# Early Option Exercise: Never Say Never

Mads Vestergaard Jensen and Lasse Heje Pedersen*

First version: March, 2012.

This version: October 10, 2014

**ABSTRACT**

A classic no-arbitrage result by Merton (1973) is that, except just before expiration or dividend payments, one should never exercise a call option and never convert a convertible bond. We show theoretically that this classic result is overturned when investors face financial frictions. Indeed, early option exercise can be optimal when it reduces 1) short-sale costs, 2) transaction costs, or 3) funding costs. We provide empirical evidence consistent with our theory, documenting billions of dollars worth of early exercise for equity call options and convertible bonds using unique data sets on actual exercise and conversion decisions and actual financial frictions. Our model and the observed frictions can explain as much as 98% of early exercises by market makers and 67% of those by customers of brokers.

*Mads Vestergaard Jensen is at Copenhagen Business School; email: mvj.fi@cbs.dk. Lasse H. Pedersen is at Copenhagen Business School, New York University, AQR Capital Management, CEPR, and NBER; email: lhp.fi@cbs.dk; web: www.lhpedersen.com. We thank Farid Aitsahlia, Mikael Reimer Jensen, Morten Karlsmark, Yang Lu, Jesper Lund, Mark Mitchell, Dmitriy Muravyev, Todd Pulvino, and Bill Silber for helpful comments and discussions, seminar participants at Copenhagen Business School, New York University, HEC Paris, the European Finance Association Conference, and especially Robert Whaley for providing data on the exercise of equity options. Both authors gratefully acknowledges support from the European Research Council (ERC grant no. 312417) and the FRIC Center for Financial Frictions (grant no. DNRF102).

# I. Introduction: Never Exercise a Call and Never Convert a Convertible?

One of the classic laws of financial economics is that call options should never be exercised, except at expiration or just before dividend payments (Merton, 1973) and, similarly, convertible bonds should not be converted early (Brennan and Schwartz (1977), Ingersoll (1977a)). This rule is ubiquitous in option theory and taught in most introductory finance classes. We challenge this classic rule — both theoretically and empirically — in light of financial frictions.

Theoretically, we show that early exercise can be optimal when agents face short-sale costs, transaction costs, or funding costs and we characterize the optimal exercise boundary under such financial frictions. Empirically, we show that investors indeed exercise equity options early and convert convertibles when facing these frictions, using unique data on actual exercise and conversion decisions.

To understand our result, first recall famous arbitrage argument of Merton (1973): Rather than exercising a call option and receive the stock price $S$ less the strike price $X$, an investor is better off shorting the stock, putting the discounted value of $X$ in the money market, and possibly exercising the option at expiration — or selling the option to another agent who can do so. However, this arbitrage argument can break down when shorting is costly (Avellaneda and Lipkin, 2009) or agents face transaction costs or funding costs.

We introduce these financial frictions in a model. We first show that Merton's no-exercise rule holds even with "mild" frictions, meaning either (i) when short-sale costs and funding costs are small (even if transaction costs are large), or (ii) when transaction costs are small and the option price is above the intrinsic value (which can be driven by other agents facing low shorting and funding costs). However, we show that early exercise in fact *is* optimal when frictions are more severe such that the option price net of transaction costs is below intrinsic value and the option owner faces sufficiently high shorting and/or funding costs.

Next, we ask whether an agent who may need to exercise an option early would ever buy one in the first place? To address this, we consider a dynamic equilibrium model in which rational agents trade options in zero net supply while facing financial frictions. In the equilibrium that we derive, some agents indeed buy options and later exercise them before expiration with positive probability.

Finally, we show how the effects of financial frictions can be quantified in a continuous-time model in which the parameters can be directly calibrated to match the data. We show that exercise is optimal when the stock price is above an optimal exercise boundary which we derive. The exercise boundary is decreasing in short-sale costs, margin requirements, and, under certain conditions, funding costs, and increasing in the stock price volatility. Said differently, exercise happens earlier (i.e., for lower stock prices) with larger short-sale costs, larger margin requirements, larger funding costs, and lower volatility.

To intuitively understand our model and to illustrate its clear quantitative implications, consider the example of options written on the Ishares Silver Trust stock (the largest early exercise day in our sample of options on non-dividend paying stocks). Figure 2 shows the stock price of an Ishares trust and the optimal exercise boundary that we derive based on the short-sale cost (or "stock lending fee") and funding costs that we observe in our data. While exercise is never optimal before expiration without frictions, we see that the optimal exercise boundary is finite due to the observed financial frictions. Furthermore, we see investors actually exercise shortly after the stock price crosses our model-implied exercise boundary.

This illustrative example provides evidence consistent with our model, but does it represent broader empirical phenomena? To address this question, we collect and combine several large datasets. For equity options, we merge databases on option prices and transaction costs (OptionMetrics), stock prices and corporate events (CRSP), short-sale costs (Data Explorers), proxies for funding costs, and actual option exercises (from the Options Clearing Corporation). Focusing only on options on non-dividend-paying stocks, we find that 1.8 billion options are exercised early (i.e., before Merton's rule) in the time period from 2003 to 2010, representing a total exercise value of $36.3 billion. (Of course, the amount of exercises before Merton's rule would be larger if we included dividend paying stocks, but for clarity we restrict attention to the most obvious violations.)

Consistent with our theory's qualitative implications, we find that early exercise is more likely when (i) the short-sale costs for the underlying stock are higher, (ii) the option's transaction costs are higher, (iii) the option is more in-the-money, and (iv) the option has shorter time to expiration. These results are highly statistically significant (due to the large amounts of data). Moreover, our data allows us to identify exercises for each of three types of agents, customers, market makers, and proprietary traders. We find that each type of agent exercises options early, including the professional market makers and proprietary traders, and that each type is more likely to do so when frictions are severe, consistent with our theory of rational exercises.

We also test the quantitative implications of the model more directly. For each option that is exercised early, we estimate the optimal exercise boundary by solving our model-implied partial differential equation (PDE) based on the observed frictions. We find that 66–84% of all early exercise decisions in our data happen when the stock price is above the model-implied exercise boundary, depending on how input variables are estimated. The behavior of market makers is most consistent with our model (their exercise decisions coincide with the model-implied prediction in 86-98% of the cases), while customers of brokers make the most exercise decisions that we cannot explain and proprietary traders are in between, consistent with the idea that market makers are the most sophisticated and face the lowest frictions while customers of brokers face the highest frictions.

Furthermore, using logit and probit regressions, we find that real-world investors are more likely to exercise early when the stock price is above the model-implied exercise boundary. This results consistent with the model is both statistically and economically significant: The estimated probability that an option contract is exercised early, cumulated over a 20-trading-day period where the stock price is above the boundary, is 20.7% (21.8%) based on logit (probit) regressions. The corresponding probability when the stock price is below the boundary is 0.4% (0.4%). The large difference in exercise behavior across these two cases is a success for the model, but the numbers also indicate that far from all options are exercised immediately when the stock price goes above our model-implied boundary. To understand the sluggish exercise behavior, note that different agents presumably have different level of frictions and different propensities to check if early exercise is optimal. A vigilant option owner may exercise as soon as the stock price crosses the boundary, leading to a single observation of an early exercise in accordance with theory. A less attentive option owner, or one with lower frictions, might not exercise for several days where the stock price is above the boundary, resulting in many observations of no early exercise. This dependency structure means that the probabilities should be interpreted with caution. A "burn-out" effect (similar to that known from refinancing of mortgages) may imply that the estimated cumulative exercise probability is lower than the unconditional probability of an option being exercised for a random option owner.

For convertible bonds, we combine data on equities and short-sale costs with the Mergent FISD database on convertible bond features and actual conversions. We find 25.4 million early conversions, representing an equity value of \$7.7 billion at conversion. The early conversion rates for convertible bonds is increasing in the short-sale cost of the stock and in the moneyness of the convertible bond, again consistent with our theory, but we note that this data set is smaller and subject to potential errors and inaccuracies.

Our paper complements the large literature following Black and Scholes (1973) and Merton (1973). Option prices have been found to be puzzlingly expensive (Longstaff (1995), Bates (2000), Bates (2003), Jackwerth (2000), Ni (2009), Constantinides, Jackwerth, and Perrakis (2009)) and several papers explain this based on frictions: Option prices are driven by demand pressure (Bollen and Whaley (2004), Garleanu, Pedersen, and Poteshman (2009)), are affected by transaction costs (Brenner, Eldor, and Hauser (2001) and Christoffersen, Goyenko, Jacobs, and Karoui (2011)), short-sale costs (Ofek, Richardson, and Whitelaw (2003), Avellaneda and Lipkin (2009)), funding constraints (Bergman (1995), Santa-Clara and Saretto (2009), Leipold and Su (2012)), embedded leverage (Frazzini and Pedersen (2011)), and interest-rate spreads and other portfolio constraints (Karatzas and Kou (1998), Piterbarg (2010)). We complement the literature on how frictions affect option prices by showing that frictions also affect option exercises.

There also exist papers on option exercises, which document irrational early exercise decisions (Poteshman and Serbin (2003)) and irrational failures of exercise of call options (Pool, Stoll, and Whaley (2008)) and put options (Barraclough and Whaley (2012)). We complement these findings by linking early exercise decisions to financial frictions, both theoretically and empirically, and by drawing a parallel to convertible bonds. Early exercise therefore exist both for rational and irrational reasons; while Poteshman and Serbin (2003) find that customers sometimes irrationally exercise early, we find that market makers and firm proprietary traders also frequently exercise early and that most exercises appear to be linked to financial frictions. Battalio, Figlewski, and Neal (2014) also find that option bid prices can be below intrinsic value, which is necessary condition for optimal early exercise, and our model helps explain why the option price can be this low.

Regarding convertible bonds, the literature has linked their prices to financial frictions (Mitchell, Pedersen, and Pulvino (2007), Agarwal, Naik, and Loon (2011)) and examined whether the companies call these bonds too late (often convertible bonds are also callable, see the literature following Ingersoll (1977b)), while we study early conversions by the owners of the convertible bonds due to financial frictions.

In summary, we characterize how frictions can lead to optimal early exercise of call options and conversion of convertibles, and we provide extensive empirical evidence consistent with our predictions. These findings overturn one of the fundamental laws of finance, providing another example that the basic workings of financial markets are affected by financial frictions with broader implications for economics.

# II. Theory

We are interested in studying when it is optimal to exercise an American call option early, that is, during times other than expiration and days before ex-dividend days of the underlying stock. Such rational early exercises must be driven by frictions since they violate Merton's rule. We first consider a simple model to illustrate how early exercise can be optimal for an investor who is long an option (Section II.A). Section II.B shows that buying and later exercising early indeed can happen in equilibrium. Finally, Section II.C presents a continuous-time model with testable quantitative predictions for early exercise.

## A. When Is Early Exercise Optimal?

Consider an economy with three securities that all are traded at times 0 and 1: a risk-free security with interest rate $r^f > 0$, a non-dividend-paying stock, and an American call option with strike price $X > 0$ that expires at time $t = 1$. The stock price at time $t$ is denoted $S_t$ and the option price $C_t$. The stock price $S_1$ at time 1 can take values in $[0, \infty)$ and is naturally unknown at time 0 . The final payoff of the option is $C_1 = \max(S_1 - X, 0)$.

All agents are rational, wealth-maximizing price takers, subject to financial frictions. Agent $i$ faces a proportional stock transaction cost of $\lambda^{i,S} \in [0,1]$ per dollar stock sold. Furthermore, agent $i$ faces a proportional transaction cost of $\lambda^{i,C} \in [0,1]$ per dollar option sold. If agent $i$ sells the stock short at time $t = 0$, agent $i$ incurs a proportional securities-lending fee of $S_0 L^i$ at time 1, $L^i \geq 0$. If $i$ is long the stock, agent $i$ can lend out the stock and receive a proportional securities-lending fee of $S_0 l^i$ at time 1, where $l^i \in [0, L^i]$. Agent $i$ also faces a funding cost of $F^i(x, y)$ at time 0 if the agent chooses to hold a value of $x \in \mathbb{R}$ of the stock and $y \in \mathbb{R}$ of the option. This funding cost could be due to an opportunity cost associated with binding capital requirement. Naturally, the funding cost is zero if the agent takes a zero position, $F^i(0, 0) = 0$, and increasing in the absolute sizes of $x$ and $y$.[1]

We are interested in whether early exercise can be optimal. We therefore analyze whether a strategy is "dominated." We say that a strategy is dominated if there exists another strategy that generates at least as high cash flows in each time period and in every state of nature, and a strictly higher cash flow in some possible state. Further, early exercise is defined as being dominated if any possible strategy that includes early exercise is dominated. We assume that there exists no pure arbitrage net of transaction costs because such a strategy

[1] Stated mathematically, the funding cost function has the property that for $x_2 \geq x_1 \geq 0$ then $F^i(x_2, y) \geq F^i(x_1, y) \geq 0$ for all $i$ and $y \in \mathbb{R}$. Similarly, if $x_2 \leq x_1 \leq 0$ then $F^i(x_2, y) \geq F^i(x_1, y) \geq 0$ for all $i$ and $y \in \mathbb{R}$, and similarly for the dependence on $y$.

would trivially dominate all other strategies (or, said differently, all strategies are either non-dominated or dominated by a non-dominated strategy).

We first show that, under certain "mild" frictions, early exercise is always dominated. This result extends Merton's classic no-early-exercise rule and shows that the rule is robust to certain frictions. All proofs are in Appendix A

**Proposition 1 (No Exercise with "Mild" Frictions)**

*Early exercise is dominated for an agent $i$ that has:*

*i. zero short-sale and funding costs, i.e., $L^i = F^i = 0$ (regardless of all transaction costs); or*

*ii. a sale revenue of the option above the intrinsic value, $C_0(1-\lambda^{i,C}) > S_0 - X$. A sufficient condition for this high sale revenue is that agent $i$ has zero option transaction costs, $\lambda^{i,C} = 0$, and the existence of another type of agents $j$ with zero short-sale costs, funding costs, and stock transaction costs, $L^j = F^j = \lambda^{j,S} = 0$.*

The first part of this proposition states that transaction costs alone cannot justify rational early exercise. The reasoning behind this is as follows: When the option is exercised it is either to get the underlying stock or to get cash. In the case where the option holder wants cash, exercising early and immediately selling the stock is dominated by hedging the option position through short-selling of the underlying stock and investing in the risk-free security. The transaction cost from selling the stock after early exercise and from selling the stock short are the same so positive transaction costs of the stock cannot in themselves make early exercise optimal.

In the case where the option holder wants stock, early exercise is dominated by holding on to the option, exercising later, and investing the strike price discounted back one period, $\frac{X}{1+r^f}$, in the risk-free asset. Thereby the investor will still get the stock, but on top of that earn interest from the risk-free asset. This strategy does not involve any direct trading with the stock and, hence, is not affected by stock transaction costs. (Note that these two alternative strategies do not involve option transactions and hence dominate early exercise even with high option transaction costs.)

The second part of the proposition states that early exercise is also dominated if the option owner's net proceeds from selling the option exceeds intrinsic value. In this case, the owner is better off by selling the option than by exercising early. If there is a type of agents, $j$, who faces no short-sale costs, no funding costs, and no stock transaction costs then these agents value the option at strictly more than its intrinsic value (as explained above). Therefore, the option holder $i$ prefers selling to $j$ over exercising early if no option transaction costs apply.

While it is important to recognize that frictions need not break Merton's rule, we next show that Merton's rule indeed break down when frictions are severe enough. Specifically, a combination of short-sale costs and transaction costs can make early exercise optimal.

**Proposition 2 (Rational Early Exercise with "Severe" Frictions)**
*Consider an agent $i$ who is long a call option which is in-the-money taking stock transaction costs into account, $S_0(1-\lambda^{i,S}) > X$. Early exercise is not dominated for $i$ if the revenue of selling the option is low, $C_0(1-\lambda^{i,C}) \leq S_0(1-\lambda^{i,S}) - X$ and one of the following holds:*

*a. the short-sale costs, $L^i$, is large enough or*

*b. the funding costs, $F^i$, is large enough.*

*The condition $C_0(1-\lambda^{i,C}) \leq S_0(1-\lambda^{i,S}) - X$ is satisfied if the option transaction cost $\lambda^{i,C}$ is large enough and/or the option price is low enough.*

To understand the intuition behind who early exercise can be optimal, consider an option owner who wants cash now (with no risk of negative cash flows at time 1). Such an agent can either (i) sell the option, (ii) hedge it, or (iii) exercise early. Option (i) is not attractive (relative to early exercise) if the sale revenue after transaction costs is low. Further, option (ii) is also not attractive if the funding costs or short-sale costs (or those in combination) make hedging very costly. Therefore, option (iii), early exercise, can be optimal.

Note that a low option price can itself be a result of frictions. For instance, the option price is expected to be low if all agents face high short-sale costs and can earn lending fees from being long stocks as we explore further in the next sections.

## B. Equilibrium Model with Rational Early Exercise

We have seen that it can be optimal to exercise early for a option owner who faces financial frictions. However, would such an agent ever buy the option in the first place? And if so, could such an equilibrium early exercise happen with "reasonable" parameter values? We show that the answer to both questions is "yes." To do so, we present a simple dynamic equilibrium model with endogenous prices, allocations, and exercise decisions. We seek to keep the example simple, providing a more quantitative analysis in the next section.

The economy has two types of competitive agents, $A$ and $B$. Each agent maximizes his expected time-2 wealth utilities subject to the constraint that wealth must almost be non-negative (no default). The representative agent $A$ has a total initial cash endowment of 200 and $B$ a cash endowment of 800.

The agents trade three assets at times $t \in \{0, 1, 2\}$, namely a stock, a risk-free asset with interest rate $r^f = 0.1\%$ per period, and an American call option with strike price $X = 100$ and expiration at time $t = 2$.

The value of the stock evolves in a binomial tree as seen in the Figure 1 below. The stock's time-2 payoff is given by its exogenous final payoff as seen in the figure. The option expires in the money only in the case of the best outcome for the stock. At time $t = 1$, there are two possible states, an "up" state $u$ with a more positive outlook for the stock relative to the "down" state $d$. The time-1 stock prices, $S_u$ and $S_d$, in each state are endogenous. The agents have differences of opinions regarding the relative probabilities of the different states. Agent $i \in \{A, B\}$ assigns a probability $p_t^i$ to the event that the stock prospects evolve favorably in the next period (an "up" move in the tree) at time $t \in \{0, 1\}$. Specifically, $p_0^A = p_1^B = 0.46$ and $p_0^B = p_1^A = 0.45$ so that agents have similar views, but agent $A$ is more optimistic at time 0 while $B$ is more optimistic at time 1.

The markets are subject to the following frictions. When the agents trade an option, the option buyer pays $C_0$ while the option seller receives $(1 - \lambda^C)C_0$, where the transaction cost is $\lambda^C = 2\%$. The transaction cost for stock trades is zero ($\lambda^S = 0$). The stock short-sale fee is $L = 0.4\%$ per period, such that to borrow a stock at time $t$, the borrower must pay $LS_t$

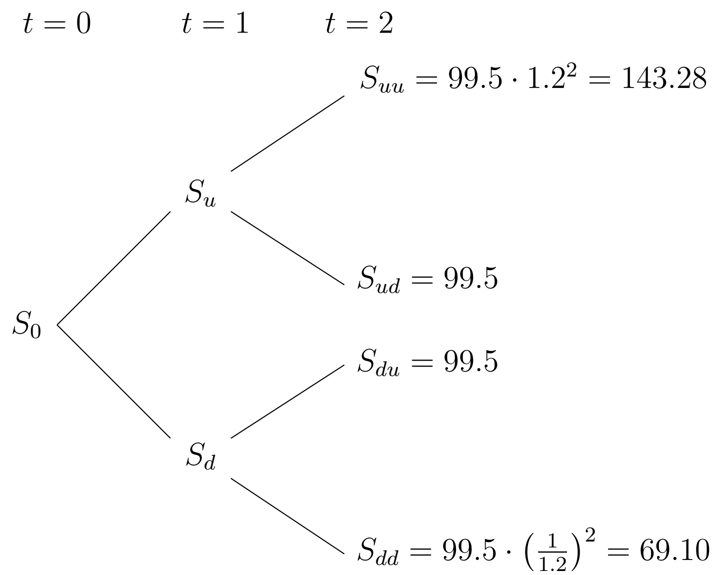


Figure 1: Evolution of the value of the stock.

at time $t + 1$. The securities lender does not receive a fee, $l = 0$.

**Proposition 3 (Early exercise in equilibrium)**

*An equilibrium exists in which agent A buys option contracts at time zero and exercises early at time 1 if the stock's prospects improve (i.e., in the "up" state).*

The intuition behind the result is as follows. At time $t = 0$, the optimistic $A$ agents want as much leverage as possible and therefore buy options written by $B$ agents. At time $t = 1$ the picture changes and $B$-types have a more optimistic view on the stock than $A$-types. This change in relative views means that agents want to shift stock exposures, which involves an early option exercise by $A$ agents. This early exercise in equilibrium is driven by the frictions. First, without short-sale costs, $A$ agents would not exercise the options, but short the stock instead. Second, option transaction costs high enough to make early exercise preferable to trading options at time $t = 1$ (while at the same time low enough to make option trade attractive at time $t = 0$).

The parameters in this example do not appear unreasonable in the sense of being extreme. To see that, note first that if we think of the period length as 0.1 years (about a month), then the annual risk-free rate is 1%. The assumed annualized stock price volatility is 58%, which is reasonable for an individual stock.[2] The annualized short-sale fee is 4%, which is a common fee for the stocks "on special" (the about 5-10% stocks that are the most difficult to sell short). Lastly, the assumptions of risk neutrality and minor differences of opinions are harder to evaluate, but, again, the difference of opinions are not extreme. This is but one example of equilibrium early exercise, it is possible to construct others, including with risk averse agents.

## C. Optimal Exercise Boundary and Comparative Statics

We have seen qualitatively how early exercise can arise when frictions are large enough. Next, we consider a model that is realistic enough – and tractable enough – that we can use its *quantitative* implications in our empirical analysis. We solve for the optimal exercise boundary in a continuous-time model in which all parameters have clear empirical counterparts. Hence, in our empirical analysis, for each option and each date, the model provides a clear prediction regarding whether early exercise is optimal or not. The model solution also allows us to derive interesting comparative statics, showing how the exercise decision depends on the moneyness of the option, short-sale costs, funding costs, margin requirements, and the volatility of the underlying stock.

The optimal exercise decision is closely connected to the rational valuation of American options in the context of financial frictions. Hence, we seek to joint solve for the value of the option and the optimal exercise decision. We start in the classic Black-Scholes-Merton framework, where agents can invest in a risk-free money-market rate of $r^f$ and a stock with price process $S$ given by:

$$\mathrm{d}S(t) = S(t)\mu \mathrm{d}t + S(t)\sigma \mathrm{d}W(t) \tag{1}$$

[2]Recall that in a standard binomial model with proportional stock price moves of $u = 1.2$ and $1/u$, it is well known that $u = e^{\sigma\sqrt{\Delta t}}$, where $\Delta t$ is the length of a period in years and $\sigma$ is the annualized volatility. Therefore, $\sigma = \log(u)/\sqrt{\Delta t} = 58\%$.

where $\mu$ is the drift, $\sigma$ is the volatility, and $W$ is a Brownian motion. The stock can be traded without cost, but we consider the following financial frictions. First, agents face short-sale costs, modeled based on standard market practices: To sell the stock short, the agent must borrow the share and leave the short-sale proceeds as collateral. The collateral account earns an earns the interest rate $r^f - L^i$, called the "rebate rate." The fact that the rebate rate is below the money-market rate reflects an (implicit) continuous short-sale cost of $L^i$ (called the "rebate rate specialness") multiplied by the value of the stock. The securities lender — the owner of the share — holds the cash and must pay a continuous interest of $r^f - l^i$ . Since he can invest the cash in the money market, this corresponds to a continuous securities-lending income of $l^i \in [0, L^i]$. We allow that the securities-lending fees depend on the agent $i$, and that lender earns less than the short-seller pays ($l^i < L^i$) since the difference is lost to intermediaries (custodians and brokers) and search costs and delays.[3]

The second friction that we consider is funding costs, modeled as a margin requirement, $M^i$ given by:

$$M^i(x, y) = m^{i,S}|x| + m^{i,C}|y| \tag{2}$$

where $x \in \mathbb{R}$ and $y \in \mathbb{R}$ are the values held in the stock and option, respectively and $m^{i,C} > m^{i,S} \geq 0$ are the respective margin requirements. The margin requirement is positive for both long and short positions and the option faces larger proportional margin requirements than the stock due to its larger risk (embedded leverage). The margin requirement matters for the option pricing and exercise when the agent faces capital costs. We assume that agent $i$'s capital cost is $\psi^i \geq 0$ in the sense that using his own equity for a risk free investment is associated with an opportunity cost of $r^f + \psi^i$. Such a capital cost can arise from costly equity financing and from a binding capital constraint.[4]

Based on these relatively realistic assumptions, we are interested in the value $C$ and

[3]The institutional details of short selling and the over-the-counter securities-lending market are described in Duffie, Gârleanu, and Pedersen (2002) who also discuss why not all investors can immediately lend their shares in equilibrium.

[4]See Garleanu and Pedersen (2011) for an equilibrium model with binding margin requirements where such implicit capitals costs arise endogenously as $\psi^i$ is the Lagrange multiplier of the margin requirement.

optimal exercise of an American option with expiration $T$ and strike price $X$. The option value depends on the time $t$ and the stock price $S$ so we apply Itô's lemma to write the option price dynamics as:

$$\mathrm{d}C(t) = \left(C_t + \frac{1}{2}\sigma^2 S^2 C_{SS}\right)\mathrm{d}t + C_s \mathrm{d}S(t) \tag{3}$$

We solve for the option value and exercise strategy that makes the agent indifferent with respect to buying an (additional) option on the margin. (As we will see, the answer depends on the agent's existing portfolio and frictions.) To do this, we consider the portfolio dynamics of buying one option, hedging by selling $C_S$ shares of the stock, and fully financing the strategy based on margin loans and the use of equity capital. The value of this fully-financed strategy evolves as according to:

$$\left(C_t + \frac{1}{2}\sigma^2 S^2 C_{SS}\right)\mathrm{d}t + C_S \mathrm{d}S(t) - (1 - m^{i,C})Cr^f \mathrm{d}t - m^{i,C}C(r^f + \psi^i)\mathrm{d}t \tag{4}$$

$$- C_S \mathrm{d}S(t) + C_S S\left(-\tilde{m}^{i,S} r^f + \tilde{m}^{i,S}(r^f + \psi^i) + (r^f - \tilde{l}^i)\right)\mathrm{d}t \tag{5}$$

Let us carefully explain each of the terms in this central expression. The first two terms simply represent the dynamics of the option (as seen in Eqn. (3)). The next two terms represent the funding of the option. Specifically, $(1 - m^{i,C})C$ can be borrowed against the option at the money-market funding cost $r^f$. The remaining option value, the margin requirement $m^{i,C}C$, must be financed as equity at a rate of $r^f + \psi^i$.

The second line of (5) represents the terms stemming from the stock position and its financing. The first term is the stock dynamics, given the $C_S$ number of shares sold. The last three terms capture the various financing costs. These terms depend on whether the agent is already long the stock or short the stock. In the former case, the stock sale arising from the hedge of the option reduces the agent's stock position, freeing up capital and reducing securities-lending income. In the latter case, the option hedge increases the size of the agent's short stock position, using capital and paying short-sale costs. To capture both these situations in one equation, Eqn. (5) uses the notation $\tilde{m}^{i,S}$ and $\tilde{l}^i$ to capture a generalized notion of margin requirements and securities lending fees, defined as follows:

$$(\tilde{m}^{i,S}, \tilde{l}^i) = \begin{cases} (m^{i,S}, l^i) & \text{if agent } i \text{ is long stock,} \\ (-m^{i,S}, L^i) & \text{if agent } i \text{ is short stock,} \end{cases} \tag{6}$$

Consider first the situation where agent $i$ is long stock. In this case, the stock sold to hedge the option implies a margin relief, reducing the need for equity financing by $C_S S m^{i,S}$, with the cost rate of $r^f + \psi^i$. The smaller stock position simultaneous reduces by $C_S S m^{i,S}$ the amount on the margin account earning rate $r^f$. Further, the shares sold decrease by $C_S S$ the amount on the securities-lending account with interest $r^f - l^i$. If agent $i$ is short the stock the sign with respect to the margin requirement flip as the short-sale represents an increase in equity need and a deposit on the margin account. Further, when the agent is already short, the stock sale leads to an increase in the short-sale cost $L^i$ (rather than $l^i$).

Since the stock position is chosen to offset the risk of the option, the stochastic terms involving $\mathrm{d}W(t)$ cancel out in (5). As the portfolio is fully financed and the change in value is deterministic, the drift must also be zero. Setting the drift equal to zero gives the following partial differential equation (PDE) for the option value:

$$C_t + \frac{1}{2}\sigma^2 S^2 C_{SS} - (r^f + m^{i,C}\psi^i)C + C_S S_t(\tilde{m}^{i,S}\psi^i + r^f - \tilde{L}^i) = 0 \tag{7}$$

The option value is subject to the conditions

$$\begin{aligned} C(T) &= \max(S(T) - X, 0) \\ C(t) &\geq \max(S(t) - X, 0) \quad \text{for } t < T \end{aligned} \tag{8}$$

The first condition is the standard boundary condition for the value of the option at the expiration date $T$. The second boundary condition is that the option value cannot fall below the intrinsic value before expiration because then it would be more optimal to exercise early. Indeed, early exercise occurs when the inequality binds.

We see that the PDE is mathematically equivalent to that arising from the pricing of an American call option in a Black-Scholes-Merton model with a modified interest rate and a continuous dividend yield. Specifically, we see that $r^f + m^{i,C}\psi^i$ plays the role of the interest rate and $\tilde{l}^i + \psi^i(m^{i,C} - \tilde{m}^{i,S})$ corresponds to a dividend yield. This means that the partial differential equation (PDE) can be solved by traditional numerical methods, as we do in our empirical analysis. Further, we can analytically characterize the optimal exercise behavior by utilizing results from Dewynne, Howison, Rupf, and Wilmott (1993) as stated in the following proposition. (The proposition focuses on the case of an agent who is short the stock, while similar intuitive results hold when the agent is long, as reported in Appendix A.)

**Proposition 4 (The Optimal Exercise Boundary and Comparative Statics)** *Consider an agent $i$ who is long an option which he cannot sell above its intrinsic value and short the stock. If his funding cost or short-sale cost is strictly positive ($\psi^i > 0$ or $L^i > 0$) then there exists a finite optimal exercise boundary, $B_{T-t} < \infty$, which solves (7)–(8). Hence, the agent optimally exercises as soon as the stock price is above the boundary, $S_t > B_{T-t}$. The following comparative statics hold for the limit boundary for long-dated options $B^* := \lim_{\tau\to\infty} B_\tau$. The exercise boundary*

*i. decreases in short-sale cost ($\frac{\partial B^*}{\partial L^i} < 0$)*

*ii. decreases in funding costs ($\frac{\partial B^*}{\partial \psi^i} < 0$)*

*iii. decreases in option margin requirements ($\frac{\partial B^*}{\partial m^{i,C}} \leq 0$, $\frac{\partial B^*}{\partial m^{i,C}} < 0$ if $\psi^i > 0$)*

*iv. decreases in marginal stock margin requirement ($\frac{\partial B^*}{\partial m^{i,S}} \leq 0$, $\frac{\partial B^*}{\partial m^{i,S}} < 0$ if $\psi^i > 0$)*

*v. increases in the volatility ($\frac{\partial B^*}{\partial \sigma} > 0$).*

The first part of the proposition establishes that there exists and optimal exercise boundary and that exercise is optimal if and only if the stock prices is higher than this level. It is natural that early exercise happens for large stock prices for several reasons. First, exercise is only relevant if the option is in the money. Second, the value of the optionality becomes smaller when the option is deeper in the money (as eventual exercise becomes ever more certain). Third, the cost of hedging increases with moneyness as both the stock price and the number of stocks needed to be shorted become larger. Fourth, the money "tied up" in the option increases in the moneyness.

Proposition 4 further provides several intuitive comparative statics with respect to the optimal exercise boundary. We further study how the exercise boundary depends on the short-sale cost, funding cost, margin requirements, and volatility in Figure 4. In each of the four panels, the time to expiration $T - t$ is on the $x$-axis and the scaled exercise boundary is on the $y$-axis. Specifically, we scale the exercise boundary $B_t$ by the strike price $X$, which makes it a more intuitive number. For example, a value of $B_t/X = 1.6$ means that early exercise is optimal when the stock price is at least 60% in the money. Said differently, early exercise happens when $S_t \geq B_t$ or, equivalently, when $S_t/X \geq B_t/X$, and the latter scaled measure is more intuitive. Clearly a lower exercise boundary corresponds to an earlier optimal exercise decision (for lower moneyness of the stock price). The figures vary the model parameters around the following base-case parameters for an agent who is long a call option and shorts the stock: the risk-free rate is $r^f = 2\%$, the volatility is $\sigma = 40\%$, the funding cost is $\psi^i = 1\%$, the short-sale cost is $L^i = 1\%$, the margin requirement for the stock is $m^{i,S} = 50\%$, and margin requirement for the option is $m^{i,C} = 100\%$.

The top left panel shows that higher short-sale costs correspond to earlier optimal exercise (consistent with the general result in Proposition 4, part i). This result is natural as short-sale costs make it costly to hedge the option. The top right panel shows that higher funding costs also make it optimal to exercise early, namely to free up capital (as in Proposition 4, part ii). The bottom left panel shows that a higher option margin requirement encourages an earlier exercise (as in Proposition 4, part iii). Finally, the bottom right panel shows that early exercise is delayed with higher volatility (as in Proposition 4, part v). To see why, recall that a higher volatility increases the value of optionality, therefore making it less attractive to exercise early.

In all the graphs we see that the optimal exercise boundary decreases in the time to expiration. In fact, it is a general result that the exercise boundary must be weakly decreasing in time to expiration. To understand this result, note that if it is optimal to exercise a longer-dated option, then it must also be optimal to exercise a shorter-dated one (since you give up less optionality).

Finally, the figures provide *quantitative* insights into when we should expect early exercise due to frictions. In the base-case, early exercise is optimal when the stock price is 1.67 times the strike price a quarter before expiration and 1.27 times the strike price 10 days before expiration. We next turn to the empirical analysis, where we also implement our model for each option in a large data set and analyze whether the real-world exercise decisions occur when the stock price is above the exercise boundary that we calculate.

# III. Data and Preliminary Analysis

This section describes our data sources, provides summary statistics, and outlines our empirical methodology. We start with the data and then turn to the summary statistics, which already shows large amounts of early exercises and early conversions both in terms of number of contracts and in terms of dollar value.

## A. Data

Our study combines a number of very large data sets as described in Table I. For equity options, we combine the OptionMetrics database on U.S. option prices and option bid-ask spreads with the CRSP tape of U.S. equity prices and corporate events. We use data on the cost of short-selling stocks from Data Explorers, focusing on their Daily Cost of Borrow Score (DCBS), which is an integer from 1 to 10 with 1 indicating a low cost of shorting and 10 indicating a high one.

We analyze actual exercise behavior using data originally from the Options Clearing Corporation (OCC).[5] This data contains the number of contract exercises for each option series each day. The daily exercises can be separated into three groups of market participants,

[5] We are very grateful to Robert Whaley for providing this data.

namely exercises done by customers of brokers (retail customers and hedge funds), market makers, and firm proprietary traders. The option exercise data runs from July 2001 to September 2008. The data is missing in the months of November 2001, January and July 2002, and January 2006.

Finally, we use the Mergent FISD database on convertible bonds. This database provides time and amount of conversions together with total outstanding amount for convertible bonds.

## B. Sample Selection

To identify option exercises that clearly violate Merton's rule, we focus on early exercise of options on stocks that do not pay dividends. In particular, we exclude any option series if OptionMetrics reports a non-zero forecast of future dividends during any day of the life time of the option. Further, we exclude observations of option series on the day of expiry to focus on early exercise. Lastly, we exclude options that do not follow the standard practice of the having an expiration date on the third Friday in a given month (this excludes only a tiny fraction).

The data from OptionMetrics is merged with the exercise data from OCC on date, ticker, option ticker, strike price, and expiry date. For each option series, we further merge the data with that of the corresponding stock from CRSP and Data Explorers based on CUSIP.

We further clean the data in a number of ways. We exclude any option series where (i) the underlying at some point in the life time of the option experienced a distribution event according to CRSP (split, dividend, exchanges, reorganizations etc.); (ii) OptionMetrics has records with different strike prices for the same series, different underlying identifiers (secid or CUSIP), or different expiry dates (indicating data errors or changes in the contract); (iii) no data is available on the last trading day before expiry (indicating some possible outside event); (iv) OptionMetrics has several records for the same series for the same day; (v) settlement is special, e.g. AM-settlement; (vi) CRSP has a missing closing price for the stock at some trading date; or (vii) lacks a matching observation of the underlying stock in

Data Explorers on any date.

We note that the OCC data only has records of exercises, meaning that option series that never experienced an exercise (before or at expiry) are not part of that sample. Hence, by requiring a match between Option Metrics and the OCC data, our sample only includes options that were exercised at some point. An alternative approach is to include our entire OptionMetrics sample and assume that option series missing in the OCC data were never exercised. Since we do not know whether the OCC data is complete, neither approach is perfect. Either way we do it, we find the same number of early exercises and these exercises are linked to financial frictions.

Convertible bond data is acquired through Mergent FISD. Our sample only includes convertible bonds that at some point in time were converted (including at a dividend, at maturity, or as a response to a call). If Mergent FISD has not been able to identify the exact day of a conversion they set the date to the end of the quarter or even fiscal year (the latter seems to be the case only rarely). This makes it difficult to identify whether the conversion happened on the day before ex-dividend or not in these cases. To avoid problems related to these issues, we only include bonds where the underlying did not have any distribution events (including dividend payments) during the sample period, using data of distribution events from CRSP. We also exclude bonds where the underlying is first observed in CRSP more than one day later than the offering date of the bond. Furthermore, we exclude bonds that at some point had an exchange offer or a tender offer, or where the underlying is not either a common stock or an American depository stock. If Mergent FISD data has no maturity date or conversion price of the bond, then it is also excluded. The original conversion prices of the bonds are recorded in FISD and through the cumulative adjustment factors provided by CRSP they can be updated to reflect any changes e.g. due to stock splits.[6] We only include observations from days where bond holder had the right to convert early. If CRSP data for the underlying is missing starting at some point in time, we exclude the observations from five days before this happened and onwards, to avoid inclusion of conversions related

[6]If e.g. a stock is split in two, a convertible bond with this stock as underlying will have its conversion price halved at the same time.

to some kind of exogenous corporate event. If the day of the initial observation of the bond in Mergent FISD is after the offering day of the bond, we exclude observations before this initial observation. If the bond has been partly or fully called at some point in time, then we exclude all observations which are less than three months earlier or three months later to this event. This measure is taken in order to avoid inclusion of conversions that are a response to a call from the issuer (and hence not early conversion initiated by the bond holder), though it is not guaranteed that we catch all such events. Likewise, if the record shows any reorganization or exchange of the bond, observations from three months before this event and onwards are excluded.

## C. Summary Statistics

Table II provides summary statistics for our final sample. We see a substantial amount of early exercises and conversions. Panel A reports the total number of early exercises of equity options, that is, exercises that violate Merton's rule. Our data contain 1,806 million early exercises, representing a total exercise value of $36.3 billion or a total intrinsic value of $22.8 billion. Naturally, the exercises are concentrated among in-the-money options, especially deep-in-the-money options. Our data does contain a small fraction of exercised out-of-the-money options, which could be due to measurement error or investors exercising to save transaction costs when they want the actual stock. Measurement error may occur for instance when options are exercised during the day and we measure the moneyness based on the end-of-day price.

Table II also shows that early exercises are concentrated among shorter-term options. This finding is consistent with our theory, since short-term options have less optionality, but it could also be driven by the simple fact that there is a larger open interest of such options. Our formal empirical tests therefore consider the number of exercises as a fraction of the open interest.

The final part of Panel A in Table II shows that all types of agents exercise early for billions of dollars. Customers of brokers exercise early with a total strike price of $16.0

billion, firm proprietary traders for a total of about \$2.0 billion, and market makers for a total of \$18.3 billion.

Panel B in Table II reports the total number of early conversions of convertible bonds. In our data, 25 million bonds are converted early, corresponding to about \$5.7 billion worth of principal or \$7.7 billion of equity value. A few conversions of out-of-the-money bonds are seen, which relates to the definition of moneyness applied. A convertible bond is considered out-of-the-money if the price of the underlying stock is less than conversion price, in-the-money if stock price is up to 25% above conversion price, and deep-in-the-money if stock price is more than 25% above conversion price. Conversion price is defined by the principal amount of bond that must be converted to get one stock. Our definition of moneyness is not perfect: Indeed, the market value of a non-convertible bond can deviate some from the face value, so it might actually be attractive to convert even when it is out-of-the-money according to the above definition. While this adds noise to our analysis, it should not drive our conclusions.

Figure 3 shows the evolution of relative share of early option exercise over time for the three different agent types observed. The picture is dominated by Customers (which includes retail customers and hedge funds) and Market Makers. Interestingly, the share of early exercise for the professional Market Makers has increased over time. We next discuss how we use this exercise and conversion data to test our theoretical predictions.

## D. Variables of Interest and Methodology

We are interested in the fraction of options that are exercised and how this relates to our theoretical predictions. For each day $t$ and each option series $i$ in our sample, we define the options that are exercised $EX$ as a fraction of the open interest $OI$ on the close of the day before.

$$EX_t^i = \frac{\#\text{exercised options}_t^i}{\max\{OI_{t-1}^i, \#\text{exercised options}_t^i\}} \tag{9}$$

We take the maximum in the denominator to ensure that $EX$ is between zero and one, including the rare instances when the number of exercises is greater than the open interest the day before (which must be due to options that are bought and exercised on the same day or data errors). Similarly, in our logit and probit regressions, we compare the number of exercises to the number of "trials" given by $\max\{OI^i_{t-1}, \#\text{exercised options}^i_t\}$.

For each daily observation of an option series, we measure the option transaction costs as the relative bid-ask spread constructed in the following way:

$$TCOST^i_t = \frac{\text{ask price}^{s(i)} - \text{bid price}^{s(i)}}{(\text{ask price}^{s(i)} + \text{bid price}^{s(i)})/2} \tag{10}$$

Here, the superscript $i$ denotes the option series and $s(i)$ is the corresponding at-the-money option series with the same underlying stock, defined as the series with the smallest absolute difference between stock price and strike price (where $s(i)$ may be equal to $i$ itself). We use the at-the-money option instead of the option itself to avoid endogeneity issues. The possibility of exercising the option early will itself affect bid and ask prices, especially for deep-in-the-money options. The bid price will in such cases often go below, but not much below, the intrinsic value. As a result the intrinsic value will somewhat be a floor for how much the bid price can go below the ask price. Our focus is to test how the general level of transaction costs of an option series affects early exercise in the first place.

We measure the short-sale fee, $L$, as follows. For each stock and each date, we observe its Daily Cost of Borrow Score (DCBS), which is an integer score from from 1 to 10. We map this DCBS score to a short-sale fee level by using the median among all stocks with this DCBS score that have data on both their DCBS score and their fee level.

The model-implied optimal exercise boundary $B^i_t$ for any option $i$ on day $t$ is computed as follows. We numerically solve the PDE (7)–(8) that takes frictions into account using the observed characteristics on any date $t$. The stock volatility $\sigma$ is set as the 60-day average historical volatility, the risk-free rate $r^f$ is the Fed Funds rate, the margin requirement of the call option is set to 50%, the funding cost $\psi$ is set as the LIBOR-OIS interest-rate spread, and the short-sale and funding costs are as defined above. We measure the stock price $S$ as the closing price on the given day.

Similarly to the exercise measure $EX$ for equity options, we are interested in the daily converted bonds as a fraction of the total outstanding amount. We define $CONV_t$ as amount converted on day $t$ divided by the sum of amount converted and outstanding amount after the conversion. (Equivalently, the denominator is the amount outstanding before the conversion.)

$$CONV_t^i = \frac{\text{Amount converted}_t^i}{\text{Amount converted}_t^i + \text{Amount outstanding after conversion}_t^i} \tag{11}$$

# IV. Empirical Results: Never Exercise a Call Option?

We turn to formally testing the link between early options exercises and financial frictions. We first sort the exercises by short-sale costs and transaction costs to analyze the connection between early exercises and financial frictions in a simple way. Next, we test the model more directly by considering whether the stock price is above or below the model-implied optimal exercise boundary at the time of exercise. Finally, we use multivariate logit and probit regressions to analyze how the propensity to exercise can be explained by the model and how it depends on the joint effects of a number of option characteristics.

Table III and Figure 5 show how the fraction of early option exercises $EX$ (defined in (9) above) varies with short-sale costs. We see that the fraction of early exercise decisions increases monotonically in the short-sale cost, consistent with our model. Among options with minimal short-sale costs, the fraction of options exercised early is only 0.17%, while among options with the highest short-sale costs, the fraction exercised is above 4%. As seen in the table, the difference in these two extreme groups is highly statistically significant.

Table III and Figure 5 further consider the exercises broken down by moneyness. Naturally, there are virtually no out-of-the-money options that are exercised (which would clearly be irrational or due to measurement error), and the exercises are concentrated among deep-in-the-money options, as we would expect. Splitting the data by expiration, the table and figure show that the exercises are more frequent for shorter-term options. This is consistent with our theory as the benefits of postponing exercising is smaller (smaller optionality) for shorter-term options. Again we see that the option exercises increase in short-sale costs within expiration group.

Lastly, we split the data by agent types, that is, across customers of brokers (retail customers and hedge funds), market makers, and firm proprietary traders. We see that all types of agents exercise early and more so when the short-sale costs are higher. We note that the absolute magnitude of the numbers should not be compared across groups for the following reason: Our data do not contain open interest by agent type, so we measure the number of exercises by each agent type as a fraction of the total open interest. Hence, the fraction of exercises by firm proprietary traders may be low simply because this agent type trades few options relative to the total open interest. In any case, the pattern of an increasing propensity to exercise as short-sale costs increase is consistent with our theory.

Table IV and and Figure 6 show how the fraction of early exercises varies with the transaction costs. We measure transaction costs as the relative bid-ask spread ($TCOST$) of the at-the-money with the same expiration as defined in (10) above. Options are classified in three groups: out-of-the-money options (with stock price below strike price), in-the-money options (with stock price up to 25% above strike price), and deep-in-the-money options (with stock price more than 25% above strike price).

We see that the fraction of options exercised increases monotonically with the transaction costs. This pattern holds overall, for each moneyness group, each expiration group, and each type of agents. The absolute of numbers in Table IV are smaller on average than the numbers in Table III. This is because 81% of the data have a short-sale cost code (DCBS) equal to 1 (as classified by Data Explorers) and, as expected, the fraction of exercises is low in this group as seen in Table III. In Table IV, our groups by transaction costs are more balanced.

In summary, consistent with our model's qualitative predictions, we have seen that option exercises increase in short-sale costs and transaction costs and that these patterns hold within groups sorted by moneyness, expiration, and agent types. Next, we seek to test our model's quantitative predictions. In particular, for each exercised option, we compare the model-implied optimal exercise boundary to the stock price at the day of the exercise as seen in

Table V. The table reports the fraction of exercises consistent with our model for each agent type and for all agents (in each row of the table). Each column of the table corresponds to a specific set of assumptions underlying the model, with increasingly conservative assumptions going from left to right. In the left-most column, we estimate each stock's volatility based on the 60-day realized volatility. Given that volatility is mean-reverting and may spike up temporarily during significantly news announcements, the next column uses a lower estimate of future volatility, namely the minimum of the current 60-day volatility and its median in the OptionMetrics sample (June 1 2001–January 31 2012). The third column uses both the conservative volatility estimate and a conservative estimate of short-sale costs, namely the 90th percentile of short-sale costs within each group (rather than the median observed cost).

In all cases, we see that the majority of option exercises happen when the stock price is above the model-implied optimal exercise boundary. The fraction of exercises consistent with our model is highest for market makers, which could be because these agents face the lowest financial frictions and are the most active market participants. Naturally, the fraction of exercises consistent with the model increases in the columns with more conservative assumptions (by construction). The model can explain 98% of the market makers' exercises with the most conservative assumptions, a very large fraction in light of the remaining noise.[7]

Lastly, we study the propensity to exercise in a logit and probit regression setting. To do so, we need in principle to compute the model-implied optimal exercise boundary for each type of option and each date, including days when no exercises are observed. The very large amount of data combined by the numerical complexity in solving our model's PDE makes such a complete analysis unfeasible. To address this issue, we look at a sub-sample only consisting of one day per month, namely 17 days before option expiry (which is the day after

[7]Recall that funding costs and margin requirements can be highly agent specific and may be more severe than our most conservative estimates for individual investors. Also, some early exercise decisions may be driven by corporate actions. For instance, studying the days with option exercises by market makers that cannot be explained by the most conservative boundary (the 1.6% cases remaining after the model has explained 98.4%), we find that the largest such early exercise happened on January 7, 2004 for options written on Univision Communications Inc. This was the day after Univision announced a plan of offering and repurchasing 15.8 million shares. The second largest of the unexplained early exercise days was Monday July 26, 2010 for options written on Americredit. Four days earlier, on July 22, 2010, a press release announced that General Motors was to acquire Americredit.

the third Friday in every month). The sub-sample analysis is sufficient to obtain statistically significant results and we have confirmed that our model independent results hold up in the full sample (i.e., by regressing the propensity to exercise on characteristics such as short-sale costs). For each of the selected 80 dates, we compute the optimal exercise boundary for each option by solving the PDE (7)–(8) that takes frictions into account.

If we run a pooled logit (or probit) regression using all options on all of the selected dates, then we get highly significant results consistent with our model (not reported). However, such standard errors would be heavily downward biased since investors usually exercise many options simultaneously, generating a strong correlation across options on a given date. To address this correlation issue, we proceed as follows. First, we run a logit (or probit) regression for each subsample of three dates (the last subsample has only two dates). This generates an estimated vector of parameters, $\hat{\theta}_s$, for each subsample $s$. Second, we estimate the full-sample parameters and their standard errors based on the insight of Fama and MacBeth (1973) that each parameter can be viewed as sampled from parameters' distribution. In particular, we estimate the full-sample parameters as the sample average, $\hat{\theta} = 1/27 \sum_s \hat{\theta}_s$, and the standard errors based on the sample standard deviation corrected for possible auto-correlation (Newey-West correction with automatic lag selection using a Bartlett-kernel, Newey and West (1987, 1994)). This estimation method is relatively immune to cross-sectional correlation in option exercises and assumes that any time-series correlation is captured by the Newey-West correction.

Table VI reports the results, where Panel A is the logit regressions and Panel B is probit. The first regression specification simply considers how the propensity to exercise depends on the indicator that the stock price is above the exercise boundary ($S > B$). We see that the estimated coefficient for $S > B$ is positive, consistent with our model, and the effect is highly statistically significant. The estimated probability of exercise on a day with $S > B$ is $1/(1 + \exp(8.54 - 4.09)) = 1.2\%$, while estimated probability on other days is $1/(1 + \exp(8.54 - 0)) = 0.02\%$. Cumulated over 20 trading days, these probabilities correspond to aggregate exercise probabilities of $1-(1-1.2\%)^{20} = 20.7\%$ in the case of $S > B$

and 0.4% otherwise. We see that the exercise behavior is very different depending on whether the stock price is above or below the exercise boundary. We note however, as mentioned in the introduction, that the cumulated exercise probabilities should be interpreted with caution due to "burn-out" effects similar to those in the mortgage markets during refinancing waves. The estimated probabilities are very similar in the probit regressions in Panel B.

The second specification shows a similar result, but where we allow the exercise probability to increase in the distance between the stock price and the boundary. The third specification shows that having a bid price for the option below the intrinsic value is also a significant predictor of early option exercise, which is also consistent with our model (and a basic trade off between selling the option vs. exercising it). Of course, the option bid price is an endogenous variable, so this regression does not address the deeper question of *why* option prices can be so low that early exercise can make sense. The first specification based on our model is not subject to the same issue since the exercise boundary is computed based on the observed financial frictions. Therefore, if all investors face these financial frictions, the model implies that all agents should exercise early and the equilibrium option price should fall to the intrinsic value.

The fourth specification includes all these variables jointly and the fifth also includes a number of control variables. The indicator that $S > B$ is positive and significant in all specifications. Several of the other variables are also significant, suggesting that the precise probability of exercise is a complicated function of the observable data, perhaps because there are differences across investors in terms of the frictions that they face and the frequency with which they observe the markets.

# V. Empirical Results: Never Convert a Convertible?

We next consider how early conversions of convertible bonds are related to financial frictions. Table VII reports our results. We see that early conversions are more frequent among companies with large short-sale costs for the equity, consistent with the theory.

The table breaks down the conversions by the moneyness of the convertible bonds. We consider a convertible bond to be out-of-the-money if the price of the underlying stock is less than conversion price, in-the-money if stock price is up to 25% above conversion price, and deep-in-the-money if stock price is more than 25% above conversion price. As expected, we see that conversions are concentrated among in-the-money and deep-in-the-money convertibles. We see a few conversions of out-of-the-money bonds, which relates to the definition of moneyness discussed in Section III.B.

More importantly, we see that conversion rates increase monotonically in short-sale costs both for in-the-money and deep-in-the-money convertibles, providing further evidence consistent with the theory. However, we note that the difference across groups is not statistically significant due to the small and noisy dataset.

# VI. Conclusion: Never Say Never Again

A classic rule in financial economics states that, except just before expiration or dividend payments, one should never exercise a call option and never convert a convertible bond. We show that this rule breaks down — theoretically and empirically — when financial frictions are introduced, just as frictions break the Modigliani-Miller Theorem, the Law of One Price, and other classic rules in financial economics.

Our theory shows that early exercise of options can be rational in light of financial frictions and, indeed, we would expect early exercise to happen in equilibrium. Consistent with our theory, the empirical propensity to exercise equity options is increasing in the short-sale costs, transaction costs, and moneyness, and decreasing in the time to expiration. We find that options are exercised by customers of brokers, market makers, and firm proprietary traders and, for each group, exercises are more prevalent when the financial frictions are more severe. Our model further implies that it can be optimal to convert a convertible bond. We document a number of early conversions of convertible bonds, especially among stocks with high short-sale costs.

# References

Agarwal, Vikas, Narayan Y. Naik, and Y.C. Loon, 2011, Risk and return in convertible arbitrage: Evidence from the convertible bond market, *Journal of Empirical Finance* 18, 175–194.

Avellaneda, M., and M. Lipkin, 2009, A dynamic model for hard-to-borrow stocks, *Risk Magazine* June, 92–97.

Barraclough, K., and R. Whaley, 2012, Early exercise of put options on stocks, *The Journal of Finance* 67, 1423–1456.

Bates, D. S., 2000, Post 87 crash fears in the s&p 500 futures options market, *Journal of Econometrics* 94, 181–238.

Bates, D. S., 2003, Empirical option pricing: A retrospection, *Journal of Econometrics* 116, 387–404.

Battalio, Robert, Stephen Figlewski, and Robert Neal, 2014, Exercise boundary violations in american-style options, *working paper, Mendoza School of Business.*

Bergman, Y., 1995, Option pricing with differential interest rates, *Review of Financial Studies* 53, 475–500.

Black, F., and M. S. Scholes, 1973, The pricing of options and corporate liabilities, *The Journal of Political Economy* 81, 637–654.

Bollen, N. P., and R. E. Whaley, 2004, Does net buying pressure affect the shape of implied volatility functions?, *The Journal of Finance* 59, 711–753.

Brennan, M.J., and E.S. Schwartz, 1977, Convertible bonds: valuation and optimal strategies for call and conversion, *The Journal of Finance* 32, 1699–1715.

Brenner, M., R. Eldor, and S. Hauser, 2001, The price of options illiquidity, *The Journal of Finance* 56, 789–805.

Christoffersen, P., R. Goyenko, K. Jacobs, and M. Karoui, 2011, Illiquidity premia in the equity options market, *Working Paper, University of Toronto.*

Constantinides, G. M., J. C. Jackwerth, and S. Perrakis, 2009, Mispricing of s&p 500 index options, *Review of Financial Studies* 22, 1247–1277.

Dewynne, J. N., S. D. Howison, I. Rupf, and P. Wilmott, 1993, Some mathematical results in the pricing of american options, *European Journal of Applied Mathematics* 4, 381–398.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen, 2002, Securities lending, shorting, and pricing, *Journal of Financial Economics* 66, 307–339.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Figlewski, S., 1989, Options arbitrage in imperfect markets, *The Journal of Finance* 44, 1289–1311.

Frazzini, A., and L.H. Pedersen, 2011, Embedded leverage, *Working paper, New York University.*

Garleanu, Nicolae, and Lasse Heje Pedersen, 2011, Margin-based asset pricing and deviations from the law of one price, *Review of Financial Studies* 24, 1980–2022.

Garleanu, Nicolae, Lasse Heje Pedersen, and Allen Poteshman, 2009, Demand-based option pricing, *Review of Financial Studies* 22, 4259–4299.

Ingersoll, J. E., 1977a, A contingent-claims valuation of convertible securities, *Journal of Financial Economics* 4, 289–321.

Ingersoll, J. E., 1977b, An examination of corporate call policy on convertible securities, *Journal of Finance* 32, 463–478.

Jackwerth, J., 2000, Recovering risk aversion from option prices and realized returns, *Review of Financial Studies* 13, 433–451.

Karatzas, Ioannis, and Steven G Kou, 1998, Hedging american contingent claims with constrained portfolios, *Finance and Stochastics* 2, 215–258.

Leipold, M., and L. Su, 2012, Collateral smile, *Working paper, University of Zurich.*

Longstaff, F. A., 1995, Option pricing and the martingale restriction, *Review of Financial Studies* 8, 1091–1124.

Merton, R., 1973, Theory of rational option pricing, *The Bell Journal of Economics and Management Science* 4, 141–183.

Mitchell, Mark, Lasse Heje Pedersen, and Todd Pulvino, 2007, Slow moving capital, *American Economic Review (Papers & Proceedings)* 97, 215–220.

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

Newey, Whitney K., and Kenneth D. West, 1994, Automatic lag selection in covariance matrix estimation, *Review of Economic Studies* 61, 631–653.

Ni, S.X., 2009, Stock option returns: A puzzle, *Working paper, Hong Kong University of Science and Technology.*

Ofek, E., M. Richardson, and Robert F. Whitelaw, 2003, Limited arbitrage and short sales restrictions: Evidence from the options markets, *Journal of Financial Economics* 74, 305–342.

Piterbarg, Vladimir, 2010, Funding beyond discounting: collateral agreements and derivatives pricing, *risk-magazine.net* pp. 97–102.

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

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

Santa-Clara, P., and A. Saretto, 2009, Option strategies: Good deals and margin calls, *Journal of Financial Markets* 12, 391–417.

# A. Appendix: Comparative Statics and Proofs

## Comparative Statics

Proposition 4 provides comparative statics for an agent who is short the stock. Here we provide the complimentary case of an agent who is long stock.

**Proposition 5 (The Optimal Exercise Boundary for an agent long stock)**
*Suppose that agent $i$ is long an option which he cannot sell above its intrinsic value.*

*If agent $i$ is long stock and either his funding cost or the stock lending fee is strictly positive ($\psi^i > 0$ or $l^i > 0$) then there exists a finite optimal exercise boundary, $B_{T-t} < \infty$, which solves* (7)–(8)*. Hence, the agent optimally exercises as soon as the stock price is above the boundary, $S_t > B_{T-t}$. The following comparative statics hold for the limit boundary for long-dated options $B^* := \lim_{\tau \to \infty} B_\tau$: The exercise boundary*

- *decreases in lending fee ($\frac{\partial B^*}{\partial l^i} < 0$)*
- *decreases in option margin requirements ($\frac{\partial B^*}{\partial m^{i,C}} \leq 0$, $\frac{\partial B^*}{\partial m^{i,C}} < 0$ if $\psi^i > 0$)*
- *increases in marginal stock margin requirement ($\frac{\partial B^*}{\partial m^{i,S}} \geq 0$, $\frac{\partial B^*}{\partial m^{i,S}} > 0$ if $\psi^i > 0$)*
- *increases in the volatility ($\frac{\partial B^*}{\partial \sigma} > 0$).*

The rest of the appendix provides proofs of our propositions.

## Proof of Proposition 1

If the agent wants cash and exercises to immediately sell the stock in the market, the net proceeds would be $S_0(1 - \lambda^{i,S}) - X$. If the investor instead places $\frac{X}{1+r^f}$ in the risk-free asset and sells short the stock the cash flow would now be $S_0(1 - \lambda^{i,S}) - \frac{X}{1+r^f}$ (as there are no funding or short-sale costs). In the next period the option can be exercised paid by the money from the risk-free asset and the received stock can be used to close the short position. This has a net payment of zero. Thereby the cash flow of the strategy dominates the one created by exercising and selling the stock (since $r^f > 0$).

If the investor wants the exposure to the stock and exercises to get the stock, the cash flow is $-X$. The alternative dominating strategy is to place $\frac{X}{1+r^f}$ in the risk-free asset and wait and exercise later. This costs less today (as $r^f > 0$), and the stock purchase can be made cash flow neutral at time 1 by the money from the risk-free asset. No value is forgone by not owning the stock in the meantime as there is no lending fee, $l^i = 0$, and no funding costs are incurred.

Consider *(ii)* where the sale revenue of the option is above the intrinsic value for agent $i$, $C_0(1-\lambda^{i,C}) > S_0 - X$. If agent $i$ wants cash, exercising the option and selling the stock obviously gives less than selling the option. If agent $i$ wants stock, selling the option and buying the stock has a net cost of $S_0 - C_0(1-\lambda^{i,C})$ which is less than what it costs to gain the stock through exercise, $X$. Hence, no matter if agent $i$ wants stock or cash, early exercise is dominated.

A sufficient condition for $C_0(1-\lambda^{i,C}) > S_0 - X$ is that agent $i$ faces no option transaction costs, $\lambda^{i,C} = 0$, and the existence of another type of agents $j$ with zero short-sale costs, funding costs, and stock transaction costs. We have just shown above that for an investor without short-sale costs, funding costs or stock transaction costs the value of the option will be at least $S_0 - \frac{X}{1+r^f}$. If such unconstrained agents $j$ exist, the option price must be at least $S_0 - \frac{X}{1+r^f}$ to avoid arbitrage. If $\lambda^{i,C} = 0$ we get: $C_0(1-\lambda^{i,C}) = C_0 \geq S_0 - \frac{X}{1+r^f} > S_0 - X$.

## Proof of Proposition 2

This proof shows that there exists a strategy involving early exercise that is not dominated by any strategy not involving early exercise if frictions are large enough.

We introduce the following notation for the portfolio held by some agent: The number of stocks, $\alpha_0 \in \mathbb{R}$, the amount invested in the risk-free asset, $\beta_0 \in \mathbb{R}$, and the number of options, $\gamma_0 \in \mathbb{R}$.

Let an agent $i$ who is long option, $\gamma_0 > 0$, want as much cash as possible at time 0 without any negative cash flow in any possible state at time 1. One way to do this is to exercise the option position early and close the positions in stock and risk-free asset. If stocks needs to be sold to close the position, stock transaction costs apply. However, in the event that more can be earned through lending out the stock and borrow the present value of the future earned lending fee than through selling the stock, this is done instead. If stock must be acquired this can be done either by buying stock or buying options and exercise immediately. The cheapest way is chosen. The cash-flow at time 0 of this strategy can be described by $K$:

$$K := \begin{cases} (\alpha_0 + \gamma_0)S_0(1 - \lambda^S) + \beta_0 - \gamma_0 X & \text{if } \alpha_0 + \gamma_0 \geq 0 \text{ and } 1 - \lambda^S \geq \frac{l^i}{1+r^f} \\ (\alpha_0 + \gamma_0)\frac{S_0 l^i}{1+r^f} + \beta_0 - \gamma_0 X & \text{if } \alpha_0 + \gamma_0 \geq 0 \text{ and } 1 - \lambda^S < \frac{l^i}{1+r^f} \\ (\alpha_0 + \gamma_0)S_0 + \beta_0 - \gamma_0 X & \text{if } \alpha_0 + \gamma_0 < 0 \text{ and } S_0 \leq C_0 + X \\ (\alpha_0 + \gamma_0)(C_0 + X) + \beta_0 - \gamma_0 X & \text{if } \alpha_0 + \gamma_0 < 0 \text{ and } S_0 > C_0 + X \end{cases} \tag{12}$$

We refer to this strategy as *the liquidating early exercise strategy.*

Any strategy without early exercise is identified by the allocation of wealth between the three assets.[8] This allocation can be denoted by $(\alpha_1, \beta_1, \gamma_1)$ for the number of stocks, amount invested in risk-free asset, and number of options respectively. The cash-flow at time 0 for an arbitrary strategy not involving early exercise can be described by the function $J_1$ where:

$$\begin{aligned} J_1(\alpha_1, \beta_1, \gamma_1) :=& (\alpha_0 - \alpha_1)(S_0 - \mathbb{1}_{(\alpha_0 - \alpha_1 > 0)} S_0 \lambda^S) + (\beta_0 - \beta_1) \\ &+ (\gamma_0 - \gamma_1)(C_0 - \mathbb{1}_{(\gamma_0 - \gamma_1 > 0)} \lambda^{i,C} C_0) - F^i(\alpha_1 S_0, \gamma_1 C_0) \end{aligned}$$

We want to show that there for agent $i$ exists a strategy involving early exercise that is not dominated by any strategy not involving early exercise if frictions are severe enough. Our candidate for such a non-dominated strategy involving early exercise is the liquidating early exercise strategy.

In order for a strategy to dominate the liquidating early exercise strategy the strategy must have non-negative net-liabilities at time 1 in all possible states.

[8] Admittedly, the same portfolio can be obtained through several ways of trading. E.g. the number of stocks held can be reduced by 1 either by selling one stock or selling two stocks and buying one back immediately, corresponding to money burning when stock transaction costs are positive. For our purpose it is sufficient to consider strategies with the cheapest way to obtain the portfolio. If such strategies cannot dominate early exercise neither can strategies where money is given up for nothing.

Specifically this means that for every stock agent $i$ is short agent $i$ must also be long at least 1 option and have no less than $\frac{X+S_0L^i}{1+r^f}$ in the risk-free asset. As the stock has unlimited upside the option is needed to secure no negative payoff, and the cash is needed to pay the possible future exercise of the option and the short-sale fee of the stock. Moreover, for every option agent $i$ is short, $i$ must be at least long 1 stock, and not use the risk-free asset to borrow more than $\frac{S_0l^i}{1+r^f}$. The stock is needed to be able to honor a possible future exercise of the option in all states. Since the stock could also turn out to be worthless at time 1 there cannot be borrowed more than $\frac{S_0l^i}{1+r^f}$, the present value of the only secure income raised from lending out the stock in the period. Only strategies where this is fulfilled can possible dominate the liquidating early exercise strategy. The maximal cash-flow that can be obtained at time $t=0$ without early exercise and without net-liabilities at time $t=1$ is then the solution to this problem:

$$\max_{\alpha_1,\beta_1,\gamma_1} J_1(\alpha_1,\beta_1,\gamma_1) \tag{13}$$

$$\text{s.t. } \alpha_1+\gamma_1 \geq 0 \tag{14}$$

$$\beta_1 + \mathbb{1}_{(\alpha_1<0)}\alpha_1\frac{X+S_0L^i}{1+r^f} + \mathbb{1}_{(\alpha_1>0)}\alpha_1\frac{S_0l^i}{1+r^f} \geq 0 \tag{15}$$

All strategies are either non-dominated or dominated by a non-dominated strategy. Since domination is a transitive property this implies that it is sufficient to show that there exists no non-dominated strategy not involving early exercise that dominates the liquidating early exercise strategy.

If there exists a non-dominated strategy involving selling the option (i.e. $\gamma_1 < \gamma_0$) and $C_0(1-\lambda^{i,C}) \leq S(1-\lambda^{i,S})-X$ then there is a corresponding non-dominated strategy involving early exercise; namely the one with the same trades except that instead of selling the options the same number of options are exercised early and the obtained stock sold immediately. And hence the statement in the proposition is fulfilled.

This means that it is sufficient to show that there are no strategies that does not involve selling the option or exercising it early that dominate the liquidating early exercise strategy.

Not selling the option corresponds to $\gamma_1 \geq \gamma_0$. Combining this with (14) we get the inequality $\gamma_1 \geq \max(-\alpha_1, \gamma_0)$. Note that $J_1$ is decreasing in $\gamma_1$ for $\gamma_1 \geq 0$ meaning that this inequality must bind in optimum. Likewise, $J_1$ is decreasing in $\beta_1$ which means that (15) must bind at the optimum. We can now substitute both $\beta_1$ and $\gamma_1$ and reformulate the problem as:

$$\max_{\alpha_1} J_2(\alpha_1) \tag{16}$$

where we define the function $J_2$ by:

$$\begin{aligned}
J_2(\alpha_1) :=& J_1(\alpha_1, -\mathbb{1}_{(\alpha_1<0)}\alpha_1 \frac{X + S_0 L^i}{1+r^f} - \mathbb{1}_{(\alpha_1>0)}\alpha_1 \frac{S_0 l^i}{1+r^f}, \max(-\alpha_1, \gamma_0)) \\
=& (\alpha_0 - \alpha_1)(S_0 - \mathbb{1}_{(\alpha_0-\alpha_1>0)} S_0 \lambda^S) + \beta_0 + \mathbb{1}_{(\alpha_1<0)}\alpha_1 \frac{X + S_0 L^i}{1+r^f} \\
&+ \mathbb{1}_{(\alpha_1>0)}\alpha_1 \frac{S_0 l^i}{1+r^f} + (\gamma_0 - \max(-\alpha_1, \gamma_0))C_0 - F^i(\alpha_1 S_0, -\max(-\alpha_1, \gamma_0)C_0)
\end{aligned} \tag{17}$$

First, we want to show part *a.* in the proposition. The proposition claims that when $S_0(1-\lambda^{i,S}) > X$ and $C_0(1-\lambda^{i,C}) \leq S_0(1-\lambda^{i,S}) - X$ then early exercise is not dominated for $i$ if the short-sale cost, $L^i$, is large enough. Specifically we show that $L^i > \frac{Xr^f}{S_0}$ is sufficient.

Given this we show that the optimal cash-flow solving (16) is smaller than the cash flow from the liquidating early exercise strategy (12), regardless of funding costs. Since funding costs, $F^i$, are non-negative it is sufficient to show that the maximum of $J_2$ is smaller than (12) for $F^i = 0$. So in the following we consider the case where $F^i = 0$. Then $J_2$ is piece-wise linear in $\alpha_1$ with a kinks at $\{-\gamma_0, \alpha_0, 0\}$. So a global maximum for the $J_2$ will either be at $-\gamma_0$, $\alpha_0$, $0$ or be non-existing. The maximum is non-existing if $J_2$ goes to infinity for $\alpha_1$ going to either $\infty$ or $-\infty$. We check the last first:

$$\begin{aligned}
\lim_{\alpha_1 \to -\infty} J_2(\alpha_1) =& \lim_{\alpha_1 \to -\infty} \left[ (\alpha_0 - \alpha_1) S_0 (1 - \lambda^{i,S}) + \beta_0 + \alpha_1 \frac{X + S_0 L^i}{1+r^f} + (\gamma_0 + \alpha_1) C_0 \right] \\
=& \alpha_0 S_0 (1 - \lambda^{i,S}) + \beta_0 + \gamma_0 C_0 + \lim_{\alpha_1 \to -\infty} \left[ \alpha_1 [-S_0(1 - \lambda^{i,S}) + \frac{X + S_0 L^i}{1+r^f} + C_0] \right]
\end{aligned} \tag{18}$$

It must hold that $C_0 + X \geq S_0(1 - \lambda^{i,S})$, otherwise there would be an arbitrage strategy by buying options, exercise immediately and sell the stock. Given this and that $L^i > \frac{Xr^f}{S_0}$ the entire expression (18) is diverging to $-\infty$. We next check $\alpha_1 \to \infty$:

$$\begin{aligned}\lim_{\alpha_1 \to \infty} J_2(\alpha_1) =& \lim_{\alpha_1 \to \infty} \left[ (\alpha_0 - \alpha_1)S_0 + \beta_0 + \alpha_1 \frac{S_0 l^i}{1 + r^f} \right] \\ =& \alpha_0 S_0 + \beta_0 + \lim_{\alpha_1 \to \infty} \left[ \alpha_1 (-S_0 + \frac{S_0 l^i}{1 + r^f}) \right]\end{aligned}$$

No-arbitrage implies that $S_0 > \frac{S_0 l^i}{1+r^f}$. Otherwise an arbitrage gain could be made through buying stocks and lending them out. This in turn implies that $J_2$ does not go to infinity as $\alpha_1$ goes to infinity.

We now evaluate the expression in $\alpha_1 = \alpha_0 \leq -\gamma_0$:

$$J_2(\alpha_0) = \beta_0 + \alpha_0 \frac{X + S_0 L^i}{1 + r^f} + (\gamma_0 + \alpha_0)C_0 \tag{19}$$

This is strictly smaller than the cash flow from early exercise in (12) (using that $L^i > \bar{L} = \frac{Xr^f}{S_0}$ and $\alpha_0 + \gamma_0 \leq 0$). Next, evaluate where $\alpha_1 = \alpha_0 \in (-\gamma_0, 0]$:

$$J_2(\alpha_0) = \beta_0 + \alpha_0 \frac{X + S_0 L^i}{1 + r^f} \tag{20}$$

This is strictly smaller than (12) (using that $\alpha_0 + \gamma_0 > 0$ and $L^i > \bar{L} = \frac{Xr^f}{S_0}$). Evaluating $J_2$ where $\alpha_1 = \alpha_0 > 0$:

$$J_2(\alpha_0) = \beta_0 + \alpha_0 \frac{S_0 l^i}{1 + r^f}$$

which is strictly smaller than (12).

Next, we evaluate where $\alpha_1 = 0$:

$$J_2(0) = \alpha_0(S_0 - \mathbb{1}_{(\alpha_0 > 0)} S_0 \lambda^S) + \beta_0 \tag{21}$$

This is also smaller than the cash flow at time $t = 0$ from the liquidating early exercise strategy described by (12).

The last evaluation of the expression is where $\alpha_1 = -\gamma_0$:

$$J_2(-\gamma_0) = (\alpha_0 + \gamma_0)(S_0 - \mathbb{1}_{(\alpha_0+\gamma_0>0)} S_0 \lambda^{i,S}) + \beta_0 - \gamma_0 \frac{X + S_0 L^i}{1 + r^f} \tag{22}$$

Again, this is strictly smaller than (12) since $L^i > \bar{L} = \frac{X r^f}{S_0}$. Hence it has been proven that for sufficiently high short-sale costs early exercise is not dominated regardless of funding costs if $C_0(1 - \lambda^{i,C}) \leq S_0(1 - \lambda^{i,S}) - X$ and $S_0(1 - \lambda^{i,S}) - X > 0$. Now, we want to show that a similar result holds for sufficiently high funding costs instead of short-sale costs.

The proposition states that if $S_0(1 - \lambda^{i,S}) > X$ and $C_0(1 - \lambda^{i,C}) \leq S_0(1 - \lambda^{i,S}) - X$ then early exercise is not dominated when the funding costs, $F^i$, is large enough. By $F^i$ being large enough we specifically mean that $F^i(x, y) \geq \bar{F}(|x| + |y|)$ where $\bar{F} > \frac{r^f X}{(1+r^f)(S_0+C_0)}$. When this is the case $J_3$ will be smaller than $J_2$ where:

$$\begin{aligned} J_3(\alpha_1) :=& (\alpha_0 - \alpha_1)(S_0 - \mathbb{1}_{(\alpha_0-\alpha_1>0)} S_0 \lambda^S) + \beta_0 + \mathbb{1}_{(\alpha_1<0)} \alpha_1 \frac{X + S_0 L^i}{1 + r^f} + \mathbb{1}_{(\alpha_1>0)} \alpha_1 \frac{S_0 l^i}{1 + r^f} \\ &+ (\gamma_0 - \max(-\alpha_1, \gamma_0)) C_0 - \bar{F}(|\alpha_1 S_0| + \max(-\alpha_1, \gamma_0) C_0) \end{aligned} \tag{23}$$

So it is sufficient to show that $J_3$ is smaller than (12) given that $\bar{F} > \frac{r^f X}{(1+r^f)(S_0+C_0)}$. In the following we assume that this condition is fulfilled. $J_3$ is piecewise linear with kinks in $\{-\gamma_0, \alpha_0, 0\}$. In order for a global maximum to exist for $J_3$ it must not go to $\infty$ as $\alpha_1$ goes to either $\infty$ or $-\infty$. We check this:

$$\begin{aligned} &\lim_{\alpha_1 \to -\infty} J_3(\alpha_1) \\ =& \lim_{\alpha_1 \to -\infty} (\alpha_0 - \alpha_1) S_0 (1 - \lambda^S) + \beta_0 + \alpha_1 \frac{X + S_0 L^i}{1 + r^f} + (\gamma_0 + \alpha_1) C_0 + \alpha_1 \bar{F}(S_0 + C_0) \\ =& \alpha_0 + \beta_0 + \gamma_0 C_0 + \lim_{\alpha_1 \to -\infty} \left[ \alpha_1 [-S_0(1 - \lambda^S) + \frac{X + S_0 L^i}{1 + r^f} + C_0 + \bar{F}(S_0 + C_0)] \right] \end{aligned}$$

The expression in the squared brackets is positive since no-arbitrage implies that $S(1-\lambda^S) \leq C + X$. If this inequality was not fulfilled an arbitrage strategy would be to buy options, exercise them immediately and sell the obtained stocks. We conclude that $J_3$ does not go to $\infty$ for $\alpha_1$ going to $-\infty$.

$$\begin{aligned}
&\lim_{\alpha_1\to\infty} J_3(\alpha_1) \\
=&\lim_{\alpha_1\to\infty}(\alpha_0-\alpha_1)S_0+\beta_0+\alpha_1\frac{S_0 l^i}{1+r^f}-\bar{F}(\alpha_1 S_0+\gamma_0 C_0) \\
=&\alpha_0+\beta_0-\bar{F}\gamma_0 C_0+\lim_{\alpha_1\to\infty}\left[\alpha_1[-S_0+\frac{S_0 l^i}{1+r^f}-\bar{F}S_0]\right]
\end{aligned}$$

Since $S_0 > \frac{S_0 l^i}{1+r^f}$ it follows that this does not go to infinity. Next, we evaluate $J_3$ where $\alpha_1 = -\gamma_0$:

$$J_3(-\gamma_0) = (\alpha_0+\gamma_0)(S_0-\mathbb{1}_{(\alpha_0+\gamma_0>0)}S_0\lambda^S)+\beta_0-\gamma_0\frac{X+S_0L^i}{1+r^f}-\bar{F}(\gamma_0 S_0+\gamma_0 C_0)$$

This is strictly smaller than the payoff from early exercise in (12) since $\bar{F} > \frac{r^f X}{(1+r^f)(S_0+C_0)}$. For $\alpha_1 = 0$ we get:

$$J_3(0) = \alpha_0(S_0-\mathbb{1}_{(\alpha_0>0)}S_0\lambda^S)+\beta_0-\bar{F}\gamma_0 C_0$$

which is strictly smaller than (12) since $S(1-\lambda^{i,S})-X>0$. Next, evaluate where $\alpha_1 = \alpha_0 \leq -\gamma_0$:

$$J_3(\alpha_0) = \beta_0+\alpha_0\frac{X+S_0L^i}{1+r^f}+(\gamma_0+\alpha_0)C_0+\alpha_0\bar{F}(S_0+C_0)$$

which is strictly smaller than (12). Evaluate where $\alpha_1 = \alpha_0 \in (-\gamma_0, 0]$:

$$\begin{aligned}
J_3(\alpha_0) &= \beta_0+\alpha_0\frac{X+S_0L^i}{1+r^f}-\bar{F}(-\alpha_0 S_0+\gamma_0 C_0) \\
&= \beta_0+\alpha_0\frac{X+S_0L^i}{1+r^f}+\alpha_0\bar{F}(S_0+C_0)-\bar{F}C_0(\alpha_0+\gamma_0)
\end{aligned}$$

which is strictly smaller than (12). Finally, we evaluate where $\alpha_1 = \alpha_0 > 0$:

$$J_3(\alpha_0) = \beta_0 + \alpha_0 \frac{S_0 l^i}{1 + r^f} - \bar{F}(\alpha_0 S_0 + \gamma_0 C_0)$$

which is also strictly smaller than (12). Hence, for funding cost large enough, i.e. $F^i(x, y) \geq \bar{F}(|x| + |y|)$ where $\bar{F} > \frac{r^f X}{(1+r^f)(S_0+C_0)}$, early exercise is not dominated (regardless of short-sale costs).

## Proof of Proposition 3

We want to show that it in fact is an equilibrium where the option is exercised early at time 1 if the stock prospects evolve positively in the first period. Let $i_j^S$ and $i_j^C$ denote the number of respectively stocks and options held by types $i$ in state $j$. Likewise, let $i_j^R$ be the cash amount invested in the risk-free asset by $i$ in state $j$. The price of the stock and option in state $j$ we denote $S_j$ and $C_j$ respectively. The net supply of stock is 1 and the net supply of option is 0. In equilibrium supply and demand meet meaning that $A_j^S + B_j^S = 1$ and $A_j^C + B_j^C = 0$ in all states $j$. Then the equilibrium is such that the option is exercised early in the $u$-state and given by

$$\begin{aligned}
S_d &= \frac{p_1^B S_{du} + (1 - p_1^B) S_{dd}}{1 + r^f} = \frac{0.46 \cdot 99.5 + 0.54 \cdot 69.097}{1.001} = 83.000 \\
B_d^S &= 1 \\
A_d^S &= C_d = A_d^C = B_d^C = A_d^R = 0 \\
C_d &= 0 \\
B_d^R &= B_0^S S_d + B_0^C C_d + B_0^R (1 + r^f) - B_d^S S_d - B_d^C C_d \\
&= 0 + 979.727 \cdot 1.001 - 1 \cdot 83.000 - 0 = 897.707 \\
S_u &= \frac{p_1^B S_{uu} + (1 - p_1^B) S_{du}}{1 + r^f} = \frac{0.46 \cdot 143.280 + 0.54 \cdot 99.5}{1.001} = 119.519 \\
A_u^S &= 0 \\
B_u^S &= 1
\end{aligned}$$

$$
\begin{aligned}
C_u &= \frac{p_1^B(S_{uu} - X)}{1 + r^f} = \frac{0.46(143.280 - 100)}{1.001} = 19.889 \\
A_u^C &= \frac{-(A_0^S S_u + A_0^C(S_u - X) + A_0^R(1 + r^f))(1 + r^f)}{S_{uu} - X - C_u(1 - \lambda^C)(1 + r^f)} \\
&= \frac{-(1 \cdot 119.519 + 20.482 \cdot 19.519 - 82.917 \cdot 1.001)1.001}{143.280 - 100 - 19.889 \cdot 0.98 \cdot 1.001} = -18.374 \\
B_u^C &= -A_U^C = 18.374 \\
A_u^R &= (A_0^S S_u + A_0^C(S_u - X) + A_0^R(1 + r^f)) - A_u^C \cdot C_u(1 - \lambda^C) \\
&= (1 \cdot 119.519 + 20.482 \cdot 19.519 - 82.917 \cdot 1.001) + 18.374 \cdot 19.889 \cdot 0.98 = 794.450 \\
B_u^R &= B_0^S S_u + B_0^C(S_u - X) + B_0^R(1 + r^f) - B_u^S S_u - B_u^C C_u \\
&= 0 - 20.482 \cdot 19.519 + 979.727 \cdot 1.001 - 119.519 - 18.374 \cdot 19.889 = 95.948 \\
S_0 &= \frac{\frac{S_u - S_d}{S_u - X - C_d}(C_0(1 + r^f) - (S_u - X)) + S_u}{1 + r^f} \\
&= \frac{\frac{119.519 - 83.000}{19.519}(8.954 \cdot 1.001 - 19.519) + 119.519}{1.001} = 99.669 \\
A_0^S &= 1 \\
B_0^S &= 0 \\
C_0 &= \frac{p_0^B(S_u - X)}{(1 + r^f)(1 - \lambda^C)} = \frac{0.45(119.519 - 100)}{1.001 \cdot 0.98} = 8.954 \\
A_0^C &= \frac{800 - A_0^S S_0 - A_0^R}{C_0} = \frac{200 - 1 \cdot 99.522 + 82.917}{8.954} = 20.482 \\
B_0^C &= -A_0^C = -20.482 \\
A_0^R &= \frac{-S_d}{1 + r^f} = \frac{-83.000}{1.001} = -82.917 \\
B_0^R &= 800 - B_0^S S_0 - B_0^C C_0(1 - \lambda^C) = 800 - 0 + 20.482 \cdot 8.954 \cdot 0.98 = 979.727
\end{aligned}
$$

Now, we verify the equilibrium in each of the three states. Remember that $A$ have a cash endowment of 200 and $B$ have 800, and observe how in all states at time 0 and 1 $B$'s no-default condition is not binding. Hence, in all states the expected utility gained from investing 1 dollar in the risk-free asset must equal the expected utility gained from investing

1 dollar in the stock and 1 dollar in the option, respectively. First, we verify the equilibrium in the 0-state: The reserve price of the stock for $B$ as potential buyer, $\tilde{S}_0^B$, is where the expected utility gain from buying stock is at least as big as from investing the cash in the risk-free asset:

$$\frac{p_0^B S_u + (1 - p_0^B) S_d}{\tilde{S}_0^B} = 1 + r^f$$

The reserve then is

$$\tilde{S_0^B} = \frac{p_0^B S_u + (1 - p_0^B) S_d}{1 + r^f} = 99.334$$

This means that $B$ is not interesting in buying as $S_0 = 99.669$. Similarly, the reserve price for $B$ as seller can be derived as

$$\frac{p_0^B S_u + (1 - p_0^B) S_d + S_0 L}{1 + r^f} = 99.732$$

where short-sale costs must be taken into account. This means that $B$ will not sell either. What is then the reserve price for $B$ as an option seller?

$$\frac{p_0^B (S_u - X)}{(1 + r^f)(1 - \lambda^C)} = 8.954$$

This means that exactly at the equilibrium price $B$ is indifferent between selling options and not selling options. In equilibrium some options are sold to $A$ at this price. Turn to $A$ at state 0: Here, the no-default constraint is binding, as the portfolio value in state $d$ is zero per construction:

$$A_0^S S_d + A_0^C C_d + A_0^R (1 + r^f) = 0$$

Would $A$ find it attractive to buy less stock and instead investing in the risk-free asset? The reserve price for the stock for $A$ is

$$\frac{p_0^A S_u + (1 - p_0^A) S_d}{1 + r^f} = 99.699$$

As this is larger than $S_0 = 99.669$ $A$ will not find it attractive to buy less stock and instead invest the risk-free asset. How about the option? Here the reserve price is

$$\frac{p_0^A (S_u - X) S_d}{1 + r^f} = 8.970$$

Thus, it is not attractive for $A$ to buy less option and invest in the risk-free asset, as this is larger than $C_0 = 8.954$. But is it attractive to buy fewer options and instead buy more stock? Or vice versa? The option price is per construction such that nothing can be gained from either of this. I.e. the option price equals the cost of replicating the option using stock and risk-free asset. This follows the classic binomial model option price given the stock price, foreseeing that the option is exercised early in the $u$-state. The replication argument is that if one at time $t = 0$ buys $\alpha$ stocks and invests $\beta$ in the risk-free asset then this replicates the value of the option at time $t = 1$ when $\alpha$ and $\beta$ solve the following equations:

$$\alpha S_u + \beta(1 + r^f) = S_u - X$$
$$\alpha S_d + \beta(1 + r^f) = C_d$$

The unique solution to this is $(\alpha, \beta) = \left( \frac{(S_u - X) - C_d}{S_u - S_d}, \frac{(S_u - X) - \frac{(S_u - X) - C_d}{S_u - S_d} S_u}{1 + r^f} \right)$. Hence, the unique replication value of the option for $A$ is

$$\alpha S_0 + \beta = \frac{(S_u - X) - C_d}{S_u - S_d} S_0 + \frac{(S_u - X) - \frac{(S_u - X) - C_d}{S_u - S_d} S_u}{1 + r^f} = 8.954$$

It is not a coincidence that this equals $C_0 = \frac{p_0^B (S_u - X)}{(1 + r^f)(1 - \lambda^C)} = 8.954$. The stock price, $S_0$, per construction makes the two equal. Hence, nothing can be gained from buying less option and instead invest in stock or vice versa.

Next, we look at the $d$–state. Here, $A$'s portfolio value is 0. The no–default constraint then implies that there is no portfolio choice. The option is also worthless. The only thing to consider is the value of the stock. What's the reserve price of the stock for $B$?

$$\frac{p_1^B S_{du} + (1-p_1^B)S_{dd}}{1+r^f} = 83.000$$

which is then the price in equilibrium as $B$'s no-default constraint is not binding.

Finally, we must look at what happens in the $u$–state. This is where the early exercise takes place. Again, we utilize that $B$'s no–default constraint is not binding to determine the reserve price for $B$ for stock and option. In equilibrium $B$ is long stock. Hence the reserve price which must also be the equilibrium price is

$$\frac{p_1^B S_{uu} + (1-p_1^B)S_{ud}}{1+r^f} = 119.519$$

The option price must be

$$\frac{p_1^B (S_{uu} - X)}{1+r^f} = 19.889$$

The option and stock prices per construction leave no arbitrage opportunities as the option price reflect the cost of replicating it using stock and risk-free asset. $B$ arrives at the $u$–state being long option and stock. The first thing we see is that it's strictly better to exercise the option early and sell the stock compared to selling the option. Selling the option gives $C_u(1-\lambda^C) = 19.491$ whereas exercising and selling the stock gives $S_u - X = 19.519$. As $A$ in the $u$–state goes from being long option to short option this means that the long position has been exercised early. Observe also that, once again, $A$'s no–default condition is binding. The portfolio value in the $uu$–state is zero (per construction):

$$A_u^S S_{uu} + A_u^C C_{uu} + A_u^R (1+r^f) = 0$$

The reserve price for $A$ as stock buyer financed through the risk-free asset is

$$\frac{p_1^A S_{uu} + (1 - p_1^A S_{ud})}{1 + r^f} = 119.082$$

which is less than $S_u = 119.519$. Hence, $A$ is not better off buying stock financed by risk-free asset. How about by selling less option? The reserve price for $A$ as option seller and investing the proceeds in the risk-free asset is

$$\frac{p_1^A (S_{uu} - X)}{1 + r^f} = 19.457$$

This is smaller than $C_0(1-\lambda^C) = 19.491$ which can be gained from selling the option. Hence, it is not attractive to sell less option and invest correspondingly less in risk-free asset. But the option could also be replicated using the stock and the risk-free asset and then extra option could be sold. Would this represent an improvement to $A$? To replicate the option costs $A$

$$S_u \frac{S_{uu} - X}{S_{uu} - S_{ud}} + \frac{(S_{uu} - X) - \frac{S_{uu} - X}{S_{uu} - S_{ud}} S_{uu}}{1 + r^f} = 19.771$$

following the binomial replication argument. This exceeds what the option can be sold for net of transaction costs. Hence it is no improvement for $A$. How about selling less option and then replicate a short option position? To do this $A$ will be subject to short-sale costs that affect the proceeds from replicating a short option position. The proceeds will be

$$S_u \frac{S_{uu} - X}{S_{uu} - S_{ud}} + \frac{(S_{uu} - X) - \frac{S_{uu} - X}{S_{uu} - S_{ud}} (S_{uu} - S_u L)}{1 + r^f} = 19.293$$

To replicate writing an option hence gives less than selling the option, $C_0(1 - \lambda^C) = 19.491$. Hence, it is not attractive to sell less option and instead replicate through short-sale of stock.

To verify the equilibrium it has been sufficient to verify that no one benefits from holding more/less stock, option or risk-free asset on the margin. This is because the utility is linear in the number of stocks, options and risk-free assets held.

In conclusion we can say that it in fact is an equilibrium, and in this equilibrium the option is exercised early in the $u$–state at time $t = 1$.

## Proof of Proposition 4 and Proposition 5

The optimal exercise boundary for the PDE (7)–(8) is the mathematical equivalent to the problem solved by Dewynne, Howison, Rupf, and Wilmott (1993), a Black-Scholes pricing model with the stock paying a continuous dividend yield. Our equivalent of the continuous dividend yield they present is $\tilde{l}^i + \psi^i(m^{i,C} - \tilde{m}^{i,S})$. Note that either $\tilde{l}^i > 0$ or $\psi^i > 0$ and both being non-negative combined with $m^{i,C} > m^{i,S}$ implies that $\tilde{l}^i + \psi^i(m^{i,C} - \tilde{m}^{i,S}) > 0$. Using their established result we then can say that if $\tilde{l}^i > 0$ or $\psi^i > 0$ then:

$$B^* = \frac{\mu_1 X}{\mu_1 - 1}$$

where

$$\begin{aligned}
\mu_1 &= \frac{1}{2}(1 - k_1 + \sqrt{(1 - k_1)^2 + 4k_2}) \\
k_1 &= \frac{2(r^f + \tilde{m}^{i,S}\psi^i - \tilde{l}^i)}{\sigma^2} \\
k_2 &= \frac{2(r^f + m^{i,C}\psi^i)}{\sigma^2}
\end{aligned}$$

Remember that

$$(\tilde{m}^{i,S}, \tilde{l}^i) = \begin{cases} (m^{i,S}, l^i) & \text{if agent } i \text{ is long stock,} \\ (-m^{i,S}, L^i) & \text{if agent } i \text{ is short stock,} \end{cases}$$

Also, observe that since $k_2$ is always positive when $\psi^i > 0$ or $\tilde{l}^i > 0$ so is $\mu_1$ the following relations hold

$$
\begin{aligned}
\frac{\partial B^*}{\partial \mu_1} &= \frac{-X}{(\mu_1 - 1)^2} < 0 \\
\frac{\partial \mu_1}{\partial k_1} &= -\frac{1}{2} + \frac{-(1-k_1)}{2\sqrt{(1-k_1)^2 + 4k_2}} < 0 \\
\frac{\partial \mu_1}{\partial k_2} &= \frac{1}{\sqrt{(1-k_1)^2 + 4k_2}} > 0
\end{aligned}
$$

We want to show $\frac{\partial B^*}{\partial L^i} < 0$ if agent $i$ is short stock and $\frac{\partial B^*}{\partial l^i} < 0$ if agent $i$ is long stock. Observe that $\frac{\partial \tilde{l}^i}{\partial L^i} > 0$ when $i$ is short stock and $\frac{\partial \tilde{l}^i}{\partial l^i} > 0$ when $i$ is long stock. Also note that $k_2$ does not depend on $\tilde{l}^i$ and that $\frac{\partial k_1}{\partial \tilde{l}^i} = \frac{-2}{\sigma^2} < 0$. Hence $\frac{\partial B^*}{\partial \tilde{l}^i} = \frac{\partial B^*}{\partial \mu_1}\frac{\partial \mu_1}{\partial k_1}\frac{\partial k_1}{\partial \tilde{l}^i} < 0$. The conclusion is that $\frac{\partial B^*}{\partial L^i} = \frac{\partial B^*}{\partial \tilde{l}^i}\frac{\partial \tilde{l}^i}{\partial L^i} < 0$ if agent $i$ is short stock and $\frac{\partial B^*}{\partial l^i} = \frac{\partial B^*}{\partial \tilde{l}^i}\frac{\partial \tilde{l}^i}{\partial l^i} < 0$ if agent $i$ is long stock.

We also want to show that $\frac{\partial B^*}{\partial \psi^i} < 0$ when agent $i$ is short the stock (ie. $\tilde{m}^{i,S} = -m^{i,S}$). We get:

$$
\begin{aligned}
\frac{\partial \mu_1}{\partial \psi^i} &= \frac{\partial \mu_1}{\partial k_1}\frac{\partial k_1}{\partial \psi^i} + \frac{\partial \mu_1}{\partial k_2}\frac{\partial k_2}{\partial \psi^i} \\
&= \left(-\frac{1}{2} + \frac{-(1-k_1)}{2\sqrt{(1-k_1)^2 + 4k_2}}\right)\frac{-2m^{i,S}}{\sigma^2} + \left(\frac{1}{\sqrt{(1-k_1)^2 + 4k_2}}\right)\frac{2m^{i,C}}{\sigma^2} \\
&= \left(1 + \frac{(1-k_1)}{\sqrt{(1-k_1)^2 + 4k_2}}\right)\frac{m^{i,S}}{\sigma^2} + \left(\frac{1}{\sqrt{(1-k_1)^2 + 4k_2}}\right)\frac{2m^{i,C}}{\sigma^2} > 0
\end{aligned}
$$

Using that $\frac{\partial B^*}{\partial \mu_1} < 0$ we get that $\frac{\partial B^*}{\partial \psi^i} = \frac{\partial B^*}{\partial \mu_1}\frac{\partial \mu_1}{\partial \psi^i} < 0$ when agent $i$ is short stock.

Next, we show that $\frac{\partial B^*}{\partial m^{i,C}} < 0$. $k_1$ does not depend on $m^{i,C}$ whereas $\frac{\partial k_2}{\partial m^{i,C}} = \frac{2\psi^i}{\sigma^2}$. This is non-negative, and strictly positive if $\psi^i > 0$. This means that $\frac{\partial B^*}{\partial m^{i,C}} = \frac{\partial B^*}{\partial \mu_1}\frac{\partial \mu_1}{\partial k_2}\frac{\partial k_2}{\partial m^{i,C}} \leq 0$ with strict inequality if $\psi^i > 0$.

We want to show that when agent $i$ is short stock then $\frac{\partial B^*}{\partial m^{i,S}} \leq 0$ and when agent $i$ is long stock then $\frac{\partial B^*}{\partial m^{i,S}} \geq 0$, both inequalities being strict if $\psi^i > 0$. First, observe that $\frac{\partial \tilde{m}^{i,S}}{\partial m^{i,S}} = -1$ when agent $i$ is short stock and $\frac{\partial \tilde{m}^{i,S}}{\partial m^{i,S}} = 1$ when agent $i$ is long stock. Hence it suffices to prove that $\frac{\partial B^*}{\partial \tilde{m}^{i,S}} \geq 0$ with strict inequality when $\psi^i > 0$. $k_2$ does not depend on $\tilde{m}^{i,S}$ while $\frac{\partial k_1}{\partial \tilde{m}^{i,S}} = \frac{2\psi^i}{\sigma^2} \geq 0$ ($>$ when $\psi^i > 0$). This leads to $\frac{\partial B^*}{\partial \tilde{m}^{i,S}} = \frac{\partial B^*}{\partial \mu_1}\frac{\partial \mu_1}{\partial k_1}\frac{\partial k_1}{\partial \tilde{m}^{i,S}} \geq 0$ with strict inequality when $\psi^i > 0$.

Finally it is to be shown that $\frac{\partial B^*}{\partial \sigma} > 0$. First, we look at the sign of

$$\begin{aligned}
\frac{\partial \mu_1}{\partial \sigma^2} &= \frac{\partial \mu_1}{\partial k_1}\frac{\partial k_1}{\partial \sigma^2} + \frac{\partial \mu_1}{\partial k_2}\frac{\partial k_2}{\partial \sigma^2} \\
&= \left(-\frac{1}{2} + \frac{-(1-k_1)}{2\sqrt{(1-k_1)^2 + 4k_2}}\right)\frac{-k_1}{\sigma^2} + \frac{1}{\sqrt{(1-k_1)^2+4k_2}}\frac{-k_2}{\sigma^2}
\end{aligned}$$

Multiplying with $2\sigma^2\sqrt{(1-k_1)^2+4k_2}$ does not change the sign and we get

$$\begin{aligned}
&k_1\sqrt{(1-k_1)^2+4k_2} + k_1(1-k_1) - 2k_2 \\
=&k_1\sqrt{(1+k_1)^2+4\delta} - k_1^2 - k_1 - 2\delta
\end{aligned} \tag{24}$$

where $\delta := k_2 - k_1 = \frac{2(\psi^i(m^{i,C} - \tilde{m}^{i,S}) + \tilde{l}^i)}{\sigma^2} > 0$. We can consider the expression (24) a continuous function in $k_1$ which is well-defined for all real values. However, since $k_2 > 0$ we get $k_1 > -\delta$. We'll now search for roots for this function. We get that all roots must fulfill:

$$\sqrt{k_1^2(1+k_1)^2 + 4\delta k_1^2} = k_1^2 + k_1 + 2\delta$$

leading to

$$k_1^4 + 2k_1^3 + k_1^2 + 4\delta k_1^2 = k_1^4 + k_1^2 + 4\delta^2 + 2k_1^3 + 4\delta k_1^2 + 4\delta k_1$$

Terms cancel out and the only candidate for a root is $k_1 = -\delta$. The continuity of (24) in $k_1$ combined with the fact that it is well-defined for all real values of $k_1$ now means that the expression has constant sign for $k_1 > -\delta$. For $k_1 = 0$ the expression is strictly negative since $\delta > 0$. Hence it must be strictly negative for all $k_1 > -\delta$, which we know is fulfilled. This leads to $\frac{\partial \mu_1}{\partial \sigma^2} < 0$. Using that $\frac{\partial B^*}{\partial \mu_1} < 0$ and $\frac{\partial \sigma^2}{\partial \sigma} = 2\sigma > 0$ we get that $\frac{\partial B^*}{\partial \sigma} = \frac{\partial B^*}{\partial \mu_1}\frac{\partial \mu_1}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial \sigma} > 0$.

Table I
**Data Sources.**

This table shows the data sources used in our study, the variables that we use, the start and end date of each data source, the number of securities, and the number of observations (which is the number of rows in the data).

| Data set | Data | Start date | End date | Number of call options/ convertible bond series | Number of underlying securities | Number of observations |
|---|---|---|---|---|---|---|
| CRSP[1] | Dividends, prices, corporate events | 30-08-1985 | 31-12-2011 | | 23,597 | 18,314,652 |
| OCC Exercises[2] | Exercises of equity options | 01-07-2001 | 31-08-2010 | 821,052 | 5,727 | 7,852,739 |
| OptionMetrics | Option prices, open interests, volatilities, expected future dividends | 01-01-1996 | 31-01-2012 | 3,949,199 | 7,509 | 355,259,334 |
| Data Explorers[3] | Short-sale costs | 19-06-2002 | 03-12-2012 | | 41,188 | 55,139,348 |
| Mergent FISD[4] | Convertible bond features and conversions | 30-08-1985 | 05-06-2012 | 4,539 | 1,731 | 14,194 |
| Bloomberg | LIBOR-OIS spreads | 02-01-1990 | 22-01-2013 | | | 8,620 |

[1] CRSP data start in 1926, but we only use it when we have option and convertible bond data.

[2] Data for the months November 2001, January and July 2002, and January 2006 are missing.

[3] We focus on the Daily Cost of Borrow Score (DCBS), which is first observed from October 22, 2003.

[4] Mergent FISD has earlier bond observations, but this is the first date a convertible bond can be observed

Table II
**Summary Statistics.**

This table summarizes the number of early exercises and early conversions in the sample used in our study. It reports the number of contracts exercised and bonds converted. The total numbers are broken into categories of Moneyness, Expiration, and Agent Type performing the exercise. Options are defined as "out of the money" if the closing stock price is below the strike price, "in the money" if the stock price is 0–25% above the strike price, and "deep in the money" if the stock price is more than 25% higher than the strike price. For convertible bonds the definition is parallel with conversion price used instead of strike price.

**Panel A: Early Exercises of Equity Call Options in sample**

| | Exercises (number of contracts, millions) | Exercises (value of strike, USD millions) | Exercises (intrinsic value, USD millions) |
|---|---|---|---|
| **All** | 1,806 | 36,250 | 22,811 |
| **By moneyness** | | | |
| Out of the money | 2 | 38 | -1 |
| In the money | 480 | 14,170 | 1,942 |
| Deep in the money | 1,324 | 22,042 | 20,870 |
| **By time to expiration** | | | |
| Less than 3 months | 1,720 | 35,422 | 21,199 |
| Between 3 and 9 months | 72 | 690 | 1,006 |
| More than 9 months | 13 | 137 | 605 |
| **By agent type** | | | |
| Customer | 808 | 16,007 | 6,338 |
| Firm | 81 | 1,955 | 998 |
| Market maker | 916 | 18,288 | 15,475 |

**Panel B: Early Conversions of Convertible Bonds in sample**

| | Conversions (number of bonds, millions) | Conversions (principal amount, USD millions) | Conversions (value of stock, USD millions) |
|---|---|---|---|
| **All** | 25.4 | 5,655 | 7,732 |
| **By moneyness** | | | |
| Out of the money | 3.8 | 2,063 | 1,255 |
| In the money | 15.4 | 1,001 | 1,127 |
| Deep in the money | 6.3 | 2,591 | 5,350 |

Table III

**Early Exercise of Equity Options by Short-Sale Costs.**

This table shows the average number of early option exercises as a fraction of open interest on the previous day for options sorted on the short-sale cost of the underlying equity. The table further classifies options by their moneyness, expiration, and agent type. Options are defined as "out of the money" if the closing stock price is below the strike price, "in the money" if stock price is 0–25% above strike price, and "deep in the money" if stock price is more than 25% higher than strike price. We note that the number of exercises for each agent type is reported as a fraction of the total open interest since our data do not include open interest by agent type. Standard errors are reported in parenthesis.

| | 1 Low cost of shorting | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 High cost of shorting |
|---|---|---|---|---|---|---|---|---|---|---|
| **All** | 0.17% | 0.31% | 0.44% | 0.58% | 0.85% | 1.21% | 1.74% | 2.57% | 3.05% | 4.28% |
| | (0.00%) | (0.00%) | (0.01%) | (0.01%) | (0.02%) | (0.02%) | (0.03%) | (0.04%) | (0.05%) | (0.07%) |
| **By moneyness** | | | | | | | | | | |
| Out of the money | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.02% |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.01%) |
| In the money | 0.10% | 0.18% | 0.24% | 0.28% | 0.31% | 0.32% | 0.38% | 0.41% | 0.50% | 0.63% |
| | (0.00%) | (0.00%) | (0.01%) | (0.01%) | (0.02%) | (0.02%) | (0.02%) | (0.03%) | (0.03%) | (0.05%) |
| Deep in the money | 0.31% | 0.55% | 0.78% | 1.06% | 1.51% | 2.23% | 3.13% | 4.52% | 5.75% | 8.44% |
| | (0.00%) | (0.01%) | (0.01%) | (0.02%) | (0.03%) | (0.04%) | (0.05%) | (0.08%) | (0.09%) | (0.14%) |
| **By time to expiration** | | | | | | | | | | |
| Less than 3 months | 0.28% | 0.51% | 0.71% | 0.99% | 1.49% | 2.10% | 2.97% | 4.51% | 5.32% | 7.58% |
| | (0.00%) | (0.01%) | (0.01%) | (0.02%) | (0.03%) | (0.04%) | (0.05%) | (0.08%) | (0.09%) | (0.14%) |
| Between 3 and 9 months | 0.02% | 0.05% | 0.09% | 0.10% | 0.21% | 0.29% | 0.48% | 0.61% | 0.88% | 1.15% |
| | (0.00%) | (0.00%) | (0.01%) | (0.01%) | (0.01%) | (0.02%) | (0.02%) | (0.03%) | (0.04%) | (0.07%) |
| More than 9 months | 0.01% | 0.03% | 0.07% | 0.08% | 0.10% | 0.14% | 0.23% | 0.53% | 0.40% | 0.29% |
| | (0.00%) | (0.00%) | (0.01%) | (0.01%) | (0.02%) | (0.03%) | (0.03%) | (0.05%) | (0.05%) | (0.05%) |
| **By agent type** | | | | | | | | | | |
| Customer | 0.11% | 0.19% | 0.24% | 0.32% | 0.43% | 0.56% | 0.76% | 0.96% | 1.24% | 1.76% |
| | (0.00%) | (0.00%) | (0.00%) | (0.01%) | (0.01%) | (0.02%) | (0.02%) | (0.02%) | (0.03%) | (0.04%) |
| Firm | 0.01% | 0.02% | 0.02% | 0.03% | 0.06% | 0.07% | 0.11% | 0.13% | 0.17% | 0.17% |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.01%) | (0.01%) | (0.01%) | (0.01%) |
| Market maker | 0.05% | 0.11% | 0.18% | 0.23% | 0.37% | 0.58% | 0.86% | 1.48% | 1.64% | 2.36% |
| | (0.00%) | (0.00%) | (0.00%) | (0.01%) | (0.01%) | (0.02%) | (0.02%) | (0.03%) | (0.03%) | (0.05%) |

Table IV

**Early Exercise of Equity Options by Transaction Costs.**

This table shows the average number of early option exercises as a fraction of open interest on the previous day for options sorted on the transaction costs. Transaction costs for options are measured daily as the bid-ask spread divided by the mid price for the at-the-money option with same underlying, expiration, and observation day. Observations are grouped into quintiles by transaction costs. The table further classifies options by their moneyness, expiration, and agent type. Options are defined as "out of the money" if the closing stock price is below the strike price, "in the money" if stock price is 0–25% above strike price, and "deep in the money" if stock price is more than 25% larger than strike price. We note that the number of exercises for each agent type is reported as a fraction of the total open interest since our data do not include open interest by agent type. Standard errors are reported in parenthesis.

| | **1 Low T-cost** | **2** | **3** | **4** | **5 High T-cost** | **5-1** |
|---|---|---|---|---|---|---|
| **All** | 0.13% | 0.16% | 0.21% | 0.28% | 0.51% | **0.38%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.00%)** |
| **By moneyness** | | | | | | |
| Out of the money | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | **0.00%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.00%)** |
| In the money | 0.02% | 0.03% | 0.05% | 0.11% | 0.37% | **0.35%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.00%)** |
| Deep in the money | 0.25% | 0.32% | 0.47% | 0.62% | 1.00% | **0.76%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.01%) | **(0.01%)** |
| **By time to expiration** | | | | | | |
| Less than 3 months | 0.29% | 0.30% | 0.36% | 0.42% | 0.63% | **0.35%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.01%)** |
| Between 3 and 9 months | 0.04% | 0.04% | 0.04% | 0.05% | 0.06% | **0.03%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.00%)** |
| More than 9 months | 0.02% | 0.02% | 0.03% | 0.04% | 0.02% | **0.00%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.00%)** |
| **By agent type** | | | | | | |
| Customer | 0.06% | 0.07% | 0.11% | 0.17% | 0.37% | **0.31%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.00%)** |
| Firm | 0.01% | 0.01% | 0.01% | 0.01% | 0.02% | **0.02%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.00%)** |
| Market maker | 0.07% | 0.07% | 0.09% | 0.10% | 0.12% | **0.05%** |
| | (0.00%) | (0.00%) | (0.00%) | (0.00%) | (0.00%) | **(0.00%)** |

Table V

**Early Exercise of Equity Options: Model-Implied Optimal Exercise Boundary.**

This table shows the fraction of the observed early option exercise decisions that can be rationalized by our model, that is, occur when the daily high of the stock price, $S$, is above the model-implied estimated exercise boundary, $B$. Here, $B$ is computed as the solution to the PDE (7)–(8) based on the estimated volatility, short-sale cost, and funding cost. The table reports this fraction for all agents and by agent type (across the different rows) and for four different implementations of the model (across the columns). The first column estimates the volatility as the 60-day historical volatility and the short-sale cost as the median short-sale cost among stock loans of this type (DCBS score). The second column uses a more conservative (i.e., lower) estimate of expected future volatility which reflects that option exercises often happen after volatile periods and volatility mean reverts, namely the minimum of the current 60-day historical volatility and the median 60-day historical volatility over the stock's Option Metrics sample. The third column also uses a conservative (i.e., higher) estimate of short-sale costs, namely the 90% percentile among stock loans of this type. The fourth column has the most conservative boundary, which is 90% of the estimated boundary based on the conservative volatility and the conservative short-sale cost.

| | Exercises w/ S>B | Exercises w/ S>B, conservative volatility estimate | Exercises w/ S>B, conservative volatility and short-sale cost estimates | Exercises w/ S>B, conservative boundary |
|---|---|---|---|---|
| **All** | 65.8% | 75.1% | 79.4% | 84.2% |
| **By agent type** | | | | |
| Customer | 41.3% | 51.3% | 58.2% | 66.8% |
| Firm | 64.2% | 76.6% | 79.7% | 84.2% |
| Market maker | 86.0% | 94.4% | 96.7% | 98.4% |

Table VI

**Early Exercise of Equity Options: Regression Analysis.**

This table shows logit (Panel A) and probit (Panel B) regressions for the determinants of early option exercises. The dependent variable is 1 for every option contract that is exercised early and 0 for every contract outstanding at end of the previous day that is not exercised early on a given day. The independent variables are as follows. The first variable is the indicator that the stock closing price $S$ is above the model-implied exercise boundary, $B$, and the second variable is $(S-B)^+/B$. The third variable is the indicator that the option's closing bid price, $C_{bid}$, is below the intrinsic value given the strike price, $X$. Moneyness is $S/X$. Short-sale cost score is the observed DCBS score, where a higher score indicates a higher cost. Bid-ask spread is the relative bid-ask spread for the option closest to at-the-money with same underlying and expiry date. Time to expiration is in 100 days. Historical volatility is estimated based on the previous 60 days stock returns. The LIBOR-OIS spread is in basis points. We report estimated $t$-statistics in parenthesis based on standard errors that account for cross-sectional and time-series correlations using the method of Fama and MacBeth (1973). First, the regressions are estimated in each 3-month subsample. Second, the full sample parameters are estimated as the sample means of the estimates and the standard errors are estimated based on the sample standard deviation of parameter estimates across subsamples, using the Newey-West correction.

Panel A: Logit

| Independent variables | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $1_{(S>B)}$ | 4.09 | | | 1.84 | 1.19 |
| | (13.42) | | | (9.84) | (6.06) |
| $[(S-B)/B]^+$ | | 2.13 | | 0.45 | 0.39 |
| | | (5.98) | | (2.72) | (1.30) |
| $1_{(C\_bid<S-X)}$ | | | 5.23 | 4.46 | 3.92 |
| | | | (20.74) | (17.93) | (23.47) |
| Moneyness | | | | | 0.03 |
| | | | | | (0.32) |
| Short-sale cost score | | | | | 0.33 |
| | | | | | (10.25) |
| Bid-ask spread | | | | | 0.26 |
| | | | | | (3.19) |
| Time to expiration | | | | | -1.61 |
| | | | | | -(6.20) |
| Historical volatility | | | | | 0.73 |
| | | | | | (5.50) |
| LIBOR-OIS spread | | | | | 0.25 |
| | | | | | (0.99) |
| Intercept | -8.54 | -7.75 | -10.63 | -10.70 | -12.89 |
| | (38.83) | -(36.70) | -(53.54) | -(50.09) | -(6.25) |
| Method | Logit | Logit | Logit | Logit | Logit |

Panel B: Probit

| Independent variables | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $1_{(S>B)}$ | 1.29 | | | 0.62 | 0.42 |
| | (11.71) | | | (9.54) | (5.91) |
| $[(S-B)/B]^{+}$ | | 0.91 | | 0.24 | 0.20 |
| | | (5.81) | | (2.89) | (1.78) |
| $1_{(C\_bid<S-K)}$ | | | 1.45 | 1.22 | 1.15 |
| | | | (17.04) | (17.85) | (18.13) |
| Moneyness | | | | | 0.03 |
| | | | | | (0.82) |
| Short-sale cost score | | | | | 0.13 |
| | | | | | (9.88) |
| Bid-ask spread | | | | | 0.12 |
| | | | | | (6.44) |
| Time to expiration | | | | | -0.57 |
| | | | | | -(6.00) |
| Historical volatility | | | | | 0.24 |
| | | | | | (6.71) |
| LIBOR-OIS spread | | | | | 0.06 |
| | | | | | (0.79) |
| Intercept | -3.54 | -3.35 | -4.06 | -4.10 | -4.73 |
| | (46.99) | -(58.35) | -(85.96) | -(79.33) | -(7.19) |
| Method | Probit | Probit | Probit | Probit | Probit |

Table VII

**Early Conversion of Convertible Bonds.**

This table shows the average amount of early conversion as a fraction of the outstanding amount for convertible bonds sorted on the short-sale costs of the underlying equity. The observations are grouped into Low, Medium, and High cost of shorting based on the DCBS score from Data Explorers. "Low" cost reflects a DCBS score of 1, "Medium" a score of 2–5, and "High" 6–10. The table further classifies options by their moneyness. Convertible bonds are defined as "out of the money" when stock price is below conversion price, "in the money" when stock price is at or up to 25% above the conversion price and "deep in the money" when stock price is more than 25% above conversion price. The conversion price defines the amount of face value of bond that must be converted to obtain one share of the underlying stock. Standard errors are reported in parenthesis.

| | **1** | **2** | **3** | **3-1** |
|---|---|---|---|---|
| | **Low cost of shorting** | | **High cost of shorting** | |
| **Panel A: All** | 0.05% | 0.05% | 0.12% | **0.07%** |
| | (0.01%) | (0.02%) | (0.07%) | **(0.09%)** |
| **Panel B: By Moneyness** | | | | |
| Out of the money | 0.02% | 0.00% | 0.00% | **-0.03%** |
| | (0.01%) | (0.00%) | (0.00%) | **(0.02%)** |
| In the money | 0.06% | 0.07% | 0.16% | **0.07%** |
| | (0.03%) | (0.07%) | (0.13%) | **(0.13%)** |
| Deep in the money | 0.07% | 0.12% | 0.29% | **0.19%** |
| | (0.02%) | (0.06%) | (0.18%) | **(0.23%)** |

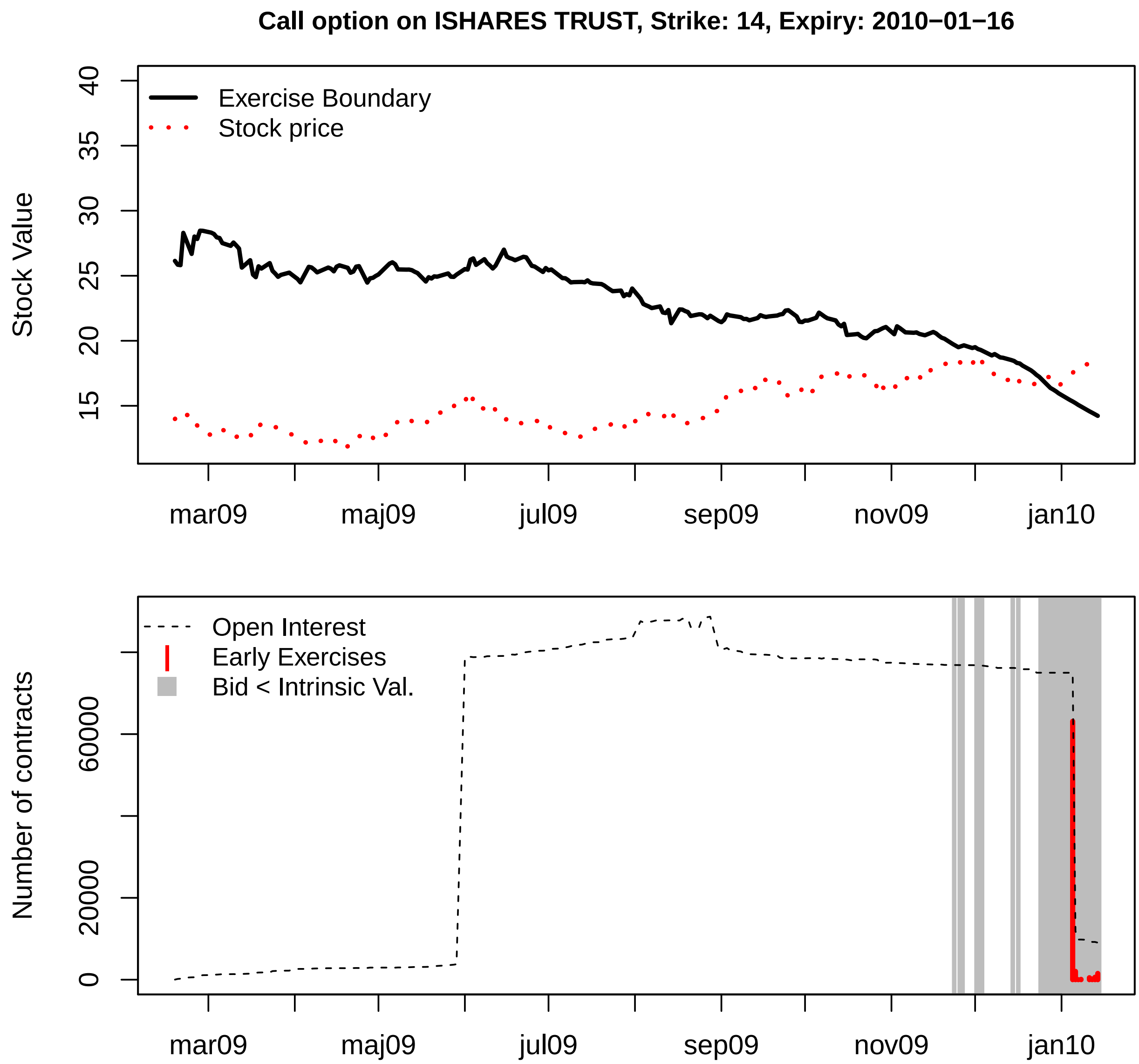


Figure 2: **Early Exercise of Call Options before Expiration: IShares Trust.** The upper panel shows the daily closing price of Ishares Trust stock (silver) and the model-implied optimal exercise boundary based on the following parameters: The risk-free rate is the Fed funds rate, the volatility is estimated as the 60-day historical volatility, the short-sale fee is from Dataexplorers, the funding cost is the LIBOR-OIS spread, and the assumed margin requirements are $m^{i,C} = 100\%$ for the option and $m^{i,S} = 50\%$ for the stock. The lower panel shows the open interest and early exercise of the option. Shortly after the stock price is above the exercise boundary, 84% of the open interest is exercised in one day. Early exercises are also observed the following days. The closing bid-price of the option is below closing bid-price of the stock minus strike price in periods with gray background.

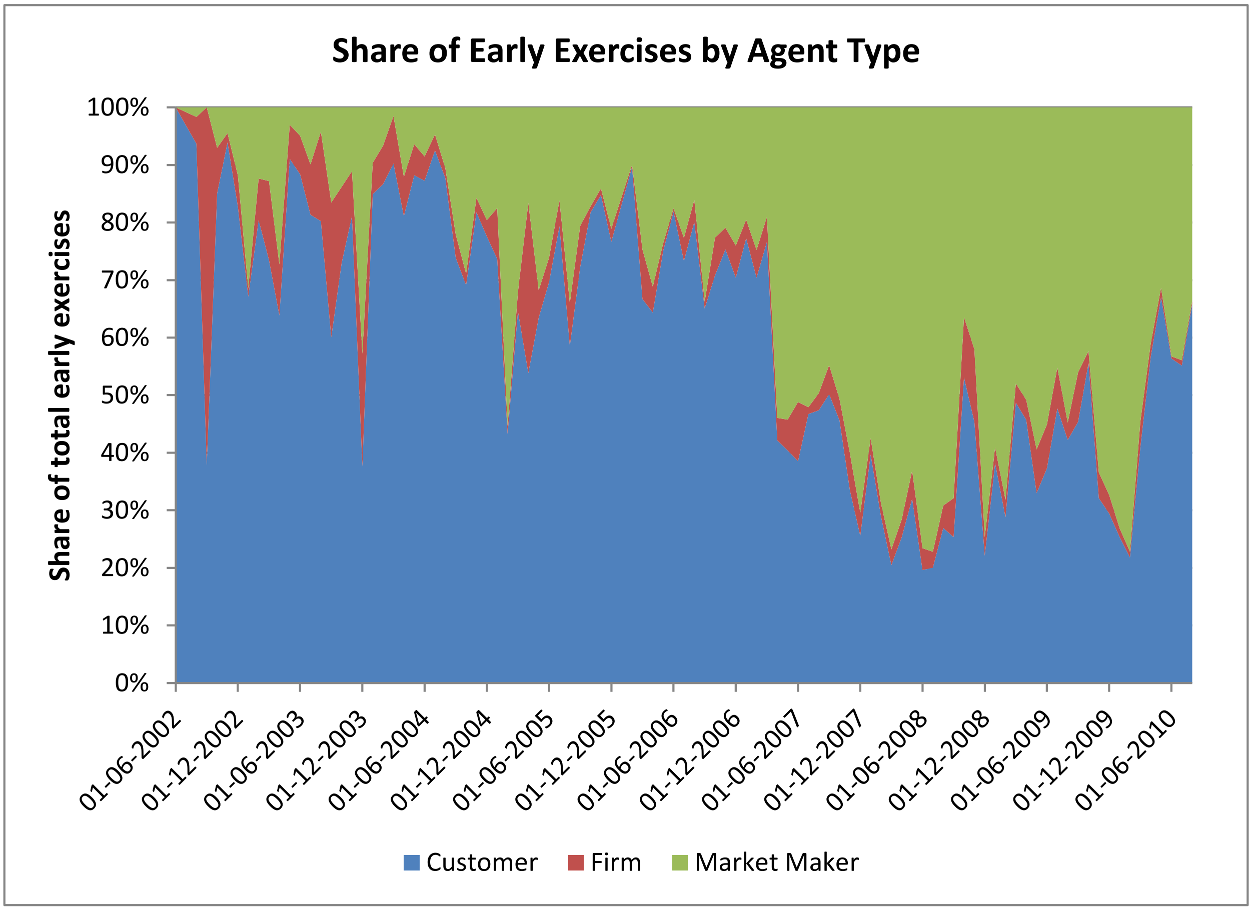


Figure 3: **Share of Early Exercises by Agent Type Over Time.** This figure shows how the monthly relative share of the number of total early exercises are distributed among the three agent types: Customers (retail costumers and hedge funds), Firms (proprietary traders), and Market Makers. We see early exercises for all three groups and an increasing share from Market Makers over time.

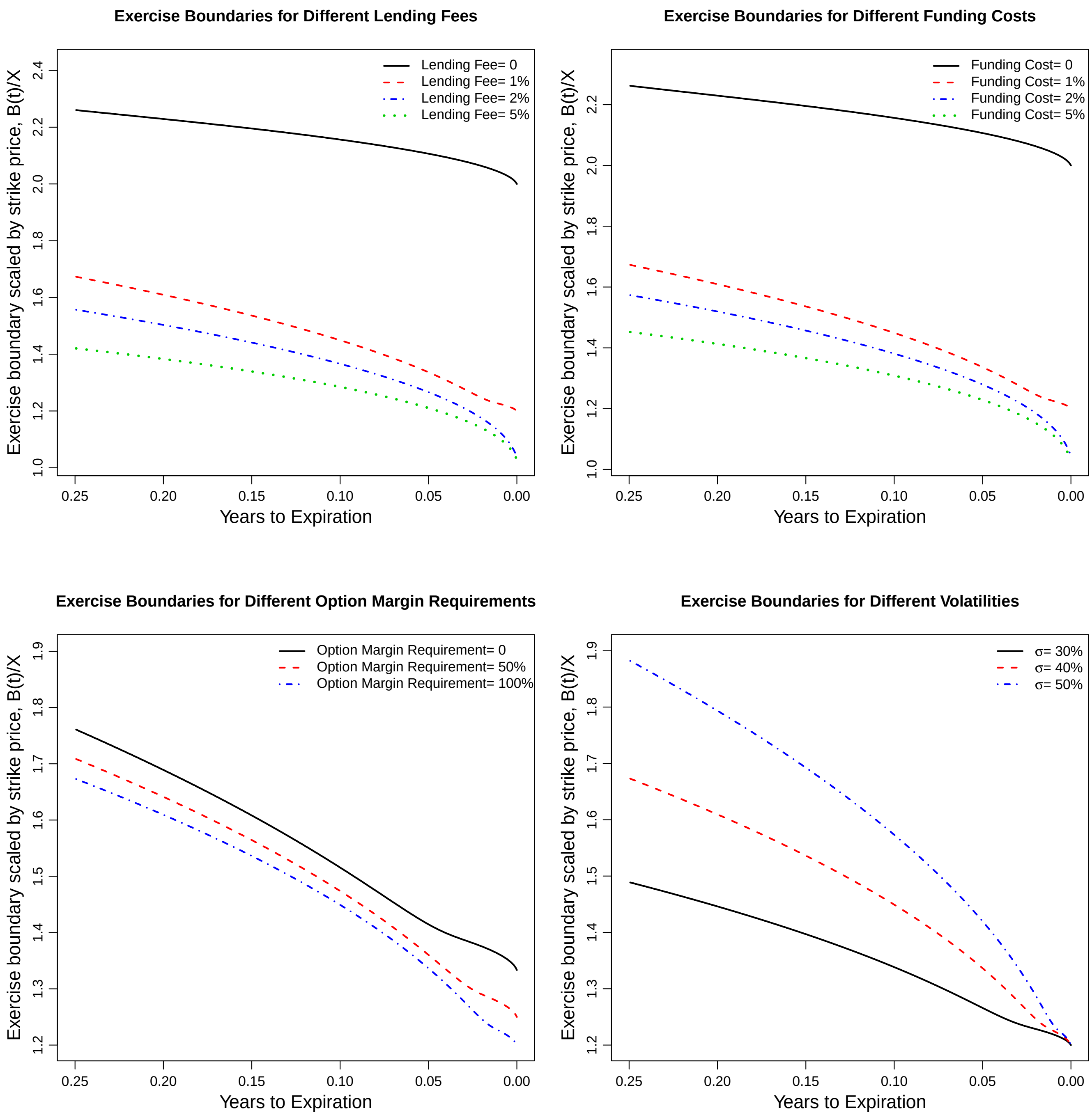


Figure 4: **The Optimal Exercise Boundary with Frictions: Comparative Statics.** This figure shows theoretical exercise boundaries for equity options with frictions for an agent who is short stock. The boundary is a solution to the PDE (7)–(8) for a stock that pays no dividend. Each graph varies the parameters around a base-case cost where the risk-free rate is $r^f = 2\%$, the lending fee is $L^i = 1\%$, the funding cost is $\psi^i = 1\%$, the volatility is $\sigma = 40\%$, and margin requirements are $m^{i,S} = 50\%$ for the stock and $m^{i,C} = 100\%$ for the option. Early exercise is seen to be increasing in lending fees, funding costs, option margin requirements, and decreasing in volatility and time to expiration.

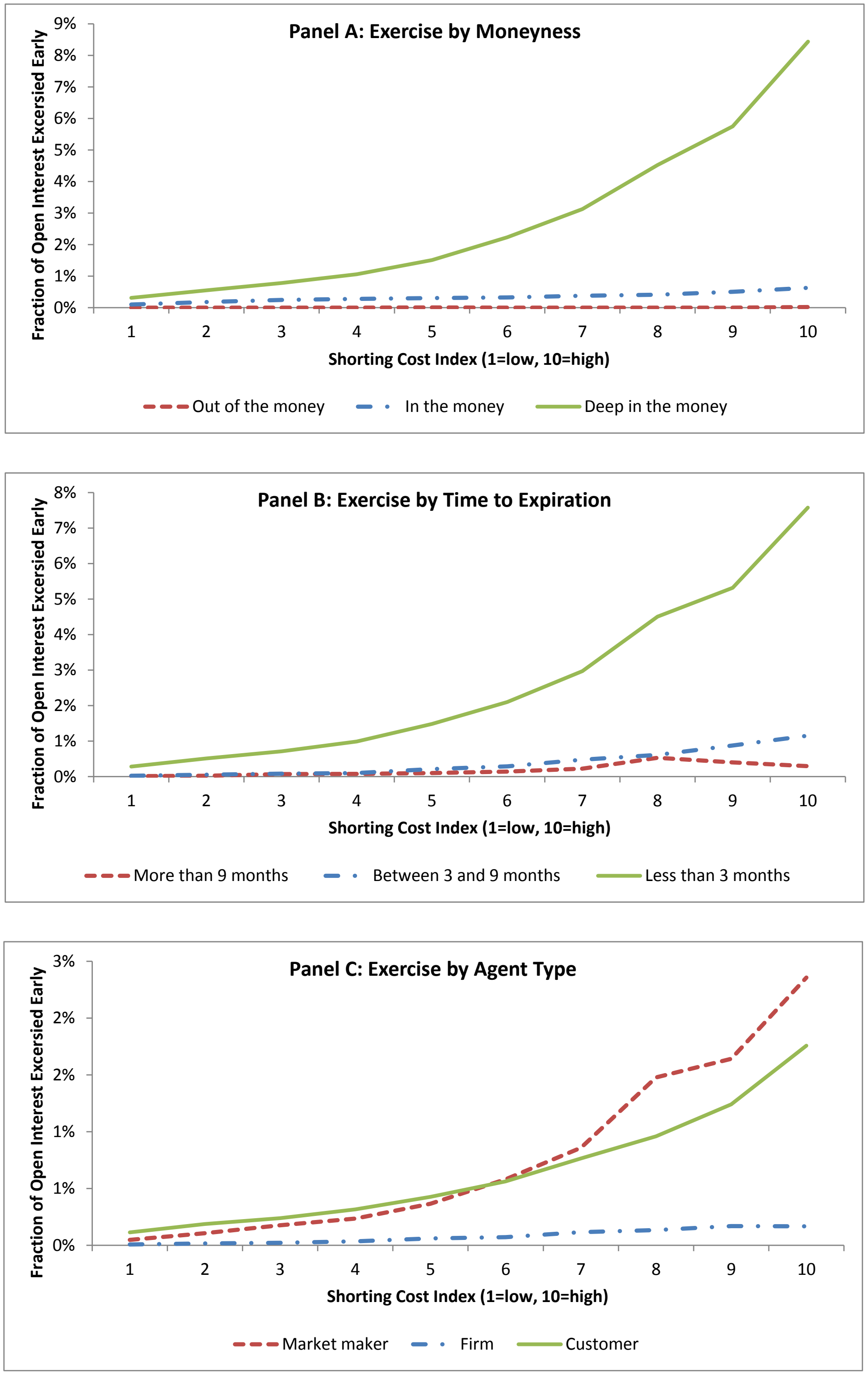


Figure 5: **Actual Early Exercise of Equity Options for Varying Short-Sale Costs.** This figure shows the empirical fraction of equity options exercised early as a function of the short-sale costs. Panel A groups the data by the moneyness of the option, Panel B by expiration, and Panel C by agent type. Consistent with our theory, early exercise is increasing in short-sale costs, increasing in moneyness, decreasing in expiration, and the exercise pattern is prevalent for all agent types including professional market makers and firm proprietary traders.

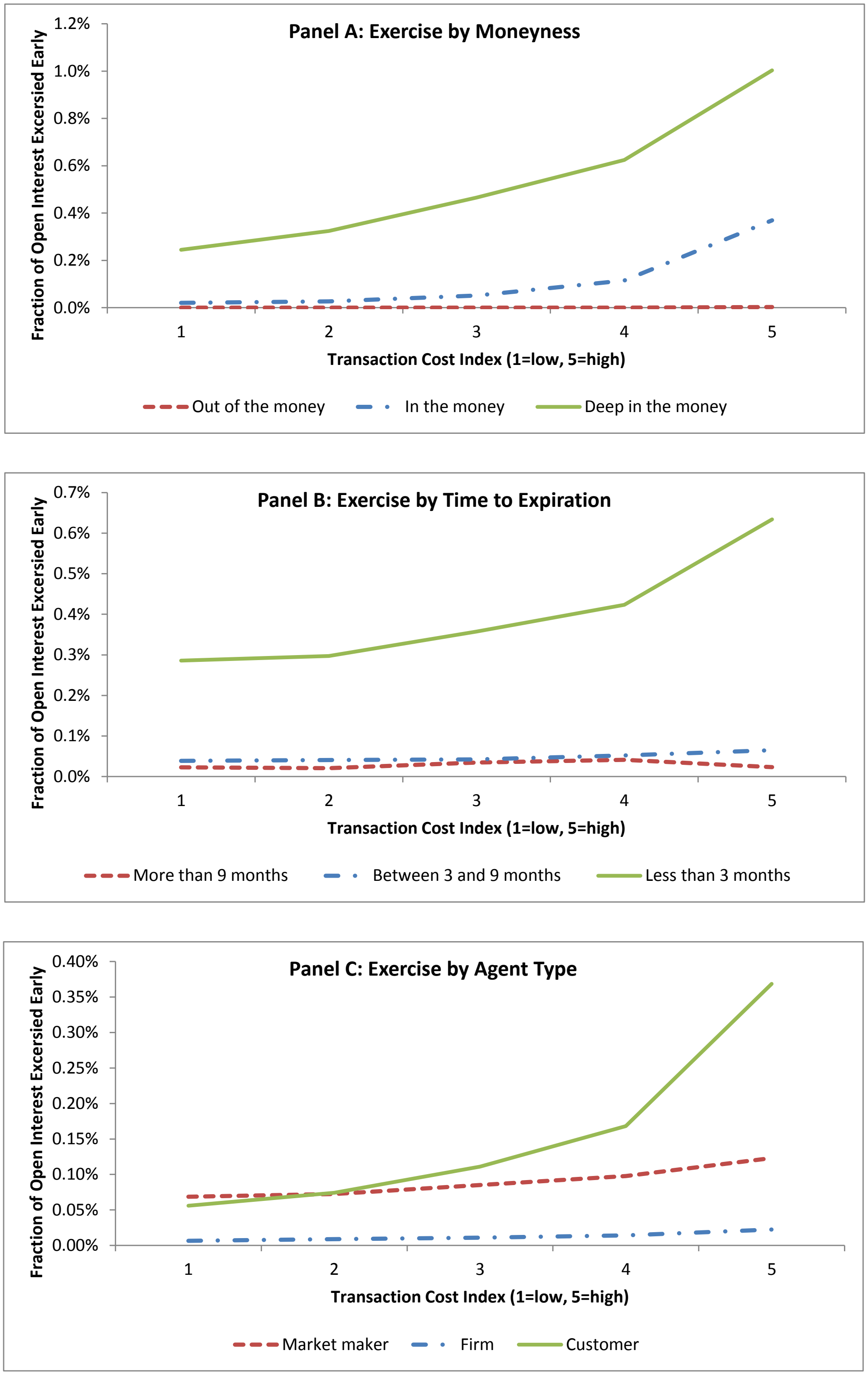


Figure 6: **Actual Early Exercise of Equity Options for Varying Transaction Costs.** This figure shows the empirical fraction of equity options exercised early as a function of transaction costs. The daily data on option series are divided in quintiles based on their transaction costs, measured each day as the bid-ask spread divided by the mid price for the corresponding at-the-money option series with the same expiration and underlying stock. Panel A groups the data by the moneyness of the option, Panel B by expiration, and Panel C by agent type. Consistent with our theory, early exercise is increasing in transaction costs, increasing in moneyness, and decreasing in time to expiration.