility and/or no asset-value jumps assumptions) could help to close the gaps between the probabilities, we deem it unlikely that it can account for the entire gaps.

TABLE 2.8 ABOUT HERE

We next investigate how the differences between the real-world and Black-Scholes (1973) early exercise probabilities relate to moneyness, time-to-maturity, and stock volatility. To do so, Table 2.9 reports the mean differences in those probabilities calculated over the strategy return period for portfolios triple-sorted according to those characteristics. We construct the portfolios as follows. At the end of month $t-1$, we independently sort our puts according to the same moneyness, time-to-maturity, and FFC idiosyncratic stock volatility breakpoints as in Tables 2.5 and 2.6, using the intersections of the univariate portfolios to create

version in volatility, the long-run volatility, the correlation between asset value and volatility, etc. Unfortunately, these additional parameters are tremendously difficult to estimate, especially at the stock level at which we only have limited amounts of data.

the triple-sorted portfolios. As before, we calculate the mean probabilities first by cross-section and then average over time.

Table 2.9 About Here

The table suggests that the difference between the probabilities tends to be more pronounced for shorter time-to-maturity puts written on lower volatility stocks. Looking at ITM puts on stocks with a third-quintile volatility in Panel B, the mean difference is, for example, 24.20% for 30-60 day puts but only 3.58% for 90-120 day puts. Conversely, looking at ITM puts with 60-90 days-to-maturity in the same panel, the mean difference is 22.96% for puts written on the lowest-volatility stocks but only 4.22% for those written on the highest-volatility stocks. The upshot is that real investors seem particularly bad in timely exercising short time-to-maturity puts on low-volatility stocks, which may explain why the risky arbitrage strategy works better on such puts. Interestingly, however, the table further suggests that the difference in the probabilities also tends to be more pronounced for higher moneyness puts. Looking at 30-60 day puts on third-quintile volatility stocks, the mean difference is, for example, 36.66% on DITM puts but only 12.07% on ATM puts. The implication is that variations in real investors' tendency to early exercise puts too late does not help to explain why the risky arbitrage strategy works better on lower-moneyness puts.

In Table 2.10, we switch to FM regressions to find out how real investors' tendency to exercise puts too late varies with moneyness, time-to-maturity, and stock volatility. While the dependent variable is the difference between the Black-Scholes (1973) and real-world early exercise probabilities, the independent variables are the characteristics defined as in Table 2.7. The regressions deliver results in alignment with the portfolio exercise results, showing that the difference in the probabilites tends to rise with moneyness and to drop with time-to-maturity and stock volatility.

Table 2.10 About Here

## 2.6 Adjusting for Transaction Costs

In this section, we adjust the profitability of the risky arbitrage strategy for bid-ask and short-selling transaction costs. As already said, it is crucial to adjust for those costs since the strategy involves potentially frequent stock purchases and sales and shorting the stock. In addition, Goyal and Saretto (2009) and Cao and Han (2013) show that high bid-ask costs in the options market greatly eat into the profitability of option trading strategies, often rendering profits insignificant or even negative. To account for transaction costs, we switch from studying the raw returns in Equations (2.11) and (2.17) to studying the transaction cost adjusted returns in Equations (2.12) and (2.21), setting $\varphi$, the proportion of the bid-ask spread representing trading costs, to either zero, 0.10, 0.25, or 0.50. When $\varphi = 0.50$, investors buy at the ask price and sell at the bid price.

Table 2.11 reevaluates the profitability of the double-sorted moneyness and time-to-maturity portfolios originally studied in Table 2.5 under transaction costs, reporting, however, only the mean strategy return (and not the mean leg returns). While Panel A looks into our full sample, Panel B focuses only on strategies executed on liquid assets. A liquid stock (put) is defined as one with an Amihud (2002) stock illiquidity estimate (bid-ask spread scaled by put price) below the median, with the liquidity proxies measured at the start of the strategy period. Conversely, columns (1) to (4) consider the $\varphi =$ zero, $0.10, 0.25,$ and $0.50$ case, with the final three columns (but not the first) also adjusting for shorting costs. As in Table 2.5, plain numbers are mean monthly returns, while those in parentheses are Newey-West (1987) $t$-statistics. Since the Markit short-selling data are available from only January 2002, our empirical work adjusting for transaction costs relies on the sample period from that date to June 2014 (leading us to lose five sample months).

TABLE 2.11 ABOUT HERE

In line with expectations, Panel A of Table 2.11 confirms that adjusting for transaction costs greatly reduces the profitability of our risky arbitrage strategy. Starting with column (1) in that panel, it is obvious that the slight change in our sample period does not materially affect mean strategy returns (compare the column with column (1)–(2) in Table 2.5). In contrast, assuming that investors incur trading costs $\varphi$ equal to 10% of an asset's bid-ask spread plus stock short-selling costs, column (2) reveals that the mean strategy return remains significantly positive only for strategies based on 30-60 day ITM or ATM puts (see Panels A.2 and A.3). In all other cases, it is either insignificant or, in one case, even significantly negative. Assuming that investors incur even higher trading costs (i.e., $\varphi = 0.25$ or 0.50), columns (3) and (4) finally show that the mean strategy return becomes highly significantly negative in the vast majority of cases.

Fortunately, Panel B of Table 2.11 suggests that the picture markedly improves once we restrict our attention to strategies executed on liquid assets. Interestingly, column (1) in that panel shows that the mean strategy return is generally higher in case of such strategies, even in the absence of transaction cost adjustments. More importantly, however, columns (2) to (4) indicate that the mean strategy return remains positive and significant even under the assumption that investors incur a 10% (25%) [50%] bid-ask trading cost plus stock short-selling costs in case of nine (six) [one] out of the nine strategies conditioned on moneyness and time-to-maturity.

In Table 2.12, we aim to decrease transaction costs further by reducing the frequency with which we rebalance the put replication portfolio, saving us costs arising from stock purchases and sales over the strategy return period. The downside is that a lower rebalancing frequency decreases the ability of a replication portfolio to track its associated put, boosting the volatility of the strategy return and making the strategy diverge even more from a textbook arbitrage strategy. To study the effects of a lower rebalancing frequency, Panels A, B, and C of Table 2.12 present the mean transaction cost adjusted returns of strategies executed on only liquid assets and using daily, weekly, and no rebalancing, respectively. The no-rebalancing case is identical to the "buy-and-hold delta-hedging strategy" of Goyal and Saretto (2009), who form replication portfolios at the start of the strategy period and hold those over the entire period. Within each panel, we consider strategies involving only puts with a price above \$1, \$2, and \$5 at the start of the strategy return period.

Table 2.12 About Here

The table suggests that a lower rebalancing frequency helps to further increase the profitability of the risky arbitrage strategy. While, surprisingly, a lower rebalancing frequency often also boosts the mean strategy return and its $t$-statistic in the absence of transaction cost adjustments (see column (1)), the improvements are far more pronounced in their presence. Looking at the case in which investors incur a 50% bid-ask trading cost plus stock short-selling costs in column (4), Panel A shows that only a single mean strategy return out of three is positive and weakly significant (with a $t$-statistic of 2.04) under daily rebalancing. In contrast, Panel B suggests that the corresponding mean returns under no rebalancing are significantly positive with $t$-statistics above 3.88.

Overall, this section shows that the risky arbitrage strategy does not only exist in theory but can profitably be exploited by real investors, at least when executed on liquid assets.

## 2.7 Concluding Remarks

We evaluate the profitability of a risky arbitrage strategy exploiting evidence that real put investors often exercise their positions far too late. Grounded in mathematical finance theory, the strategy consists of a long position in a stock-and-riskfree-asset portfolio replicating a put and a short position in the put. Our empirical work suggests that the strategy yields a highly significant positive mean return, with a low total as well as systematic risk. Consistent with theory, the mean strategy return rises with both the strike price and the interest rate. Interestingly, however, it drops with moneyness, time-to-maturity, and stock volatility, aligning with further evidence that investors do not appear to understand the implications of those characteristics for the optimal early exercise decision. Crucially, we finally show that the strategy survives accounting for both bid-ask trading costs and shorting costs, at least when it is executed on liquid assets. The upshot is that the strategy does not only exist on paper but can be pursued by real investors.

# References


Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5(1), pp.31-56.

An, B.J., Ang, A., Bali, T.G. and Cakici, N., 2014. The joint cross section of stocks and options. *The Journal of Finance*, 69(5), pp.2279-2337.

Aretz, K. and Gazi, A., 2020. The early exercise risk premium. *Working Paper*.

Barraclough, K. and Whaley, R.E., 2012. Early exercise of put options on stocks. *The Journal of Finance*, 67(4), pp.1423-1456.

Battalio, R., Figlewski, S. and Neal, R., 2020. Option investor rationality revisited: The role of exercise boundary violations. *Financial Analysts Journal*, 76(1), pp.82-99.

Bauer, R., Cosemans, M. and Eichholtz, P., 2009. Option trading and individual investor performance. *Journal of Banking & Finance*, 33(4), pp.731-746.

Black, F., Jensen, M.C., and Scholes, M.S., 1972. The capital asset pricing model: Some empirical tests. *Studies in the Theory of Capital Markets*, New York: Praeger, pp.79-121.

Black, F. and Scholes, M.S., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), pp.637-654.

Brennan, M.J. and Schwartz, E.S., 1977. The valuation of American put options. *The Journal of Finance*, 32(2), pp.449-462.

Cao, J. and Han, B., 2013. Cross section of option returns and idiosyncratic stock volatility. *Journal of Financial Economics*, 108(1), pp.231-249.

Cao, J., Han, B., Tong, Q. and Zhan, X., 2020. Option return predictability. *Working Paper*.

Carhart, M.M., 1997. On persistence in mutual fund performance. *The Journal of Finance*, 52(1), pp.57-82.

Coval, J.D. and Shumway, T., 2001. Expected option returns. *The Journal of Finance*, 56(3), pp.983-1009.

Fama, E.F. and French, K.R., 1992. The cross-section of expected stock returns. *The Journal of Finance*, 47(2), pp.427-465.

Fama, E.F., and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), pp.3-56.

Fama, E.F. and French, K.R., 2015. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), pp.1-22.

Fama, E.F. and MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), pp.607-636.

Figlewski, S., 2019. An American call is worth more than a European call: The value of American exercise when the market is not perfectly liquid. *Working Paper*.

Finucane, T.J., 1997. An empirical analysis of common stock call exercise: A note. *Journal of Banking & Finance*, 21(4), pp.563-571.

Gondzio, J., Kouwenberg, R. and Vorst, T., 2003. Hedging options under transaction costs and stochastic volatility. *Journal of Economic Dynamics and Control*, 27(6), pp.1045-1068.

Goyal, A., and Saretto, A., 2009. Cross-section of option returns and volatility. *Journal of Financial Economics*, 94(2), pp.310-326.

Heston, S.L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2), pp.327-343.

Jacka, S.D., 1991. Optimal stopping and the American put. *Mathematical Finance*, 1(2), pp.1-14.

Jensen, M.V., and Pedersen, L.H., 2016. Early option exercise: Never say never. *Journal of Financial Economics*, 121(2), pp.278-299.

Karatzas, I., 1988. On the pricing of American options. *Applied Mathematics and Optimization*, 17(1), pp.37-60.

Lakonishok, J., Shleifer, A. and Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. *The Journal of Finance*, 49(5), pp.1541-1578.

Leland, H.E., 1985. Option pricing and replication with transactions costs. *The Journal of Finance*, 40(5), pp.1283-1301.

Longstaff, F.A. and Schwartz, E.S., 2001. Valuing American options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 14(1), pp.113-147.

Merton, R.C., 1973. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1), pp.141-183.

Newey, W.K., and West, K.D., 1987. A simple positive-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, pp.703-708.

Overdahl, J.A. and Martin, P.G., 1994. The exercise of equity options: Theory and empirical tests. *The Journal of Derivatives*, 2(1), pp.38-51.

Pástor, L., and Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3), pp.642-685.

Perrakis, S. and Lefoll, J., 2000. Option pricing and replication with transaction costs and dividends. *Journal of Economic Dynamics and Control*, 24(11-12), pp.1527-1561.

Pool, V.K., Stoll, H.R. and Whaley, R.E., 2008. Failure to exercise call options: An anomaly and a trading game. *Journal of Financial Markets*, 11(1), pp.1-35.

Poteshman, A.M. and Serbin, V., 2003. Clearly irrational financial market behavior: Evidence from the early exercise of exchange traded stock options. *The Journal of Finance*, 58(1), pp.37-70.

Sharpe, W.F., 1966. Mutual fund performance. *The Journal of Business*, 39(1), pp.119-138.

Shreve, S.E., 2004. *Stochastic calculus for finance II: Continuous-time models* (Vol. 11). Springer Science & Business Media.

Vasquez, A., 2017. Equity volatility term structures and the cross section of option returns. *Journal of Financial and Quantitative Analysis*, 52(6), pp.2727-2754.

## Table 2.1: The Performance of the Risky Arbitrage Strategy

The table presents descriptive statistics for the monthly returns on daily-rebalanced stock-and-riskfree-asset portfolios replicating American puts (column (1)), on the puts (column (2)), and on spread portfolios long a replication portfolio and short the corresponding put (column (1)–(2)). It further presents descriptive statistics on the moneyness (column (3)) and time-to-maturity (column (4)) of the puts. The descriptive statistics are the mean, the standard deviation (StDev), the $t$-statistic for the mean (*Mean/StError*), the monthly and annual Sharpe (1966) ratio (Monthly and Annual Sharpe Ratio, respectively), several percentiles, and the number of observations. Each observation used in column (1) corresponds to one observation used in column (2). We calculate moneyness as the strike-to-stock price ratio and time-to-maturity as the number of calendar days until maturity, both at the start of the return period. With the exception of the $t$-statistic and the Sharpe ratio, we calculate each statistic as the time-series mean of the cross-sectional statistic. The $t$-statistic is the mean scaled by the standard error of the mean. The Sharpe ratio is the difference between mean return and riskfree rate scaled by the standard deviation of the mean in column (1) and (2), and the mean scaled by the standard deviation of the mean in column (1)–(2).

| | Monthly Replication Portfolio Return (in %) | Monthly American Put Option Return (in %) | Monthly Spread Portfolio Return (in %) | Money-ness | Days to Maturity |
|---|---|---|---|---|---|
| | (1) | (2) | (1)–(2) | (3) | (4) |
| Mean | −9.08 | −13.20 | 4.11 | 1.07 | 73 |
| StDev | 56.70 | 51.16 | 30.18 | | |
| Mean/StError | [−4.09] | [−4.88] | [5.21] | | |
| Monthly Sharpe Ratio | | | 0.47 | | |
| Annual Sharpe Ratio | | | 1.64 | | |
| Percentile 1 | −105.50 | −90.63 | −39.39 | 0.98 | 49 |
| Percentile 5 | −80.26 | −79.94 | −17.27 | 0.99 | 49 |
| Quartile 1 | −43.35 | −48.79 | −2.87 | 1.02 | 50 |
| Median | −14.99 | −20.04 | 2.72 | 1.07 | 62 |
| Quartile 3 | 15.89 | 12.17 | 8.73 | 1.12 | 96 |
| Percentile 95 | 78.65 | 76.95 | 21.66 | 1.18 | 111 |
| Percentile 99 | 169.58 | 152.01 | 67.66 | 1.20 | 111 |
| Observations | 5,612 | 5,612 | 5,612 | 5,612 | 5,612 |

**Table 2.2: Adjusting the Strategy's Performance for Systematic Risk**

The table shows the results from time-series regressions of the month-$t$ return of a spread portfolio long a stock-and-riskfree-asset portfolio replicating a put and short the put on several sets of benchmark factors measured over the same month and a constant. The factors include the market return minus the risk-free rate of return (MKT); the return of a spread portfolio long small and short large stocks (SMB); the return of a spread portfolio long value and short growth stocks (HML); the return of a spread portfolio long winner and short loser stocks (MOM); the return of a spread portfolio long profitable and short unprofitable stocks (PRF); the return of a spread portfolio long non-investing and short investing stocks (INV); the change in the VIX option-implied volatility index (VIX); the three-month LIBOR rate minus the Treasury bill rate (TED); and the return of a spread portfolio long high-liquidity and short low-liquidity stocks (LIQ). Plain numbers are estimates, while the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length.

| | Time-Series Regression Model: | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| MKT | 0.76 | 0.67 | 0.82 | 0.81 | 0.03 |
| | [6.89] | [5.67] | [6.31] | [5.63] | [0.17] |
| SMB | | 0.48 | 0.45 | 0.48 | 0.34 |
| | | [2.19] | [2.09] | [2.21] | [1.74] |
| HML | | −0.03 | 0.03 | 0.15 | 0.64 |
| | | [−0.12] | [0.14] | [0.65] | [3.04] |
| MOM | | | 0.27 | 0.27 | 0.16 |
| | | | [2.54] | [2.41] | [1.58] |
| PRF | | | | 0.08 | −0.25 |
| | | | | [0.26] | [−0.96] |
| INV | | | | −0.46 | −0.52 |
| | | | | [−1.42] | [−1.85] |
| VIX | | | | | −0.16 |
| | | | | | [−4.67] |
| TED | | | | | −3.03 |
| | | | | | [−1.72] |
| LIQ | | | | | 0.44 |
| | | | | | [3.65] |
| Constant | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| | [7.51] | [7.25] | [7.18] | [6.96] | [8.70] |

## Table 2.3: Conditioning the Strategy on the Strike Price

The table presents the mean returns of portfolios of stock-and-riskfree-asset portfolios replicating a put and of the associated puts sorted on the strike price. It further presents the mean returns of the corresponding spread portfolios long the portfolios of put replication portfolios and short the put portfolios. At the end of each sample month $t-1$, we first sort each of those assets into portfolios according to the quintile breakpoints of the associated strike price, without controlling for the strike-to-stock price ratio ("moneyness;" Panel A). Within each strike price portfolio, we next sort them into further portfolios according to the quintile breakpoints of moneyness in month $t-1$. We then form equally-weighted portfolios of those portfolios within the same strike price classification, averaging out the effect of moneyness (Panel B). We also form a spread portfolio long the top and short the bottom quintile ("High–Low"). We equally-weight the portfolios and hold them over month $t$. The put and replication portfolio observations are matched, so each put observation corresponds to one replication portfolio observation. Plain numbers are mean monthly portfolio returns (in %) and the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length.

| | Strike Price, $K$ | | | | | |
|---|---|---|---|---|---|---|
| | 1 (Low) | 2 | 3 | 4 | 5 (High) | High–Low |
| Panel A: Univariate Strike-Price Portfolios | | | | | | |
| Replication Portfolio | −8.52 | −10.89 | −11.83 | −11.60 | −6.32 | 2.21 |
| | [−4.19] | [−4.95] | [−5.28] | [−4.89] | [−2.26] | [1.26] |
| American Put Option | −10.86 | −12.87 | −14.22 | −14.85 | −15.33 | −4.47 |
| | [−4.31] | [−4.66] | [−5.06] | [−5.07] | [−5.06] | [−2.80] |
| Spread Portfolio | 2.34 | 1.99 | 2.40 | 3.25 | 9.02 | 6.68 |
| | [3.59] | [2.98] | [3.68] | [4.70] | [5.22] | [4.57] |
| Panel B: Strike-Price Portfolios Controlling for Moneyness | | | | | | |
| Replication Portfolio | −8.49 | −10.72 | −11.73 | −11.51 | −6.23 | 2.26 |
| | [−4.20] | [−4.90] | [−5.26] | [−4.86] | [−2.23] | [1.27] |
| American Put Option | −10.87 | −12.74 | −14.18 | −14.81 | −15.34 | −4.47 |
| | [−4.33] | [−4.63] | [−5.05] | [−5.06] | [−5.06] | [−2.77] |
| Spread Portfolio | 2.38 | 2.02 | 2.44 | 3.29 | 9.10 | 6.72 |
| | [3.62] | [3.01] | [3.72] | [4.72] | [5.24] | [4.58] |

## Table 2.4: Conditioning the Strategy on the Strike Price and Interest Rate

The table presents the mean returns of spread portfolios long an equally-weighted portfolio of put replication portfolios and short an identically-weighted portfolio of the associated puts sorted on the strike price and separately calculated over the July-2001 to December-2007 (column (1)) and the January-2008 to June-2014 (column (2)) subsample periods. The table also reports the differences in mean spread portfolio returns across the subsample periods (column (1)–(2)). At the end of each sample month $t-1$ within a subsample period, we first sort the spread portfolios into portfolios according to the quintile breakpoints of the strike price of the associated put. We also form a spread portfolio long the top and short the bottom quintile ("High–Low"). We equally-weight the portfolios and hold them over month $t$. The replication portfolio and put observations are matched, so that each replication portfolio observation corresponds to one put observation. Plain numbers are mean monthly portfolio returns (in %), and the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length.

| Strike Price, $K$ | Until 2007 | From 2008 | Difference |
|---|---|---|---|
| | (1) | (2) | (1)–(2) |
| 1 (Low) | 2.75 | 1.99 | 0.75 |
| | [5.95] | [1.85] | [0.79] |
| 2 | 2.00 | 1.98 | 0.03 |
| | [4.57] | [1.74] | [0.03] |
| 3 | 2.57 | 2.24 | 0.33 |
| | [5.55] | [2.04] | [0.32] |
| 4 | 3.72 | 2.85 | 0.87 |
| | [6.48] | [2.52] | [0.81] |
| 5 (High) | 13.32 | 5.38 | 7.94 |
| | [6.33] | [3.97] | [4.87] |
| High–Low | 10.57 | 3.39 | 7.18 |
| | [5.39] | [4.90] | [5.38] |

## Table 2.5: Conditioning the Strategy on Moneyness and Maturity Time

The table presents the mean returns of portfolios of stock-and-riskfree-asset portfolios replicating a put (column (1)) and of the associated puts (column (2)) sorted on the puts' moneyness and time-to-maturity. It further presents the mean returns of the corresponding spread portfolios long the portfolios of put replication portfolios and short the put portfolios (column (1)–(2)). At the end of each sample month $t-1$, we first sort each of these assets into portfolios according to whether the strike-to-stock price ratio ("moneyness") of the associated put lies above 1.10 (Panel A), between 1.025 and 1.10 (Panel B), or between 0.975 and 1.025 (Panel C). Within each moneyness portfolio, we next sort them into portfolios according to whether their days-to-maturity are below 60, between 60 and 90, or above 90 days. We equally-weight the portfolios and hold them over month $t$. The observations used in columns (1) and (2) are matched, so that each observation in column (1) corresponds to one observation in column (2). Plain numbers are mean monthly portfolio returns (in %) and the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a lag length of twelve months.

| Days-to-Maturity | Monthly Replication Portfolio Return (in %) | Monthly American Put Option Return (in %) | Monthly Spread Portfolio Return (in %) |
|---|---|---|---|
| | (1) | (2) | (1)–(2) |
| **Panel A: Deep In-The-Money (Strike-to-Stock Price > 1.10)** | | | |
| 30-60 | −15.31 | −17.80 | 2.49 |
| | [−7.61] | [−7.94] | [5.28] |
| 60-90 | −8.67 | −10.46 | 1.79 |
| | [−5.17] | [−5.31] | [3.72] |
| 90-120 | −5.80 | −7.28 | 1.48 |
| | [−3.61] | [−3.85] | [3.90] |
| **Panel B: In-The-Money (Strike-to-Stock Price 1.025 to 1.10)** | | | |
| 30-60 | −13.33 | −18.27 | 4.94 |
| | [−5.23] | [−6.08] | [6.03] |
| 60-90 | −7.10 | −10.42 | 3.32 |
| | [−3.29] | [−3.92] | [3.95] |
| 90-120 | −4.58 | −7.16 | 2.58 |
| | [−2.30] | [−2.90] | [4.16] |
| **Panel C: At-The-Money (Strike-to-Stock Price 0.975 to 1.025)** | | | |
| 30-60 | −7.36 | −17.01 | 9.66 |
| | [−2.49] | [−4.55] | [6.53] |
| 60-90 | −5.83 | −11.19 | 5.36 |
| | [−2.17] | [−3.34] | [3.92] |
| 90-120 | −4.43 | −8.03 | 3.60 |
| | [−1.86] | [−2.65] | [4.12] |

## Table 2.6: Conditioning the Strategy on Idiosyncratic Stock Volatility

The table presents the mean returns of portfolios of stock-and-riskfree-asset portfolios replicating a put and of the associated puts sorted on the idiosyncratic volatility of the stock. It further presents the mean returns of the corresponding spread portfolios long the portfolios of put replication portfolios and short the put portfolios. At the end of each sample month $t-1$, we sort each of these assets into portfolios according to the quintile breakpoints of a stock's market-model (Panel A) or Fama-French-Carhart-model (Panel B) idiosyncratic volatility. We estimate the models over the prior 60 months, defining idiosyncratic volatility as the volatility of the residual. We also form a spread portfolio long the top and short the bottom quintile ("High–Low"). We equally-weight the portfolios and hold them over month $t$. The replication portfolio observations and put observations are matched, so that each replication portfolio observation corresponds to one put observation. Plain numbers are mean monthly portfolio returns (in %), and the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length.

| | Idiosyncratic Stock Volatility | | | | | |
|---|---|---|---|---|---|---|
| | 1 (Low) | 2 | 3 | 4 | 5 (High) | High–Low |
| **Panel A: Market Model Idiosyncratic Volatility** | | | | | | |
| Replication Portfolio | −8.99 | −9.70 | −9.92 | −9.50 | −7.30 | 1.69 |
| | [−3.27] | [−4.43] | [−4.57] | [−4.36] | [−3.38] | [1.07] |
| American Put Option | −15.06 | −14.37 | −13.99 | −12.47 | −10.11 | 4.95 |
| | [−4.71] | [−5.10] | [−5.26] | [−4.77] | [−4.02] | [3.25] |
| Spread Portfolio | 6.07 | 4.67 | 4.07 | 2.96 | 2.81 | −3.26 |
| | [4.70] | [4.83] | [5.22] | [4.38] | [5.36] | [−3.19] |
| **Panel B: FFC Model Idiosyncratic Volatility** | | | | | | |
| Replication Portfolio | −9.04 | −9.78 | −9.63 | −9.56 | −7.41 | 1.63 |
| | [−3.29] | [−4.46] | [−4.24] | [−4.45] | [−3.57] | [1.05] |
| American Put Option | −14.94 | −14.56 | −13.63 | −12.67 | −10.21 | 4.73 |
| | [−4.63] | [−5.16] | [−5.03] | [−4.86] | [−4.21] | [3.03] |
| Spread Portfolio | 5.90 | 4.77 | 4.00 | 3.11 | 2.80 | −3.10 |
| | [4.76] | [4.82] | [5.12] | [4.40] | [5.57] | [−3.18] |

## Table 2.7: Fama-MacBeth (1973) Regressions

The table presents the results of Fama-MacBeth (1973) regressions of the month-$t$ return of a spread portfolio long a stock-and-riskfree-asset portfolio replicating a put and short the put (Panel A), of the replication portfolio (Panel B), and of the put (Panel C) on subsets of stock and option characteristics plus a constant. The characteristics are the log strike price, the strike-to-stock price ratio ("moneyness"), time-to-maturity as fraction of a year, and idiosyncratic stock volatility, all measured at the start of month $t$. We calculate idiosyncratic volatility using the Fama-French-Carhart model estimated over the prior 60 months. The replication portfolio observations and put observations are matched, so that each replication portfolio observation corresponds to one put observation. The plain numbers are premium estimates, and the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length.

| | Regression Model: | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | | | Panel A: Spread Portfolio Return | | | |
| Strike Price | | 0.04 | 0.04 | | | 0.04 |
| | | [8.34] | [7.91] | | | [8.03] |
| Moneyness | | | −0.26 | | −0.29 | −0.26 |
| | | | [−6.70] | | [−7.46] | [−6.96] |
| Time-to-Maturity | | | −0.17 | | −0.18 | −0.16 |
| | | | [−13.85] | | [−14.06] | [−13.78] |
| Volatility | | | | −0.03 | −0.02 | 0.04 |
| | | | | [−3.81] | [−2.59] | [5.02] |
| Constant | 0.04 | −0.10 | 0.22 | 0.05 | 0.40 | 0.20 |
| | [7.33] | [−6.26] | [4.63] | [7.25] | [8.07] | [4.04] |
| | | | Panel B: Replication Portfolio Return | | | |
| Strike Price | | 0.02 | 0.02 | | | 0.03 |
| | | [2.27] | [2.07] | | | [3.31] |
| Moneyness | | | −0.32 | | −0.37 | −0.35 |
| | | | [−5.00] | | [−5.79] | [−5.53] |
| Time-to-Maturity | | | 0.49 | | 0.49 | 0.50 |
| | | | [14.21] | | [14.18] | [14.64] |
| Volatility | | | | 0.05 | 0.08 | 0.11 |
| | | | | [2.17] | [3.24] | [4.11] |
| Constant | −0.09 | −0.16 | 0.09 | −0.12 | 0.17 | 0.04 |
| | [−4.38] | [−4.91] | [1.02] | [−4.89] | [1.86] | [0.38] |
| | | | Panel C: Put Return | | | |
| Strike Price | | −0.02 | −0.02 | | | −0.01 |
| | | [−3.67] | [−3.45] | | | [−1.99] |
| Moneyness | | | −0.06 | | −0.08 | −0.08 |
| | | | [−0.77] | | [−0.93] | [−1.00] |
| Time-to-Maturity | | | 0.66 | | 0.67 | 0.67 |
| | | | [18.68] | | [18.92] | [19.00] |
| Volatility | | | | 0.08 | 0.09 | 0.07 |
| | | | | [3.39] | [3.94] | [2.62] |
| Constant | −0.13 | −0.06 | −0.13 | −0.17 | −0.22 | −0.16 |
| | [−5.52] | [−1.81] | [−1.07] | [−6.21] | [−1.92] | [−1.33] |

### Table 2.8: Mean Black-Scholes Vs. Real-World Early Exercise Probabilities

The table presents the mean optimal Black-Scholes (1973; column (1)) and real-world (column (2)) early exercise probabilities of our sample puts separately by moneyness. We calculate both probabilities over the remaining time-to-maturity of the puts. At the end of each sample month $t-1$, we first sort the puts into moneyness portfolios according to the decile breakpoints of their Black-Scholes (1973) probabilities. We then calculate means first by portfolio and sample month and then average over our sample period by portfolio. We also report the mean differences between the two probabilities (column (1)–(2)) and their associated $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length (column (3)).

| | Mean Black-Scholes Early Exercise Probability (in %) | Mean Real-World Early Exercise Probability (in %) | Difference (in %) | $t$-statistic of the Difference |
|---|---|---|---|---|
| | (1) | (2) | (1)–(2) | (3) |
| 1 (Low) | 16.35 | 6.27 | 10.08 | [4.62] |
| 2 | 25.93 | 8.28 | 17.65 | [6.82] |
| 3 | 32.51 | 10.77 | 21.74 | [8.52] |
| 4 | 38.19 | 13.34 | 24.85 | [10.09] |
| 5 | 43.36 | 15.11 | 28.25 | [10.33] |
| 6 | 48.31 | 17.30 | 31.01 | [11.78] |
| 7 | 53.06 | 20.81 | 32.25 | [11.75] |
| 8 | 58.04 | 22.35 | 35.69 | [11.30] |
| 9 | 64.15 | 24.28 | 39.87 | [11.45] |
| 10 (High) | 76.74 | 27.05 | 49.69 | [13.50] |

## Table 2.9: Mean Difference Between Black-Scholes and Real-World Early Exercise Probabilities By Moneyness, Maturity Time, and Volatility

The table presents the mean difference between the optimal Black-Scholes (1973) and the real-world early exercise probability of our sample puts by moneyness, time-to-maturity, and idiosyncratic stock volatility. We calculate both probabilities over the strategy return period. At the end of each month $t-1$, we first sort our sample puts into portfolios according to whether their strike-to-stock price ratio ("moneyness") lies above 1.10 (Panel A), between 1.025 and 1.10 (Panel B), or between 0.975 and 1.025 (Panel C). Within each moneyness portfolio, we next sort into portfolios according to whether their days-to-maturity are below 60, between 60 and 90, or above 90 days. Within each moneyness-maturity sorted portfolio, we finally sort into portfolios according to the quintile breakpoints of the Fama-French-Carhart-model idiosyncratic stock volatility. See the captions of Tables 2.5 and 2.6 for details on the sorting variables. We then calculate means first by portfolio and sample month and then average over our sample period by portfolio. Plain numbers are the mean early exercise probability differences (in %), while the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length.

| | Idiosyncratic FFC Stock Volatility | | | | |
|---|---|---|---|---|---|
| Days-to-Maturity | 1 (Low) | 2 | 3 | 4 | 5 (High) |
| Panel A: Deep In-The-Money (Strike-to-Stock Price > 1.10) | | | | | |
| 30-60 | 54.83 | 44.17 | 36.66 | 31.47 | 25.28 |
| | [23.64] | [19.23] | [17.07] | [16.12] | [16.96] |
| 60-90 | 48.86 | 32.17 | 21.98 | 14.29 | 7.54 |
| | [13.25] | [10.67] | [10.15] | [8.75] | [8.72] |
| 90-120 | 38.02 | 19.93 | 10.61 | 5.47 | 1.98 |
| | [8.09] | [5.75] | [4.92] | [4.56] | [4.87] |
| Panel B: In-The-Money (Strike-to-Stock Price 1.025 to 1.10) | | | | | |
| 30-60 | 35.00 | 28.08 | 24.20 | 21.12 | 18.48 |
| | [20.18] | [29.06] | [35.83] | [30.90] | [37.09] |
| 60-90 | 22.96 | 13.93 | 9.49 | 6.75 | 4.22 |
| | [7.54] | [8.69] | [9.34] | [11.74] | [16.37] |
| 90-120 | 14.39 | 6.32 | 3.58 | 1.76 | 0.89 |
| | [4.27] | [4.07] | [3.84] | [3.84] | [6.73] |
| Panel C: At-The-Money (Strike-to-Stock Price 0.975 to 1.025) | | | | | |
| 30-60 | 13.50 | 12.58 | 12.07 | 12.05 | 12.07 |
| | [35.12] | [47.26] | [43.46] | [58.22] | [50.32] |
| 60-90 | 4.89 | 3.86 | 3.41 | 3.05 | 2.29 |
| | [7.36] | [10.54] | [10.93] | [14.06] | [14.05] |
| 90-120 | 1.69 | 0.86 | 0.56 | 0.37 | 0.24 |
| | [3.04] | [2.59] | [2.61] | [2.49] | [6.69] |

**Table 2.10: Fama-MacBeth Regressions Explaining the Difference Between Black-Scholes and Real-World Early Exercise Probabilities**

The table presents the results of Fama-MacBeth (1973) regressions of the difference between Black-Scholes (1973) and real-world early exercise probabilities for our sample puts on subsets of stock and option characteristics plus a constant. We calculate the two probabilities over the strategy return period. The characteristics include the strike-to-stock price ratio ("moneyness"), time-to-maturity as fraction of a year, and idiosyncratic stock volatility, and they are measured until the start of the strategy return period. We calculate idiosyncratic stock volatility from the Fama-French-Carhart model estimated over the prior 60 months. The plain numbers are the average estimates, while the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a twelve-month lag length.

| | Regression Model: | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Moneyness | 1.37 | | | 1.47 | 1.61 |
| | [30.97] | | | [33.68] | [35.75] |
| Time-to-Maturity | | −1.05 | | −1.15 | −1.18 |
| | | [−35.32] | | [−39.52] | [−41.43] |
| Volatility | | | −0.24 | | −0.32 |
| | | | [−16.57] | | [−20.18] |
| Constant | −1.29 | 0.39 | 0.24 | −1.16 | −1.20 |
| | [−29.41] | [50.65] | [32.59] | [−27.39] | [−29.27] |

## Table 2.11: Adjusting the Strategy for Transaction Costs

The table presents the mean transaction-cost-adjusted returns of spread portfolios long stock-and-riskfree-asset portfolios replicating a put and short the puts sorted on the puts' moneyness and time-to-maturity. To adjust for transaction costs, we assume that investors always buy (sell) at the midpoint price plus (minus) $\varphi$ times the bid-ask spread. We furthermore assume that investors can borrow stocks at Markit's indicative rate. Panel A considers the full sample, while Panel B restricts attention to only those strategies involving puts with a bid-ask spread below the median and stocks with an Amihud (2002) illiquidity value below the median, both measured at the start of the spread return period. At the end of each sample month $t-1$, we first sort the spread portfolios into portfolios according to whether the strike-to-stock price ratio ("moneyness") of the associated put lies above 1.10 (Panels A.1 and B.1), between 1.025 and 1.10 (Panels A.2 and B.2), or between 0.975 and 1.025 (Panels A.3 and B.3). Within each moneyness portfolio, we next sort them into portfolios according to whether their days-to-maturity are below 60, between 60 and 90, or above 90 days. We equally-weight the portfolios and hold them over month $t$. Plain numbers are mean monthly portfolio returns (in %) and the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a lag length of twelve months.

| | Borrowing Cost Adjustment/Bid-Ask Spread Fraction $\varphi$: | | | |
|---|---|---|---|---|
| Days-to-Maturity | No/0.00 | Yes/0.10 | Yes/0.25 | Yes/0.50 |
| | (1) | (2) | (3) | (4) |
| Panel A: Full Sample | | | | |
| *Panel A1: Deep In-The-Money (Strike-to-Stock Price > 1.10)* | | | | |
| 30-60 | 2.48 | −0.08 | −3.30 | −10.05 |
| | [4.92] | [−0.16] | [−4.42] | [−6.26] |
| 60-90 | 1.76 | −0.70 | −3.94 | −10.45 |
| | [3.78] | [−1.37] | [−5.11] | [−6.62] |
| 90-120 | 1.41 | −1.02 | −4.38 | −11.39 |
| | [3.48] | [−2.15] | [−5.81] | [−6.67] |
| *Panel A2: In-The-Money (Strike-to-Stock Price 1.025 to 1.10)* | | | | |
| 30-60 | 4.83 | 1.68 | −2.17 | −10.37 |
| | [5.87] | [1.96] | [−2.05] | [−5.71] |
| 60-90 | 3.15 | 0.19 | −3.64 | −11.67 |
| | [4.18] | [0.24] | [−3.63] | [−6.83] |
| 90-120 | 2.47 | −0.37 | −4.21 | −12.64 |
| | [3.72] | [−0.52] | [−4.52] | [−6.84] |
| *Panel A3: At-The-Money (Strike-to-Stock Price 0.975 to 1.025)* | | | | |
| 30-60 | 9.49 | 5.56 | 1.04 | −8.93 |
| | [6.25] | [3.62] | [0.60] | [−3.79] |
| 60-90 | 4.72 | 1.36 | −2.93 | −12.29 |
| | [4.20] | [1.14] | [−2.10] | [−6.15] |
| 90-120 | 3.32 | 0.10 | −4.24 | −14.16 |
| | [3.31] | [0.10] | [−3.40] | [−6.72] |
| Panel B: High-Liquidity Put and Stock Sample | | | | |
| *Panel B1: Deep In-The-Money (Strike-to-Stock Price > 1.10)* | | | | |
| 30-60 | 3.93 | 2.65 | 1.71 | 0.14 |
| | [3.75] | [2.93] | [1.83] | [0.14] |
| 60-90 | 3.27 | 1.93 | 1.02 | −0.52 |
| | [3.43] | [2.57] | [1.32] | [−0.65] |
| 90-120 | 2.57 | 1.61 | 0.69 | −0.85 |
| | [4.41] | [2.86] | [1.18] | [−1.34] |

*(continued on next page)*

**Table 2.11: Adjusting the Strategy for Transaction Costs (Cont.)**

| | Borrowing Cost Adjustment/Bid-Ask Spread Fraction $\varphi$: | | | |
|---|---|---|---|---|
| Days-to-Maturity | No/0.00 | Yes/0.10 | Yes/0.25 | Yes/0.50 |
| | (1) | (2) | (3) | (4) |
| *Panel B2: In-The-Money (Strike-to-Stock Price 1.025 to 1.10)* | | | | |
| 30-60 | 6.87 | 5.09 | 3.92 | 1.95 |
| | [4.93] | [4.24] | [3.20] | [1.51] |
| 60-90 | 5.85 | 4.22 | 3.10 | 1.22 |
| | [3.96] | [3.35] | [2.44] | [0.93] |
| 90-120 | 3.79 | 2.60 | 1.52 | −0.29 |
| | [4.01] | [2.76] | [1.56] | [−0.28] |
| *Panel B3: At-The-Money (Strike-to-Stock Price 0.975 to 1.025)* | | | | |
| 30-60 | 12.15 | 9.37 | 7.98 | 5.66 |
| | [4.93] | [4.39] | [3.66] | [2.48] |
| 60-90 | 8.32 | 5.79 | 4.52 | 2.40 |
| | [3.95] | [3.52] | [2.72] | [1.38] |
| 90-120 | 5.70 | 4.35 | 3.14 | 1.10 |
| | [3.73] | [2.86] | [2.00] | [0.66] |

**Table 2.12: Adjusting the Strategy for Transaction Costs Under Different Replication Portfolio Rebalancing Schemes**

The table presents the mean transaction-cost-adjusted returns of spread portfolios long stock-and-riskfree-asset portfolios replicating a put and short the puts using different rebalancing schemes. To adjust for transaction costs, we assume that investors always buy (sell) at the midpoint price plus (minus) $\varphi$ times the bid-ask spread. We further assume investors can borrow stocks at Markit's indicative rate. In Panels A, B, and C, we consider the daily, weekly, and no rebalancing cases, respectively. The table restricts attention to only those strategies involving puts with a bid-ask spread below the median and stocks with an Amihud (2002) illiquidity value below the median, both measured at the start of the spread return period. Within each panel, we further separately look into strategies involving only puts with a price above \$1, \$2, and \$5 at the start of the strategy return period. We equally-weight the portfolios and hold them over month $t$. Plain numbers are mean monthly portfolio returns (in %) and the numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) formula with a lag length of twelve months.

| | Borrowing Cost Adjustment/Bid-Ask Spread Fraction $\varphi$: | | | |
|---|---|---|---|---|
| Days-to-Maturity | No/0.00 | Yes/0.10 | Yes/0.25 | Yes/0.50 |
| | (1) | (2) | (3) | (4) |
| **Panel A: Daily Rebalancing** | | | | |
| Put Price ≥\$1 | 5.00 | 3.51 | 2.40 | 0.53 |
| | [4.90] | [3.86] | [2.55] | [0.52] |
| Put Price ≥\$2 | 5.12 | 3.64 | 2.61 | 0.87 |
| | [4.65] | [3.79] | [2.64] | [0.83] |
| Put Price ≥\$5 | 6.54 | 5.06 | 4.25 | 2.89 |
| | [4.14] | [3.72] | [3.08] | [2.04] |
| **Panel B: Weekly Rebalancing** | | | | |
| Put Price ≥\$1 | 5.29 | 4.14 | 3.11 | 1.39 |
| | [6.08] | [5.48] | [3.98] | [1.67] |
| Put Price ≥\$2 | 5.32 | 4.21 | 3.25 | 1.64 |
| | [5.67] | [5.19] | [3.89] | [1.86] |
| Put Price ≥\$5 | 6.12 | 4.97 | 4.21 | 2.93 |
| | [4.73] | [4.49] | [3.73] | [2.51] |
| **Panel C: No Rebalancing** | | | | |
| Put Price ≥\$1 | 6.80 | 5.61 | 4.72 | 3.21 |
| | [8.27] | [9.19] | [7.47] | [4.78] |
| Put Price ≥\$2 | 6.56 | 5.33 | 4.50 | 3.10 |
| | [7.37] | [8.39] | [6.88] | [4.49] |
| Put Price ≥\$5 | 6.74 | 5.35 | 4.69 | 3.59 |
| | [5.48] | [6.10] | [5.25] | [3.88] |

# Chapter 3

# Why Does the Implied Volatility Spread Predict Future Stock Returns?

***Keywords***: Empirical asset pricing; cross-sectional option pricing; implied volatility spread; put options; early exercise; frictions.

## 3.1 Introduction

Recent studies have shown that the difference in call and equivalent put implied volatilities, known as the implied volatility spread, can predict cross-sectional stock returns (see Bali and Hovakimian (2009), Cremers and Weinbaum (2010) and An, Ang, Bali and Cakici (2014) among others). Focusing on exchange-traded single-stock American options, these studies obtain option implied volatility values from the Optionmetrics Ivy DB database. Optionmetrics calculates the implied volatility for an American option under the Black-Scholes (1973) framework using the Cox-Ross-Rubinstein (CRR) binomial pricing model. In a Black-Scholes world where trading frictions play no role, any probable early exercises are conditioned upon characteristics of the underlying stock and option, characteristics such as whether the underlying stock pays a dividend, moneyness, time to maturity and so forth. Thus, in a Black-Scholes world early exercises do not drive any implied volatility spread between an equivalent American call/put option pair. A number of recent studies, however, have documented that the presence of frictions alters the optimal early exercise behaviour of holders of options (see Jensen and Pedersen (2016), Cao, Ederington and Yadav (2017) and Figlewski (2020), for example.) Allowing for trading frictions, the Black-Scholes set-up would underestimate the attractiveness of exercising the option early and thus its value. More specifically, given Optionmetrics uses a Black-Scholes framework to price American options, the pricing method used by Optionmetrics will not capture the value added of being able to exercise the option early in order to minimize trading-friction costs. This skews implied volatility upwards so that the model price matches the traded value of an option, in turn leading to an implied volatility spread between equivalent American call and put option pairs.

In this paper we investigate what drives the ability of implied volatility spreads to predict stock returns that Cremers and Weinbaum (2010) document. In particular, we examine the role of the early exercise premium, defined as the difference between the market price of an American put option and its equivalent synthetic European price, in predicting cross-sectional stock returns. In a Black-Scholes world, the early exercise premium should not carry any information about the corresponding implied volatility spread. However, if, as discussed above, the presence of trading frictions affects optimal early exercise behaviour it is possible that the early exercise premium does contain information about the corresponding implied volatility spread. Decomposing the implied volatility spread into the early exercise premium and residual frictions, which capture the impact of friction-driven optimal call early exercises on implied volatility spreads along with any other factors not related to early exercise, we find that the implied volatility spread and the early exercise premium are highly correlated with a mean cross-sectional correlation of $-0.63$. This is suggestive of a strong relationship between the implied volatility spread and the (friction-driven) early exercise premium, and that the early exercise premium has a significant role to play in explaining movements in the implied volatility spread. Our empirical evidence appears to confirm this. We find that when implied volatility spreads are decomposed into the early exercise premium and residual frictions, only the early exercise premium significantly predicts stock returns; the contribution of residual frictions are mostly insignificant. We also show that the predictive ability of the early exercise premium can not be solely explained by the informed trading activities of option investors suggested by Cremers and Weinbaum (2010).

In our empirical analysis we use exchange-traded single-stock options written on underlying assets not paying any cash within these options' maturity periods ("zero-dividend stocks"). We start by calculating the implied volatility spread and the early exercise premium for each individual option in our sample. To calculate the implied volatility spread for each option, we simply take the difference in call implied volatility and its equivalent put volatility; both of these implied volatilities are taken from the Optionmetrics Ivy DB database. To calculate the early exercise premium, we require equivalent European put option prices. These are not readily available as single-stock options are only American by design. To obtain European put prices we use Merton's (1973) insight that it is never optimal to early exercise an American call written on a zero-dividend asset, allowing us to treat these calls as *quasi*-European call options. Prompted by Zivney (1991), we next recognize that from European put-call parity a European put option can be synthetically replicated using a portfolio long the equivalent European call option, long an investment of the discounted strike price into a money market account, and short the underlying asset. This gives us a *synthetic* European put price. We then calculate the early exercise premium by taking the difference between the exchange-traded American put option price and its equivalent synthetic European price.

Armed with the early exercise premium and the implied volatility spread for each option, we proceed to calculate the early exercise premium and the implied volatility spread at the stock level. To obtain the early exercise premium for a particular stock, we weight the early exercise premium for each option written on that stock by each option's outstanding open interest. The early exercise premium at the stock level is then calculated as the open-interest-weighted average of the individual option early exercise premia. To obtain implied volatility spreads at the stock level, we follow a similar procedure, calculating open-interest-weighted average call and put implied volatilities separately and then taking the difference to arrive at a measure of the implied volatility spread for the stock. To obtain our measure of residual frictions, we run a cross-sectional regression each month of the stock-level implied volatility spread on the stock-level early exercise premium. The error term from this regression then gives us our measure of residual frictions at the stock level. The cross-sectional-average adjusted $R^2$ is 38.47%, indicating that a sizable proportion of the variation in the implied volatility spread is explained by the variation in the early exercise premium.

Forming portfolios sorted on the implied volatility spread, the early exercise premium and frictions respectively, we find that, consistent with Cremers and Weinbaum (2010), there is a positive relationship between implied volatility spreads and future stock returns. Sorting on early exercise premia, we find that there is a negative relationship between early exercise premia and future stock returns, which in turn implies that there is a negative association between implied volatility spreads and early exercise premia. Importantly, the early exercise premium significantly predicts stock returns: the "High–Low" spread portfolio sorted on early exercise premia, for example, generates an excess mean return of −0.71%, *per month* ($t$-statistic: −4.34) with a mean monthly Fama-French-Carhart 4-factor-adjusted alpha of −0.91% ($t$-statistic: −5.34). Interestingly, however, the "High–Low" portfolio sorted on residual frictions generates an insignificant excess mean return of 0.13% per month ($t$-statistic: 0.62). This suggests that the predictive ability of the implied volatility spread that Cremers and Weinbaum (2010) identify comes from the (friction-induced) early exercise premium rather than frictions more generally. To further investigate this finding, we conduct a double portfolio sort on the early exercise premium and frictions to check whether the early exercise premium truly carries incremental predictive information over residual frictions. We find that all of the "High–Low" portfolios sorted on early exercise premia while controlling for frictions generate significant mean excess returns. In contrast, for portfolios sorted on frictions while controlling for early exercise premia, all but one of the "High-Low" portfolios generate insignificant mean returns. This additional evidence suggests that the ability of implied volatility spreads to predict stock returns comes largely from early exercise premia rather than residual frictions. Fama and MacBeth (1973) cross-sectional regressions, where we control for a number of stock and option characteristics, confirm our portfolio sort results that the early exercise premium is a significant predictor of stock returns.

Our empirical evidence relies crucially on it never being optimal to early exercise an American call option on a zero-dividend asset and put-call parity holding. However, as indicated in Jensen and Pedersen (2016) and Figlewski (2020), both these conditions can be violated due to friction-driven early exercises for single-stock call options. To establish whether such violations confound our empirical results, we condition our tests on the option bid-ask spread, which we use as a measure of option illiquidity, and the Daily-Cost-to-Borrowing Score (DCBS) from Markit which is a measure of short-sale constraints on the underlying stock. Our results suggest that neither of these exert any meaningful effect on our findings. We further show that our results remain even if we change the definition of stock-level early exercise premia or control for the nonsynchronicity between stock and option markets.

Our work adds to the empirical strand of the literature studying the implications of option or option-implied variables for the behaviour of future stock returns. Manaster and Rendleman (1982) show that model stock prices implied by the Black-Scholes (1973) universe can predict future market prices and subsequent returns on a daily basis. Easley, O'Hara and Srinivas (1998), and Pan and Poteshman (2006) show that stock returns are predictable using raw option trading volume. Cremers and Weinbaum (2010) use implied volatility spreads to predict stock returns while An et al. (2014) also document that call and put implied volatilities separately have predictive ability. The predictor variable in Bali and Hovakimian (2009) is calculated as the difference between underlying stock's realized volatility and its option implied volatility. Finally, Conrad, Dittmar and Ghysels (2013) show that option-implied risk neutral moments, calculated using the frameworks in Bakshi and Madan (2000) and Bakshi, Kapadia and Madan (2003), also predict future stock returns. We contribute to this literature by identifying another option-based factor that predicts the cross section of stock returns: the early exercise premium which we measure as the difference between the American put option market price and its equivalent synthetic European price.

We also contribute to the literature looking into the drivers of option-to-stock market predictability. Using single-stock options, Cremers and Weinbaum (2010) show that the ability of the implied volatility spread to predict stock returns is due to investors' informed trading activities in the options market, ahead of the stock market, although Shang (2017) documents that informed trading is not the sole driver of the return predictability documented in Cremers and Weinbaum (2010). Gonclaves-Pinto, Grundy, Hameed, van der Heiden and Zhu (2020), and Hiraki and Skiadopoulos (2020) show that trading frictions drive implied volatility spreads and therefore it is ultimately trading frictions that predicts stock returns. Our findings complement this literature by showing that if we decompose the implied volatility spread into the (friction-induced) early exercise premium and residual frictions, it is the friction-induced early exercise premium that predicts stock returns.

The rest of the chapter is organised as follows. Section 3.2 discusses the data and methodology. In section 3.3, we document the ability of the early exercise premium to predict returns while controlling for frictions and other cross-sectional stock and option characteristics. Section 3.4 presents robustness tests of the results in section 3.3 while in section 3.5, we examine the extent to which investors' informed trading activities explain the results in section 3.3. Section 3.6

summarizes and concludes.

## 3.2 Data and Methodology

In this section, we describe our data sources, filters, and the calculation of our early exercise premia and implied volatility spreads, both at the option and stock levels.

### 3.2.1 Data Sources and Filters

We obtain daily data from the beginning of January 1996 until the end of April 2016 on American call and put options written on zero-dividend stocks, on the stocks underlying the options, and on the term structure of the risk-free rate of return from the Optionmetrics IvyDB database. We also extract implied volatilities for our option sample from this database. Market data on the underlying stocks and the return on the market comes from CRSP, firm fundamental data comes from Compustat and data on the size, book-to-market, momentum, profitability and investment factors come from Kenneth French's website.[1] Finally, data on stock short-sale constraints comes from Markit.

The filters we impose on our option data are similar to those imposed by Cremers and Weinbaum (2010). Specifically, we exclude an option pair from an observation month if either the call or put option of this pair has zero open interest, missing implied volatility, zero bid price or a price that violates the standard arbitrage bounds (e.g. the bound that an American call option's price must lie between the maximum of zero and the value of the equivalent long forward, and the stock price) at the end of each calendar month. We also exclude observations where the option price is below $1 or the stock price is below $5 to avoid any market micro-structure issues.

[1] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. We would like to thank Kenneth French for making this data available.

### 3.2.2 Calculating The Early Exercise Premium and Implied Volatility Spread

To calculate the early exercise premium, we need the American put option price and its equivalent European price. However, an issue we face is that exchange-traded single-stock options are exclusively American options so there is no market European price. To address this issue, we synthetically create European put options where we trade in American options, the underlying stocks, and the risk-free asset. As our sample only consists of options on zero-dividend stocks, then based on Merton's (1973) insight that it is never optimal to exercise an American call option written on a zero-dividend-paying stock early, the American calls in our sample are effectively European calls. Given this, then from European put-call parity a European put option can be synthetically replicated using a portfolio long the equivalent European call option, which given Merton's (1973) result and given we only use options on non-dividend-paying stocks we take as the equivalent American call option, long an investment of the discounted strike price in the risk-free asset, and short the underlying stock. We can therefore make use of put-call parity to write,

$$P_{i,j}^{synE} = C_{i,j}^{A} - S_j + Ke^{-r_f T}, \tag{3.1}$$

where $P_{i,j}^{synE}$ is the price of a synthetic European put option $i$ written on stock $j$ with strike price $K$ and time-to-maturity $T$, $C_{i,j}^{A}$ is the market price of the exchange-traded American call option written on the same stock with the same strike price and time-to-maturity, $S_j$ is stock $j$'s price, and $r_f$ is the risk-free return over the option's time to maturity.

We calculate the early exercise premium at the option level as the difference between the exchange-traded American put price and its equivalent *synthetic* European price:

$$EEP_{i,j} = P_{i,j}^{A} - P_{i,j}^{synE}, \tag{3.2}$$

where $EEP_{i,j}$ and $P_{i,j}^{A}$ are respectively the early exercise premium and the market price of an American put option $i$ written on stock $j$ with strike price $K$ and time-to-maturity $T$.

The implied volatility spread for each American option pair is the difference between the implied volatilities of equivalent call and put options written on the same underlying stock with the same strike price and time-to-maturity. We write the option-level implied volatility spread as,

$$IVSprd_{i,j} = IV_{i,j}^{CA} - IV_{i,j}^{PA}, \tag{3.3}$$

where $IVSprd_{i,j}$ is the implied volatility spread for an American option pair $i$ written on stock $j$ with strike price $K$ and time-to-maturity $T$, and $IV_{i,j}^{CA}$ and $IV_{i,j}^{PA}$ are respectively the implied volatilities of exchange-traded American call and put options written on the same stock with the same strike price and time-to-maturity.

Armed with option-level implied volatility spreads, we can calculate the implied volatility spread for a stock. To calculate stock-level spreads, we follow Cremers and Weinbaum (2010) and take a weighted-average of the option-level call and put implied volatilities for each stock where the weights are calculated as the relative open interest of these options:

$$IVSprd_j = IV_j^C - IV_j^P = \sum_{i=1}^{N} \left( w_{i,j}^C \times IV_{i,j}^C \right) - \sum_{i=1}^{N} \left( w_{i,j}^P \times IV_{i,j}^P \right), \tag{3.4}$$

where $IVSprd_j$, $IV_j^C$ and $IV_j^P$ are the implied volatility spread, the call implied volatility and the put implied volatility for stock $j$, respectively, and $w_{i,j}^C$ and $w_{i,j}^P$ are the call and put option weights, respectively. The weights are calculated as $w_{i,j}^C = \frac{OI_{i,j}^C}{\sum_{i=1}^{N} OI_{i,j}^C}$ and $w_{i,j}^P = \frac{OI_{i,j}^P}{\sum_{i=1}^{N} OI_{i,j}^P}$ where $OI$ is the option's open interest and $N$ is the number of option pairs written on a particular stock.

We follow a similar procedure to calculate the stock-level early exercise premium although as only the put option is involved in calculating the early exercise premium, we weight option-level early exercise premia using the relative open interest of the American put only:

$$EEP_j = \sum_{i=1}^{N} \left( w_{i,j}^{P} \times EEP_{i,j} \right) = \sum_{i=1}^{N} \left[ w_{i,j}^{P} \times \left( P_{i,j}^{A} - P_{i,j}^{synE} \right) \right] \tag{3.5}$$

where $EEP_j$ is the weighted-average early exercise premium for stock $j$.

## 3.3 Empirical Results

In this section we present our main results on whether early exercise premia predict future stock returns. We begin by presenting descriptive statistics for option-level early exercise premia and other option-specific variables and a number of pre-formation characteristics for the stock portfolios we form by sorting on stock-level early exercise premia. We then provide the results from various portfolio sorts and Fama and MacBeth (1973) regressions studying the relationship between early exercise premia and future stock returns while controlling for other factors.

### 3.3.1 Option Data Characteristics

Table 3.1 reports descriptive statistics for option-level early exercise premia and implied volatility spreads (columns (1) and (2)); option moneyness, defined as the ratio of the strike price to the underlying stock price, and days-to-maturity (columns (3) and (4)); and outstanding open interest for call and put options (columns (5) and (6)). Observations are taken at the end of each calendar month. The option pairs in columns (1) and (2) are matched along moneyness and time-to-maturity dimensions so that each option pair in column (1) is associated with exactly one pair in column (2) with identical moneyness and time-to-maturity. With the exception of the $t$-statistic testing the null hypothesis that the mean of the variable is zero, the descriptive statistics are calculated each sample month and then averaged over time.

Table 3.1 About Here

Table 3.1 shows that the month-end option-level early exercise premia are generally positive in our sample, with a mean of 7.82% ($t$-statistic: 6.11), indicating that the price of the American put option is significantly higher than the price of an equivalent European put. As our implied volatility spread is simply the difference between equivalent call and put implied volatilities, any positive early exercise premium induced by frictions should skew the put implied volatility upward, leading to a negative implied volatility spread for each American option pair. We should therefore expect a negative association between these two spreads. In our sample, the mean implied volatility spread at the option level is $-1.06\%$ ($t$-statistic: $-9.12$), implying a significantly higher implied volatility for the American put option on average compared to the implied volatility of the equivalent American call. The reason for the (in absolute terms) large $t$-statistic for the implied volatility spread relative to the early exercise premium is that the implied volatility spreads are far less volatile than the early exercise premia, as can be seen from their standard deviations and percentiles.

Looking at the percentiles we can see that, consistent with the findings in Cremers and Weinbaum (2010), the implied volatility spread in our sample ranges from negative to positive. The percentiles also show that the early exercise premium ranges from positive to negative in our sample. This is interesting because in frictionless markets we should only observe positive early exercise premia. However, the presence of negative values for the early exercise premium raises the question of what role, if any, market frictions play and whether the empirical results are driven by the presence of negative early exercise premia in our sample. We will return to this point later.

The moneyness and days-to-maturity statistics in columns 3 and 4 suggest that the average option pair is close to at-the-money and has slightly less than two months to maturity. We can also see that with respect to open interest for both call and put options, there are a high number of open interests in the 99th percentile and a low number in the 1st percentile. Overall, an exchange-traded American call option has higher open interest compared to its equivalent put counterpart.

### 3.3.2 Pre-formation Stock Portfolio Characteristics

To begin our empirical analysis, we form stock portfolios by sorting on early exercise premia calculated at the stock level. At the end of each sample month $t-1$, we sort the universe of stocks in our sample into quintile portfolios according to stock-level early exercise premia. Table 3.2 reports mean values for different pre-formation characteristics of these portfolios evaluated at or over month $t-1$. The bottom quintile (1, or "Low") contains stocks with low early exercise premia while the top quintile (5, or "High") contains stocks with high early exercise premia. With the exception of the lagged stock return, the figures in Table 3.2 are time series averages of the equally-weighted cross-sectional means calculated each month. For lagged stock returns, we calculate the value-weighted mean.

TABLE 3.2 ABOUT HERE

Consistent with Cremers and Weinbaum (2010), we observe that stocks with lower average market capitalization are located in the "Low" and "High" portfolios. Average betas show little change across the portfolios while stock volatility remains essentially unchanged across the portfolios. Average monthly lagged stock returns, calculated over month $t-1$, increase near-monotonically across the portfolios from 0.65% per month for the "Low" portfolio to 2.21% per month for the "High" portfolio.

The bottom two rows of Table 3.2 report average early exercise premia and implied volatility spreads calculated at the stock level. While the increase in early exercise premia from the "Low" to "High" portfolios is mechanical given we sort the portfolios on the early exercise premium, interestingly we observe a decrease in average implied volatility spreads across these portfolios. More importantly, the negative early exercise premia on average are only associated with positive implied volatility spreads, a pattern which, as discussed earlier, we would expect to observe.

### 3.3.3 Does The Early Exercise Premium Predict Future Stock Returns?

We now turn our attention to investigating the drivers of the ability of the implied volatility spread to predict stock returns that Cremers and Weinbaum (2010) identify. Recall from the discussion earlier that in a Black-Scholes world where markets are frictionless, the early exercise premium should carry no information about the corresponding implied volatility spread since any early exercise is conditioned upon the characteristics of the underlying stock and option, such as whether the stock pays dividends. However, several recent papers (Jensen and Pedersen (2016), Cao et al. (2017) and Figlewski (2020), for example) have documented that the presence of trading frictions can alter the optimal early exercise behavior of option investors. In particular, the presence of frictions means it may be optimal for investors to exercise options early when it otherwise would not be. Consequently, in the presence of trading frictions Black-Scholes-model-based implied volatilities will contain a component related to friction-induced early exercises. We can therefore decompose such a model-based implied volatility spread into that part of the spread that comes from (friction-induced) early exercise, captured by the early exercise premium, and that part due to residual (other) frictions.

To this end, we start by decomposing stock-level market implied volatility spreads into early exercise premia and residual frictions. While early exercise premia capture the effect of friction-induced put early exercises, residual frictions include the impact of call early exercises along with additional factors not related to early exercise. For each sample month $t-1$, we run the following cross sectional regression of stock-level implied volatility spreads on stock-level early exercise premia:

$$IVSprd_j = \alpha_j + \beta_j^{EEP} EEP_j + \epsilon_j, \tag{3.6}$$

where $\alpha_j$ and $\beta_j^{EEP_j}$ are the regression parameters and $\epsilon_j$ is the error term, which serves as our measure of residual frictions.

Table 3.3 About Here

Results from these cross sectional regressions are reported in Table 3.3. The results show that there is a strong negative relationship between the early exercise premium and the implied volatility spread in our data with $\beta_j^{EEP_j}$ being a highly significant $-0.19$ ($t$-statistic: $-46.99$). The adjusted $R^2$ of 38.47% indicates that a sizable proportion of the cross sectional variation in implied volatility spreads is explained by the early exercise premium.

Given the results in Table 3.3 provide evidence that a reasonable proportion of the implied volatility spread can be explained by the early exercise premium, we examine the extent to which the implied volatility spread, the early exercise premium and residual frictions are separately related to future excess stock returns. To do this, we conduct independent portfolio sorts using each of the these (at the stock level) separately as follows. At the end of sample month $t-1$, we split the stock universe into quintile portfolios based on each factor. The bottom ("Low") quintile contains stocks with low values of the factor on which the portfolios are formed while the top ("High") quintile contains stocks with high values of the factor. We also form a "High−Low" spread portfolio that is long the top quintile and short the bottom quintile. We then hold the portfolios over month $t$. Table 3.4 reports value-weighted returns in excess of the three month Treasury Bill rate from these univariate portfolio sorts.

Table 3.4 About Here

Taking the results for the implied volatility spread first we find that, consistent with Cremers and Weinbaum (2010), future stock returns increase as the implied volatility spread increases. Indeed, a strategy that is long the "High" portfolio and short the "Low" portfolio delivers a very significant average monthly return in excess of the three month Treasury Bill rate of 1.08% ($t$-statistic: 4.41; see the column labelled "High−Low" in Table 3.4). From the table, we also observe that future stock returns decrease as the early exercise premium increases. The "High−Low" portfolio from the early exercise premium sort generates average monthly excess returns of −0.71% ($t$-statistic: −4.34). That we observe an opposing pattern between stock returns and implied volatility spreads, and stock returns and early exercise premia is perhaps not surprising given our earlier finding that there is a significant negative relationship between the implied volatility spread and the early exercise premium. In contrast to the clear relationship between stock returns and the implied volatility spread, and stock returns and the early exercise premium, stock portfolios sorted on residual frictions show little change across the portfolios, with the "High−Low" portfolio providing an insignificant average monthly excess return of 0.13% ($t$-statistic: 0.62).

The results in Table 3.4 strongly suggest that much of the relationship between future stock returns and the implied volatility spread that Cremers and Weinbaum (2010) identify comes from the early exercise premium rather than frictions. Table 3.5 examines whether the findings in Table 3.4 remain once we have controlled for commonly used risk factors. To this end, Table 3.5 reports the monthly average intercept terms (the portfolio $\alpha$s) from regressing the portfolio returns from the various sorts in Table 3.4 on the market (the market model), the four factors in Carhart (1997), which we label FFC (Fama-French-Carhart), and the five fators in Fama and French (2015), which we label FF5 [2] Panel A reports results for the implied volatility spread sorts, Panel B reports results for the early exercise premium sorts and Panel C reports results for the (residual) frictions sort.

<u>TABLE 3.5 ABOUT HERE</u>

Looking at the "High−Low" column in Table 3.5, we find a positive and

[2]The four factors in the FFC model are the market, size (SMB) and book-to-market (HML) factors from Fama and French (1993) and the momentum factor from Carhart (1997) while the FF5 model includes the three factors from Fama and French (1993) along with a profitability and an investment factor.

statistically significant portfolio $\alpha$ for the implied volatility spread sort and a negative and statistically significant $\alpha$ for the early exercise premium sort, regardless of the model used. For instance, the mean monthly FFC-model-adjusted $\alpha$ is a significant 1.27% ($t$-statistic: 6.59) for the implied volatility sort, and a significant $-0.91\%$ ($t$-statistic: $-5.34$) for the early exercise premium sort; the $\alpha$s are also of similar orders of magnitude to the average excess returns reported for these portfolios in Table 3.4. For the residual frictions sort, the $\alpha$s for the "High−Low" portfolios are all statistically insignificant. These results suggest that our findings in Table 3.4 are not due to a failure to adequately control for the usual asset pricing risk factors. To summarize thus far, Tables 3.4 and 3.5 provide evidence confirming the predictive ability of the implied volatility spread that Cremers and Weinbaum (2010) document but also suggesting that such predictability is driven by the early exercise premium rather than frictions.

While the results above provide encouraging evidence about the role of the early exercise premium in predicting stock returns, an important question still remains: do we still have that predictability after controlling for frictions and other characteristics in addition to the usual asset pricing factors? As a first step in answering this question, we undertake two bivariate portfolio sorts. At the end of each sample month $t-1$, we first split the stock universe into quintile portfolios according to the first sorting variable. We then split each portfolio from the first sort into further quintile portfolios based on the second sorting variable. The bottom quintile for each sort contains stocks with low factor values ("Low"), while the top contains stocks with high values ("High"). We also form a spread portfolio long the top quintile and short the bottom quintile ("High−Low") along the second sort dimension. We then hold the portfolios over month $t$ and calculate value-weighted returns on these portfolios.

<u>TABLE 3.6 ABOUT HERE</u>

Table 3.6 presents results from our portfolio double sorts. Panel A reports

results with residual frictions as the first sorting variable and the early exercise premium as the second. Panel B reports results from reversing the sorting order. Looking at the results in panel A, we observe significant changes in mean monthly excess stock returns across different early exercise premium portfolios even after controlling for stock frictions. Average excess returns on the "High−Low" spread portfolios in this case range from −1.47% ($t$-statistic: −4.06) to −0.51% ($t$-statistic: −2.15), suggesting that early exercise premia contain predictive information about stock returns over and above any predictive information frictions contain. In panel B, we check whether frictions contain any incremental predictive information over that contained in early exercise premia. We find average excess returns for the "High−Low" portfolios are only significant for the "High" column, that is, mean monthly excess returns vary significantly with frictions only when frictions are high, delivering a mean excess return of 1.28% ($t$-statistic: 3.77). In all other cases when frictions is the second sorting variable, average excess returns on the "High−Low" portfolios are statistically indistinguishable from zero.

In Table 3.7 we report the results of Fama-MacBeth (1973) regressions to test whether a number of commonly used cross-sectional stock and option-specific factors can explain the apparent return predictability we document in Tables 3.5 and 3.6. In particular, we control for stock characteristics such as firm size, the book-to-market ratio, the stock's beta, the stock's idiosyncratic volatility, momentum, reversal and illiquidity as captured by Amihud's (2002) illiquidity measure, while our option characteristics include open interest and bid-ask spreads for exchange-traded call and put options. Since we decompose the implied volatility spread into the early exercise premium and residual frictions, we begin by examining whether the implied volatility spread itself still contains predictive information for stock returns once we control for stock and option characteristics. Column (1) in Table 3.7 reports these results.

TABLE 3.7 ABOUT HERE

The regression results in column (1) clearly show that the positive and significant relationship between the implied volatility spread and stock returns remains even after controlling for various option- and stock-related characteristics. This confirms our earlier portfolio-sort findings and reinforces the findings in Cremers and Weinbaum (2010). The next step in our analysis is to determine whether the relationship between the implied volatility spread and stock returns is driven by the (friction-induced) early exercise premium or residual frictions. Column 2 reports the results from decomposing the implied volatility spread into the early exercise premium and residual frictions and re-running the regression. The coefficient on the early exercise premium is a statistically significant $-0.21$ ($t$-statistic: $-3.71$) while the coefficient on residual frictions is statistically zero. Taken together, then, the results in this section strongly suggest that the ability of the implied volatility spread to predict stock returns that Cremers and Weinbaum (2010) identify is mainly driven by the early exercise premium, with any residual frictions being insignificant. The question that remains is how robust this finding is. We turn our attention to this in the next section.

## 3.4 Robustness Tests

Our results in the previous section rely heavily on our use of Merton's (1973) result that it is never optimal to early-exercise American call options written on zero-dividend stocks, which in turn allows us to use put-call parity to calculate synthetic European put prices. In this section we investigate how sensitive our results are to deviations from these. We also examine whether our results are robust when we group options into various moneyness and time-to-maturity categories, and whether such factors as nonsynchronous closing times between the stock and options markets affect our findings.

### 3.4.1 Violations of Early Exercise Rules and Put-Call Parity

In calculating the early exercise premium that we have used in our empirical work thus far, we make use of the result in Merton (1973) that it is never optimal to exercise American call options written on zero-dividend stocks early. This result allows us to treat such options as if they were European rather than American. This then allows us to make use of put-call parity to synthetically create European put option prices. This approach is potentially problematic because a number of recent papers (see for example, Jensen and Pedersen (2016), Cao et al. (2017) and Figlewski (2020) among others) show that short-selling constraints in the stock market, and transaction costs in options market (the bid-ask spread in particular) can theoretically make it optimal to exercise American call options on non-dividend-paying stocks early. This casts doubt on our assumption that American call options on zero-dividend stocks can be treated as being equivalent to European ones. In addition, the discussion in Cremers and Weinbaum (2010) shows that in the presence of such market imperfections deviations from put-call parity can widen, casting doubt on whether we are always able to use put-call parity to generate meaningful European put prices.

We begin by looking at the effect of stock short-selling constraints on our results. We follow Jensen and Pedersen (2016) and use the Daily-Cost-to-Borrowing Score (DCBS) as our measure of short-selling constraints. The DCBS is an integer ranging from one to ten, with a high value indicating that a stock has greater short-selling constraints while stocks with a DCBS below five are generally considered to be easier to short. To see why the DCBS is useful as a measure of short-selling constraints here, it is interesting to note that according to Jensen and Pedersen (2016), far less than one percent of all deep in-the-money American call options written on zero-dividend stocks with a DCBS value below five are exercised early, whereas almost ten percent of those same options on stocks with a DCBS value of ten are early-exercised.

TABLE 3.8 ABOUT HERE

Table 3.8 reports the results from quintile portfolio sorts on stock-level early exercise premia conditional on the DCBS value at the start of the return period. We calculate early exercise premia at the stock level using options written on stocks with a non-missing DCBS value and a value less than or equal to eight, seven and five respectively. Since the DCBS value is only widely available from the start of 2004, our sample for the results in Table 3.8 starts in 2004. The results in Table 3.8 suggest that the return on the "High−Low" spread portfolio is not dependent on the level of short selling constraint: the spread portfolio return from sorting on the early exercise premium regardless of a stock's DCBS value, for example, is −0.59% per month ($t$-statistic: −3.02) while for stocks with a DCBS value less than or equal to five, it is −0.54% ($t$-statistic: −2.76). It seems that short-selling constraints have little impact on our results.

We next study the impact of American call option illiquidity along with stock short-selling constraints on our results. Following Cao and Han (2013), and Christoffersen, Goyenko, Jacobs and Karoui (2018), we measure call illiquidity as the option's bid-ask spread scaled by its price, calculated at the end of each month $t-1$. We take the observations for options written on zero-dividend stocks with a DCBS value less than or equal to five at the end of month $t-1$ and sort them into deciles, quintiles and terciles based on illiquidity. The reason for using deciles, quintiles and terciles is to see whether the results are robust to narrower and broader levels of illiquidity. We then exclude those options with the highest levels of illiquidity (the most illiquid decile, the most illiquid quintile and so forth) and calculate stock-level early exercise premia using the remaining observations. We then, as in our earlier analysis, form value-weighted quintile stock portfolios using these early exercise premia as the sort variable and hold these portfolios over month $t$. Table 3.9 reports average excess returns for the portfolios formed from sorting on the early exercise premium conditioning on stock short-selling constraints and American call option illiquidity.

Table 3.9 About Here

The results in Table 3.9 suggest that returns on the spread portfolio formed from stocks considered relatively easy to short remain largely unchanged even after we exclude observations where the American call has a higher probability of being exercised early due to its illiquidity. The "High−Low" portfolio return when we use deciles for illiquidity with very illiquid options then filtered out, for example, stands at −0.61% per month ($t$-statistic: −2.71) while the corresponding return when we use terciles for illiquidity is −0.60 ($t$ statistic: −2.61.) This is similar to what we observe in Table 3.8 where the return for the sample of stocks with a DCBS value less than or equal to five is −0.54% ($t$-statistic: −2.76) without filtering out very illiquid options.

### 3.4.2 Option Moneyness and Time-to-Maturity

In our earlier empirical tests, we calculated value-weighted stock-level early exercise premia using open interest as the weighting variable, regardless of the individual option's moneyness and time-to-maturity. In this subsection we examine whether our earlier findings on the ability of the early exercise premium to predict stock returns is driven by moneyness and/or time-to-maturity. To do this, at the end of each sample month $t-1$ we split the universe of American-European put option pairs into independently double-sorted portfolios according to moneyness and time-to-maturity. Specifically, we first sort the option pairs into in-the-money (ITM), at-the-money (ATM) and out-of-the-money (OTM) portfolios.[3] We then split them into three portfolios according to whether option time-to-maturity is less than 30 days, lies between 30 and 60 days, or is above 60 days. The intersection yields the double-sorted portfolios. We then use option observations

[3] We define an option as ITM if the ratio of the strike price to the stock price is greater than 1.05. An option is treated as ATM if this ratio is greater than or equal to 0.95 and less than or equal to 1.05, and as OTM if the ratio is less than 0.95.

from each of these portfolios to calculate stock-level early exercise premia from which we then form value-weighted quintile portfolios which we hold over month $t$. Table 3.10 reports the results.

<u>Table 3.10 About Here</u>

The results suggest that not all early exercise premia can significantly predict the cross section of future stock returns. Looking at the results for ITM options (Panel A of Table 3.10), the average monthly return for the "High−Low" portfolio is only statistically different from zero for options that are between 60 and 120 days from maturity; returns on the spread portfolios for ITM options that are nearer to maturity are statistically zero. For ATM options, the average monthly return for the "High−Low" portfolio is only statistically different from zero for options that are near maturity; returns on the spread portfolios for ATM options that are further away from maturity are statistically zero. For OTM options, average returns on the spread portfolios are all statistically zero. The results in Table 3.10, then, suggest that the predictive ability of the early exercise premium appears to be confined to near-maturity ATM options and far-from-maturity ITM options.

### 3.4.3 Further Checks

We also conduct a number of additional tests to confirm the robustness of our main empirical results. We start by examining whether our findings are driven by the fact that the stock and option markets have separate closing times.[4] This nonsynchronicity, as shown in the literature (see, for example, Battalio and Schultz (2006)), can drive violations of put-call parity in the market, thereby allowing for wider implied volatility spreads and early exercise premia for exchange-traded American options. To examine whether nonsynchronicity drives our results, we sort stocks into portfolios using stock-level early exercise premia at the end of each month $t-1$. However, we only start cumulating stock returns from the next

[4] The options market closes two minutes after the stock market.

day, avoiding the overnight return as we move from month $t-1$ into month $t$, the holding month. These results are reported in Panel A of Table 3.11. We also examine whether an alternative definition of the stock-level early exercise premium can impede the return predictability that we observe in our main test. To this end, we follow Shang (2017) and calculate weighted average stock-level spreads using the dollar value of open interest rather than the number of contracts outstanding. These results are reported in Panel B of Table 3.11. Panel C examines whether our results are driven by the inclusion of negative early exercise premia in our sample while Panel D examines whether the predictability we find is persistent by using a two-month holding period for the portfolios.

Table 3.11 About Here

The results in Panel A suggests that our empirical evidence is not due to nonsynchronicity between stock and option market closures. Excluding the overnight returns still delivers a mean value-weighted return of −0.71% per month ($t$-statistic: −4.36) on the "High−Low" portfolio when sorting on the early exercise premium while the return when sorting on frictions remains insignificant. Using the dollar value of open interest to calculate the weights used in the stock-level early exercise premia does not change our earlier findings about the predictive ability of early exercise premia: we observe a mean monthly "High−Low" portfolio return of −0.74% ($t$-statistic: −3.73) when using the dollar value of open interest compared to a significant −0.71% using the number of contracts outstanding as the measure of open interest (see Table 3.4 above). The results in Panel C show that our findings are not driven by the presence of negative early exercise premia while the results in Panel D suggest that the predictability we observe does not extend beyond one-period-ahead.

## 3.5 The Role of Informed Trading

In this section, we examine whether informed trading can explain the ability of the early exercise premium to predict stock returns. The motivation for this comes from several papers in the literature (see Cremers and Weinbaum (2010) and An et al. (2014), for example) that, based on the sequential trade model of Easley et al. (1998), explain the ability of the implied volatility spread to predict stock returns through option investors' informed trading activities. Informed investors, having access to private information, first trade in the options market, ahead of the stock market. The result of this is that because information first appears in the options market, the options market leads the stock market and hence explains why the implied volatility spread predicts stock returns. Our results establish that for exchange-traded American options, early exercise premia form a significant part of implied volatility spreads. A natural question to ask, therefore, is whether the return predictability we observe for the early exercise premium can also be explained by informed trading.

Following Shang (2017), we create a sub-sample of options only including observations with zero trading volumes, keeping in mind that if informed trading explains our results we should not observe any return predictability from early exercise premia in this sub-sample. We then repeat the analysis in Table 3.4 on this sample. The results from this analysis are presented in Table 3.12.

<u>Table 3.12 About Here</u>

The results show that, even when the sample consists solely of options with zero trading volume, stock returns significantly decrease as the early exercise premium increases. For instance, the "High−Low" spread portfolio formed from the early exercise premium sort produces a value-weighted mean excess return of −0.63% ($t$-statistic: −3.65) per month, indicating that the ability of the early exercise premium to predict stock returns that we identify is not solely attributable to the informed trading activities of option investors. Interestingly, we also observe somewhat significant predictive ability for frictions, with the friction-sorted "High−Low" spread portfolio delivering a mean monthly excess return of 0.39% ($t$-statistic: 2.11). All in all, the evidence in Table 3.12 suggests that the informed trading activities of option investors cannot fully explain the ability of the early exercise premium to predict returns that we observe.

## 3.6 Concluding Remarks

In this paper we investigate the drivers of the ability of the implied volatility spread to predict stock returns documented in the literature. A number of studies have described this predictability as arising through the informed trading activities of option investors and/or frictions in the underlying stock. We decompose the implied volatility spread into a friction-induced early exercise premium and any remaining (residual) frictions and examine the extent to which the ability of the implied volatility spread to predict returns is driven by the early exercise premium, defined as the difference between the American put price and it's equivalent European price. Given that all individual stock options are American, we make use of put-call parity and Merton's (1973) finding that it is never optimal to exercise an American call on a non-dividend-paying stock early to calculate a synthetic European put price from which we calculate the early exercise premium that we use in our empirical tests. The error from a regression of implied volatility spreads on the early exercise premium provides our measure of residual frictions. Using a barrage of empirical tests and robustness checks we show that it is the early exercise premium that is a significant predictor of stock returns while frictions rarely possess any significant predictive power. Further tests show that our empirical conclusions are not driven by stock short-sale constraints, options illiquidity, the definition of the stock-level early exercise premium or nonsynchronicity between the stock and options markets. We finally show that the ability of the early exercise premium to predict stock returns is not solely driven by option investors' informed trading activities.

# References

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5(1), pp.31-56.

An, B. J., Ang, A., Bali, T. G., and Cakici, N., 2014. The joint cross section of stocks and options. *The Journal of Finance*, 69(5), pp.2279-2337.

Bakshi, G., Kapadia, N., and Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1), pp.101-143.

Bakshi, G., and Madan, D., 2000. Spanning and derivative-security valuation. *Journal of Financial Economics*, 55(2), pp.205-238.

Bali, T. G., and Hovakimian, A., 2009. Volatility spreads and expected stock returns. *Management Science*, 55(11), pp.1797-1812.

Battalio, R., and Schultz, P., 2006. Options and the bubble. *The Journal of Finance*, 61(5), pp.2071-2102.

Black, F., and Scholes, M., 1973. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), pp.637-654.

Cao, W., Ederington, L.H., and Yadav, P.K., 2020. What is the value of being American? *Working Paper*.

Cao, J., and Han, B., 2013. Cross section of option returns and idiosyncratic stock volatility. *Journal of Financial Economics*, 108(1), pp.231-249.

Carhart, M.M., 1997. On persistence in mutual fund performance. *The Journal of Finance*, 52(1), pp.57-82.

Christoffersen, P., Goyenko, R., Jacobs, K., and Karoui, M., 2018. Illiquidity premia in the equity options market. *The Review of Financial Studies*, 31(3), pp.811-851.

Conrad, J., Dittmar, R. F., and Ghysels, E., 2013. Ex ante skewness and expected stock returns. *The Journal of Finance*, 68(1), pp.85-124.

Cremers, M., and Weinbaum, D., 2010. Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis*, 45(2), pp.335-367.

Easley, D., O'hara, M., and Srinivas, P. S., 1998. Option volume and stock prices: Evidence on where informed traders trade. *The Journal of Finance*, 53(2), pp.431-465.

Fama, E.F., and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), pp.3-56.

Fama, E.F., and French, K.R., 2015. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), pp.1-22.

Fama, E.F., and MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), pp.607-636.

Figlewski, S., 2020. An American call is worth more than a European call: The value of American exercise when the market is not perfectly liquid. *Working Paper*.

Goncalves-Pinto, L., Grundy, B. D., Hameed, A., van der Heijden, T., and Zhu, Y., 2020. Why do option prices predict stock returns? The role of price pressure in the stock market. *Management Science*.

Hiraki, K., and Skiadopoulos, G. S., 2020. The contribution of frictions to expected returns. *Working Paper*.

Hull, J. C., 2012. *Options, futures and other derivatives*. Prentice Hall.

Jensen, M. V., and Pedersen, L. H., 2016. Early option exercise: Never say never. *Journal of Financial Economics*, 121(2), pp.278-299.

Manaster, S., and Rendleman Jr, R. J., 1982. Option prices as predictors of equilibrium stock prices. *The Journal of Finance*, 37(4), pp.1043-1057.

Merton, R.C., 1973. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1), pp.141-183.

Newey, W.K., and West, K.D., 1987. A simple positive-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.

Pan, J., and Poteshman, A. M., 2006. The information in option volume for future stock prices. *The Review of Financial Studies*, 19(3), pp.871-908.

Roll, R., Schwartz, E., and Subrahmanyam, A., 2010. O/S: The relative trading activity in options and stock. *Journal of Financial Economics*, 96(1), pp.1-17.

Shang, D., 2017. Is trading what makes prices informative? Evidence from option markets. *Working Paper*.

Zivney, T.L., 1991. The value of early exercise in option prices: An empirical investigation. *Journal of Financial and Quantitative Analysis*, 26(1), pp.129-138.

## Table 3.1: Descriptive Statistics for the Option Data

The table presents descriptive statistics on the percentage month-end spreads between American put options and their equivalent synthetic European options (the Early Exercise Premium, column (1)), and between American call implied volatilities and their equivalent American put volatilities (the Implied Volatility Spread, column (2)). The table further reports the moneyness (column (3)) and time-to-maturity (column (4)) of the option pairs along with the open interest, measured as the number of contracts outstanding, for American call and put options (columns (5) and (6), respectively). The descriptive statistics include the mean, the standard deviation (StDev), the $t$-statistic testing the null hypothesis that the mean is zero (Mean/StError), several percentiles, and the number of observations. The observation-pairs used in columns (1) and (2) are matched along the moneyness and time-to-maturity dimension, so that each observation-pair in column (1) corresponds to exactly one pair in column (2) with the same moneyness and time-to-maturity. We calculate moneyness as the ratio of the option strike price to the stock price. We measure time-to-maturity in terms of calendar days. With the exception of the $t$-statistic, we calculate each statistic as the time-series average of the respective cross-sectional statistic.

| | Early Exercise Premium (in %) | Implied Volatility Spread (in %) | Moneyness (by Option Pair) | Days to Maturity (by Option Pair) | Call Open Interest | Put Open Interest |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Mean | 7.82 | −1.06 | 1.03 | 54 | 1483 | 842 |
| StdDev | 31.35 | 18.64 | 0.08 | 32 | 4730 | 3388 |
| Mean/StdErr | [6.11] | [−9.12] | | | | |
| Percentile 1 | −56.34 | −20.74 | 0.83 | 19 | 3 | 1 |
| Percentile 5 | −24.64 | −9.16 | 0.88 | 19 | 12 | 5 |
| Quartile 1 | −3.27 | −2.46 | 0.98 | 22 | 85 | 32 |
| Median | 4.41 | −0.53 | 1.03 | 49 | 313 | 131 |
| Quartile 3 | 14.43 | 1.03 | 1.09 | 78 | 1103 | 533 |
| Percentile 95 | 49.79 | 5.52 | 1.17 | 111 | 6176 | 3485 |
| Percentile 99 | 113.30 | 12.50 | 1.19 | 111 | 19178 | 11540 |
| Observations | 6,932 | 6,932 | 6,932 | 6,932 | 6,932 | 6,932 |

## Table 3.2: Average Preformation Characteristics of Portfolios Sorted on the Early Exercise Premium

The table presents preformation characteristics of the stock portfolios sorted on the early exercise premium (the spread between the American put option and its equivalent synthetic European option), measured at the stock level. At the end of each sample month $t$, we first sort stocks into portfolios according to the quintile breakpoints of the stock-level early exercise premium. We then calculate various stock and option related characteristics of these portfolios based on information available at time $t$. The figures reported are averages of the relevant characteristic. The stock characteristics include: market capitalization of the stocks (Market Size); their beta, calculated over the prior 60 months (Stock Beta); the value-weighted portfolio return over month $t-1$ to $t$ (Lag Stock Return); and idiosyncratic volatility, calculated from the standard deviation of the residuals from the Fama-French-Carhart regression model (Carhart, 1997) estimated over the prior 60 months (Stock Volatility). The option characteristics are the percentage early exercise premium and percentage implied volatility spread, both measured at the stock level.

| | Early Exercise Premium-Sorted Quintile Portfolios | | | | |
|---|---|---|---|---|---|
| | 1 (Low) | 2 | 3 | 4 | 5 (High) |
| Market Size (in $m) | 7,438 | 8,958 | 9,475 | 9,163 | 7,297 |
| Stock Beta | 1.29 | 1.33 | 1.32 | 1.27 | 1.17 |
| Lag Stock Return (in %) | 0.65 | 0.63 | 1.08 | 1.59 | 2.21 |
| Stock Volatility | 0.39 | 0.39 | 0.38 | 0.37 | 0.37 |
| Early Exercise Premium (in %) | −14.65 | −1.66 | 3.04 | 8.33 | 27.50 |
| IV Spread (in %) | 4.28 | 0.84 | −0.88 | −2.52 | −6.62 |

## Table 3.3: Decomposing the Implied Volatility Spread

The table presents the results of a Fama-MacBeth (1973) regression of the month $t$ stock-level implied volatility spread on the month $t$ stock-level early exercise premium. Observations on the implied volatility spread and the early exercise premium are matched, so that both correspond to the same underlying stock at $t$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12. Adj. $R^2$ is the adjusted $R^2$.

| Dependant Variable: Implied Volatility Spread | |
|---|---|
| | (1) |
| Early Exercise Premium | $-0.19$ |
| | [$-46.99$] |
| Intercept | $-1.92\times10^{-3}$ |
| | [$-3.65$] |
| Adj. $R^2$ | 38.47% |

## Table 3.4: Univariate Portfolio Sorts

The table presents the mean returns of value-weighted stock portfolios in excess of the three-month Treasury Bill rate, sorted on the implied volatility spread, the early exercise premium and residual frictions, all measured at the stock level. At the end of month $t-1$, we sort stocks into portfolios according to the quintile breakpoints of each of the three sorting variables. We also form a spread portfolio long the top and short the bottom quintile ("High−Low"). We then hold the portfolios over month $t$. Observations on the implied volatility spread and the early exercise premium are matched, so that both correspond to the same underlying stock at $t-1$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

| | Value-weighted Quintile Portfolios | | | | | |
|---|---|---|---|---|---|---|
| Sorting Variables | 1 (Low) | 2 | 3 | 4 | 5 (High) | High−Low |
| | Mean Monthly Excess Portfolio Return (in %) | | | | | |
| Implied Volatility Spread | −0.07 | 0.56 | 0.37 | 0.75 | 1.01 | 1.08 |
| | [−0.16] | [1.40] | [0.95] | [2.26] | [3.06] | [4.41] |
| Early Exercise Premium | 0.91 | 0.78 | 0.53 | 0.40 | 0.19 | −0.71 |
| | [2.67] | [2.07] | [1.41] | [1.04] | [0.48] | [−4.34] |
| Residual Frictions | 0.53 | 0.51 | 0.62 | 0.42 | 0.66 | 0.13 |
| | [1.06] | [1.15] | [1.90] | [1.15] | [1.95] | [0.62] |

## Table 3.5: Portfolio Alphas

The table presents the mean $\alpha$ for each of the stock portfolios sorted separately on the implied volatility spread (Panel A), the early exercise premium (Panel B) and residual frictions (Panel C), all measured at the stock level. The $\alpha$ is the intercept estimated from three different regression models: the Market-model (Market), where portfolio returns in excess of the three month treasury bill rate are regressed on the market factor; the Fama-French-Carhart four-factor model (FFC), where portfolio excess returns are regressed on the market, size, book-to-market and momentum factors (Carhart (1997)); and the Fama and French (2015) five-factor model (FF5), where portfolio excess returns are regressed on the market, size, book-to-market, profitability and investment factors. At the end of each sample month $t-1$, we sort stocks into portfolios according to the quintile breakpoints of each of the three sorting variables. We also form a spread portfolio long the top and short the bottom quintile ("High−Low"). We then hold the portfolios over month $t$. Observations on the implied volatility spread and the early exercise premium are matched, so that both correspond to the same underlying stock at $t-1$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

| | Value-weighted Quintile Portfolios | | | | | |
|---|---|---|---|---|---|---|
| | Mean Monthly Portfolio Alpha (%) | | | | | |
| | 1(Low) | 2 | 3 | 4 | 5(High) | High−Low |
| | Panel A: Sorting Variable - Implied Volatility Spread | | | | | |
| Market | −0.51 | 0.18 | 0.04 | 0.40 | 0.65 | 1.16 |
| | [−1.53] | [0.66] | [0.14] | [1.50] | [2.04] | [5.82] |
| FFC | −0.30 | 0.38 | 0.21 | 0.58 | 0.97 | 1.27 |
| | [−0.93] | [1.44] | [0.84] | [2.25] | [3.23] | [6.59] |
| FF5 | −0.03 | 0.63 | 0.46 | 0.75 | 1.17 | 1.20 |
| | [−0.09] | [2.34] | [1.79] | [2.80] | [3.75] | [5.98] |
| | Panel B: Sorting Variable - Early Exercise Premium | | | | | |
| Market | 0.53 | 0.43 | 0.17 | 0.05 | −0.18 | −0.71 |
| | [1.76] | [1.55] | [0.60] | [0.20] | [−0.64] | [−3.85] |
| FFC | 0.84 | 0.66 | 0.40 | 0.22 | −0.07 | −0.91 |
| | [3.00] | [2.51] | [1.47] | [0.87] | [−0.25] | [−5.34] |
| FF5 | 1.10 | 0.89 | 0.64 | 0.41 | 0.16 | −0.94 |
| | [3.84] | [3.27] | [2.29] | [1.54] | [0.58] | [−5.27] |
| | Panel C: Sorting Variable - Frictions | | | | | |
| Market | 0.07 | 0.14 | 0.30 | 0.09 | 0.31 | 0.25 |
| | [0.20] | [0.47] | [1.18] | [0.33] | [1.09] | [1.22] |
| FFC | 0.41 | 0.36 | 0.48 | 0.24 | 0.45 | 0.04 |
| | [1.30] | [1.32] | [1.98] | [0.96] | [1.60] | [0.21] |
| FF5 | 0.78 | 0.67 | 0.64 | 0.39 | 0.73 | −0.05 |
| | [2.43] | [2.41] | [2.55] | [1.50] | [2.52] | [−0.28] |

## Table 3.6: Portfolios Double-Sorted on the Early Exercise Premium and Residual Frictions

The table presents the mean returns of stock portfolios in excess of the three-month Treasury Bill rate, double-sorted on the early exercise premium and residual frictions, both measured at the stock level. Panel A reports results from first sorting on residual frictions and then sorting on the early exercise premium. At the end of each sample month $t-1$, we sort stocks into portfolios according to the quintile breakpoints of the frictions variable and within each friction-sorted portfolio, we then further sort based on quintile breakpoints of the early exercise premium. Panel B reports results from reversing the order of the sorts, that is, at $t-1$, we sort stocks into portfolios according to the quintile breakpoints of the early exercise premium and within each of these portfolios, we then further sort based on quintile breakpoints of the residual frictions variable. We also form spread portfolios long the "High" and short the "Low" quintile ("High−Low") along the second sort dimension in both panels. We hold the portfolios over month $t$. Observations on the implied volatility spread and the early exercise premium are matched, so that both correspond to the same underlying stock at $t-1$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

Panel A: 1st Sort - Residual Frictions; 2nd Sort - Early Exercise Premium

| | Value-weighted Quintile Portfolios | | | | |
|---|---|---|---|---|---|
| | Friction | | | | |
| Premium | 1(Low) | 2 | 3 | 4 | 5(High) |
| | Mean Monthly Excess Portfolio Return (in %) | | | | |
| 1(Low) | 0.79 | 0.62 | 0.96 | 1.13 | 1.09 |
| | [1.55] | [1.53] | [2.74] | [3.53] | [2.55] |
| 2 | 0.68 | 0.37 | 0.83 | 0.62 | 1.14 |
| | [1.18] | [0.72] | [2.13] | [1.70] | [3.29] |
| 3 | 0.02 | 0.71 | 0.57 | 0.55 | 0.75 |
| | [0.04] | [1.52] | [1.68] | [1.36] | [2.00] |
| 4 | 0.18 | 0.64 | 0.44 | 0.34 | 0.76 |
| | [0.36] | [1.21] | [1.28] | [0.83] | [1.93] |
| 5(High) | −0.68 | 0.00 | 0.44 | 0.11 | 0.40 |
| | [−1.35] | [0.00] | [1.19] | [0.28] | [0.99] |
| High−Low | −1.47 | −0.62 | −0.51 | −1.02 | −0.69 |
| | [−4.06] | [−2.66] | [−2.15] | [−4.05] | [−2.54] |

Panel B: 1st Sort - Early Exercise Premium; 2nd Sort - Residual Frictions

| | Value-weighted Quintile Portfolios | | | | |
|---|---|---|---|---|---|
| | Premium | | | | |
| Friction | 1(Low) | 2 | 3 | 4 | 5(High) |
| | Mean Monthly Excess Portfolio Return (in %) | | | | |
| 1(Low) | 0.91 | 0.50 | 0.68 | −0.12 | −0.81 |
| | [2.24] | [0.94] | [1.32] | [−0.22] | [−1.65] |
| 2 | 0.68 | 0.83 | 0.67 | 0.36 | 0.04 |
| | [1.71] | [1.94] | [1.34] | [0.71] | [0.07] |
| 3 | 1.05 | 0.96 | 0.29 | 0.59 | 0.16 |
| | [3.39] | [3.28] | [0.71] | [1.36] | [0.34] |
| 4 | 1.22 | 0.93 | 0.57 | 0.15 | 0.22 |
| | [3.34] | [2.36] | [1.75] | [0.35] | [0.62] |
| 5(High) | 1.22 | 0.70 | 0.50 | 0.43 | 0.48 |
| | [2.87] | [1.78] | [1.10] | [1.21] | [1.09] |
| High−Low | 0.31 | 0.20 | −0.19 | 0.55 | 1.28 |
| | [0.91] | [0.56] | [−0.61] | [1.45] | [3.77] |

## Table 3.7: Fama-MacBeth (1973) Regressions

The table presents the results of Fama-MacBeth (1973) regressions in which the dependent variable is excess stock returns over month $t$. The first regression (Column (1)) includes the implied volatility spread as the main explanatory variable along with several control variables capturing various stock and option characteristics. The second regression (Column (2)) decomposes the implied volatility spread into the early exercise premium and residual frictions. The stock characteristics that we include as control variables are: Size, as measured by the natural logarithm of the market capitalization of the stock; the book-to-market ratio; the beta of the stock, estimated over the prior 60 months; idiosyncratic stock volatility, calculated from the standard deviation of the residuals from the Fama-French-Carhart regression model (Carhart, 1997) estimated over the prior 60 months; momentum, measured as the cumulative stock return over months $t-11$ to $t$; reversal, measured as the stock return over months $t-1$ to $t$; and the Amihud (2002) stock illiquidity measure, calculated as the absolute stock return divided by its dollar volume. The option characteristics we use as control variables are: put and call open interest; and the bid-ask spreads of the put and call options. All of the independent variables are dated $t-1$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

| | Regression Model: | |
|---|---|---|
| | (1) | (2) |
| Dependant Variable: Monthly Excess Stock Return | | |
| Early Exercise Premium | | −0.21 |
| | | [−3.71] |
| Residual Frictions | | 0.08 |
| | | [0.42] |
| Implied Volatility Spread | 0.05 | |
| | [4.23] | |
| Size | 0.16 | 0.16 |
| | [2.87] | [2.89] |
| Book-to-Market | 0.35 | 0.37 |
| | [1.81] | [1.87] |
| Stock Beta | −0.82 | −0.93 |
| | [−0.36] | [−0.41] |
| Idiosyncratic Volatility | 0.21 | 0.20 |
| | [2.49] | [2.45] |
| Momentum | 0.46 | 0.46 |
| | [1.68] | [1.68] |
| Reversal | −0.23 | −0.23 |
| | [−3.44] | [−3.37] |
| Stock Illiquidity | 0.50 | 0.51 |
| | [2.32] | [2.34] |
| Put Open Interest | 0.74 | 0.74 |
| | [1.75] | [1.75] |
| Call Open Interest | −0.12 | −0.12 |
| | [−0.63] | [−0.66] |
| Put Bid-Ask Spread | −0.13 | −0.15 |
| | [−0.98] | [−1.15] |
| Call Bid-Ask Spread | 0.24 | 0.25 |
| | [1.07] | [1.12] |
| Intercept | $-2.23\times10^{-3}$ | $-2.23\times10^{-3}$ |
| | [−2.77] | [−2.77] |

## Table 3.8: Controlling for Stock Short-Selling Constraints

The table presents the mean excess returns of value-weighted stock portfolios sorted on the early exercise premium, measured at the stock level. The sample consists of observations on stocks with either non-missing Daily-Cost-to-Borrowing Score (DCBS) values or ones equal to or below the DCBS values of eight, seven and five at the end of month $t-1$. At the end of month $t-1$, we sort stocks into portfolios according to the quintile breakpoints of the stock-level early exercise premium. We also form a spread portfolio long the top and short the bottom quintile ("High−Low"). We then hold the portfolios over month $t$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

| | Value-weighted Quintile Portfolios | | | | | |
|---|---|---|---|---|---|---|
| | Sorting Variable: Early Exercise Premium | | | | | |
| | 1(Low) | 2 | 3 | 4 | 5(High) | High−Low |
| | Mean Monthly Excess Portfolio Return (in %) | | | | | |
| With All DCBS Values | 0.93 | 0.71 | 0.57 | 0.47 | 0.34 | −0.59 |
| | [2.11] | [1.90] | [1.32] | [1.11] | [0.70] | [−3.02] |
| With DCBS Value $\leq 8$ | 0.92 | 0.72 | 0.55 | 0.48 | 0.37 | −0.55 |
| | [2.08] | [1.93] | [1.26] | [1.13] | [0.77] | [−2.73] |
| With DCBS Value $\leq 7$ | 0.93 | 0.72 | 0.56 | 0.47 | 0.38 | −0.55 |
| | [2.10] | [1.92] | [1.30] | [1.10] | [0.80] | [−2.73] |
| With DCBS Value $\leq 5$ | 0.92 | 0.72 | 0.60 | 0.48 | 0.39 | −0.54 |
| | [2.10] | [1.91] | [1.40] | [1.11] | [0.81] | [−2.76] |

## Table 3.9: Controlling for Stock Short-Selling Constraints and Option Illiquidity

The table presents the mean excess returns of value-weighted stock portfolios sorted on the early exercise premium, measured at the stock level. The sample consists of observations on stocks with a Daily-Cost-to-Borrowing Score (DCBS) equal to or below five at the start of the stock return period. At the end of month $t-1$, we first sort American and synthetic European option pairs into decile, quintile and tercile portfolios based on the scaled bid-ask spread (our proxy for option illiquidity) of the American call option. We then exclude the highest illiquidity option portfolio from the sample and calculate stock-level early exercise premia using the remaining observations. We then sort stocks into value-weighted portfolios according to the quintile breakpoints of these early exercise premia. We also form a spread portfolio long the top and short the bottom quintile ("High−Low"). We hold the portfolios over month $t$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

| Sample With DCBS Value $\leq 5$ | | | | | | |
|---|---|---|---|---|---|---|
| | Value-weighted Quintile Portfolios | | | | | |
| | Sorting Variable: Early Exercise Premium | | | | | |
| | 1(Low) | 2 | 3 | 4 | 5(High) | High−Low |
| Excluding Highest Illiquidity | Mean Monthly Excess Portfolio Return (in %) | | | | | |
| in Decile Sort | 0.97 | 0.72 | 0.50 | 0.55 | 0.36 | −0.61 |
| | [2.20] | [1.91] | [1.14] | [1.32] | [0.74] | [−2.71] |
| in Quintile Sort | 0.98 | 0.67 | 0.54 | 0.56 | 0.34 | −0.64 |
| | [2.25] | [1.76] | [1.26] | [1.35] | [0.69] | [−2.89] |
| in Tercile Sort | 0.95 | 0.73 | 0.55 | 0.56 | 0.34 | −0.60 |
| | [2.21] | [1.83] | [1.27] | [1.37] | [0.71] | [−2.61] |

## Table 3.10: Controlling for Option Moneyness and Time-to-Maturity

The table presents the mean excess returns of value-weighted stock portfolios sorted on the early exercise premium, measured at the stock level. We calculate a number of early exercise premia for each individual stock using observations on options from various moneyness and time-to-maturity categories. At the end of each sample month $t-1$, we first sort options into portfolios according to whether their strike-to-stock price ratio ("moneyness") lies above 1.05 (Panel A), between 0.95 and 1.05 (Panel B), or below 0.95 (Panel C). Within each moneyness portfolio, we next sort them into portfolios according to whether their days-to-maturity are between 10 and 30, between 30 and 60, or above 60 days. For each stock, we then use option observations from each of these moneyness and time-to-maturity-sorted portfolios separately to calculate stock-level early exercise premia. We finally sort stocks into portfolios separately according to the quintile breakpoints of each version. We also form separate spread portfolios long the top and short the bottom quintile ("High−Low"). We then hold the portfolios over month $t$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

| | Value-weighted Quintile Portfolios | | | | | |
|---|---|---|---|---|---|---|
| | Sorting Variable: Early Exercise Premium | | | | | |
| Days-to-Maturity | 1(Low) | 2 | 3 | 4 | 5(High) | High−Low |
| | Mean Monthly Excess Portfolio Return (in %) | | | | | |
| Panel A: In-The-Money (Strike-to-Stock Price > 1.05) | | | | | | |
| 10 - 30 | 0.59 | 0.99 | 0.56 | 0.54 | 0.41 | −0.18 |
| | [1.40] | [2.44] | [1.25] | [1.22] | [0.81] | [−0.70] |
| 30 - 60 | 0.45 | 0.48 | 0.68 | 0.65 | 0.14 | −0.31 |
| | [0.95] | [1.04] | [1.49] | [1.76] | [0.30] | [−1.42] |
| 60 - 120 | 1.03 | 0.57 | 0.60 | 0.47 | 0.41 | −0.61 |
| | [2.28] | [1.21] | [1.16] | [1.14] | [0.81] | [−2.46] |
| Panel B: At-The-Money (Strike-to-Stock Price 0.95 to 1.05) | | | | | | |
| 10 - 30 | 0.92 | 0.58 | 0.64 | 0.45 | 0.15 | −0.77 |
| | [2.33] | [1.55] | [1.69] | [1.15] | [0.36] | [−3.83] |
| 30 - 60 | 0.63 | 0.63 | 0.45 | 0.44 | 0.29 | −0.33 |
| | [1.47] | [1.51] | [1.26] | [1.18] | [0.65] | [−1.31] |
| 60 - 120 | 0.94 | 0.90 | 0.48 | 0.34 | 0.49 | −0.45 |
| | [2.02] | [2.22] | [1.05] | [0.79] | [0.88] | [−1.59] |
| Panel C: Out-Of-The-Money (Strike-to-Stock Price < 0.95) | | | | | | |
| 10 - 30 | 0.81 | 1.33 | 0.53 | 0.33 | 0.11 | −0.70 |
| | [1.04] | [1.86] | [0.77] | [0.52] | [0.14] | [−1.08] |
| 30 - 60 | 0.72 | 0.45 | 0.44 | 0.55 | 0.39 | −0.32 |
| | [1.31] | [0.90] | [0.88] | [1.01] | [0.70] | [−1.05] |
| 60 - 120 | 0.88 | 0.85 | 0.40 | 0.26 | 0.57 | −0.31 |
| | [1.80] | [1.96] | [0.78] | [0.50] | [0.98] | [−1.27] |

## Table 3.11: Additional Robustness Tests

The table presents the mean excess returns of value-weighted stock portfolios sorted on the implied volatility spread, the early exercise premium and residual frictions, all measured at the stock level. At the end of month $t-1$, we sort stocks into portfolios according to the quintile breakpoints of each of the three sorting variables. We also form a spread portfolio long the top and short the bottom quintile ("High−Low"). We then cumulate stock returns for each portfolio starting from the next day rather than starting from the portfolio formation day, that is, we exclude the overnight return from the portfolio formation day. These results are presented in Panel A. Panel B reports results from using the dollar value of open interest when calculating the weights used to calculate the implied volatility spreads and early exercise premia. Panel C reports the returns for portfolios sorted separately on positive and negative values of the stock-level early exercise premia. For these three panels, we hold the portfolios over month $t$. Finally, Panel D reports returns for early exercise premium-sorted portfolios over the second month after portfolio formation. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

| | Value-weighted Quintile Portfolios | | | | | |
|---|---|---|---|---|---|---|
| Sorting Variables | 1(Low) | 2 | 3 | 4 | 5(High) | High−Low |
| | Mean Monthly Excess Portfolio Return (in %) | | | | | |
| Panel A: Excluding Overnight Returns | | | | | | |
| Implied Volatility Spread | −0.09 | 0.58 | 0.37 | 0.76 | 1.00 | 1.09 |
| | [−0.19] | [1.49] | [0.94] | [2.29] | [3.06] | [4.34] |
| Early Exercise Premium | 0.91 | 0.78 | 0.51 | 0.43 | 0.19 | −0.71 |
| | [2.67] | [2.10] | [1.38] | [1.11] | [0.48] | [−4.36] |
| Residual Frictions | 0.52 | 0.54 | 0.60 | 0.42 | 0.67 | 0.15 |
| | [1.04] | [1.23] | [1.85] | [1.16] | [1.99] | [0.68] |
| Panel B: Using Dollar Value of Open Interests | | | | | | |
| Implied Volatility Spread | −0.06 | 0.31 | 0.59 | 0.68 | 0.96 | 1.01 |
| | [−0.13] | [0.81] | [1.70] | [1.73] | [2.69] | [5.91] |
| Early Exercise Premium | 0.97 | 0.70 | 0.54 | 0.39 | 0.22 | −0.74 |
| | [3.06] | [1.83] | [1.47] | [1.01] | [0.55] | [−3.73] |
| Residual Frictions | 0.44 | 0.53 | 0.65 | 0.46 | 0.66 | 0.22 |
| | [1.00] | [1.44] | [1.99] | [1.19] | [1.75] | [1.22] |
| Panel C: Using Positive/Negative Early Exercise Premium | | | | | | |
| with '+' Premium | 0.67 | 0.58 | 0.45 | 0.38 | 0.21 | −0.46 |
| | [1.95] | [1.55] | [1.07] | [0.95] | [0.49] | [−2.38] |
| with '−' Premium | 0.80 | 0.75 | 0.77 | 0.60 | 0.58 | −0.22 |
| | [1.90] | [1.73] | [1.94] | [1.45] | [1.42] | [−1.19] |
| Panel D: Mean Excess Return Over Month $t+2$ (in %) | | | | | | |
| Early Exercise Premium | 0.60 | 0.63 | 0.65 | 0.63 | 0.40 | −0.20 |
| | [1.66] | [1.95] | [1.61] | [1.77] | [0.89] | [−0.97] |

## Table 3.12: Only Including Observations on Options With Zero Trading Volume

The table presents the mean excess returns of value-weighted stock portfolios sorted on the implied volatility spread, the early exercise premium and residual frictions, all measured at the stock level. We only include option observations with zero trading volume to calculate the implied volatility spread, the early exercise premium and residual frictions for each stock in our sample. At the end of month $t-1$, we sort stocks into portfolios according to the quintile breakpoints of each of the three sorting variables. We also form a spread portfolio long the top and short the bottom quintile ("High−Low"). We hold the portfolios over month $t$. Observations on the implied volatility spread and the early exercise premium are matched, so that both correspond to the same underlying stock at $t-1$. The numbers in square parentheses are $t$-statistics calculated using Newey and West's (1987) autocorrelation and heteroscedasticity-consistent covariance matrix with the lag length for the autocorrelation set at 12.

| | Value-weighted Quintile Portfolios | | | | | |
|---|---|---|---|---|---|---|
| Sorting Variables | 1(Low) | 2 | 3 | 4 | 5(High) | High−Low |
| | Mean Monthly Excess Portfolio Return (in %) | | | | | |
| Implied Volatility Spread | 0.09 | 0.36 | 0.55 | 0.80 | 0.88 | 0.79 |
| | [0.20] | [0.95] | [1.49] | [2.18] | [2.43] | [4.12] |
| Early Exercise Premium | 0.90 | 0.78 | 0.59 | 0.40 | 0.27 | −0.63 |
| | [2.70] | [1.99] | [1.62] | [1.08] | [0.62] | [−3.65] |
| Residual Frictions | 0.35 | 0.51 | 0.42 | 0.56 | 0.74 | 0.39 |
| | [0.79] | [1.32] | [1.18] | [1.71] | [1.81] | [2.11] |