Abstract

We study how short-sale constraints on stock lending affect asset prices in an equilibrium model with multiple assets. We endow investors with heterogeneous beliefs in order to generate short selling demand. We obtain a CAPM-like equation that links asset-specific excess returns with the market equity premium. In the presence of short selling constraints in the market, the model gives rise to asset-specific alphas that are explained by both asset-specific and market-wide short-sale constraints; unconstrained stocks have higher risk-adjusted expected returns relative to the market portfolio, whereas the opposite holds for constrained stocks. In the absence of short-sale constraints, the model reduces to the standard CAPM. We test the model using extensive data on short interest and borrow rates. The model is able to empirically explain asset prices for 10 borrow-fee-sorted portfolios, as opposed to CAPM and factor models which produce unexplained alphas that are significantly different from zero for some low and high borrow fee portfolios.
1 Introduction

Classical asset pricing models, such as the capital asset pricing model (CAPM), assume a frictionless market with homogeneous investors. More recent theories have acknowledged that investors are heterogeneous and that efficiency of financial markets depends on the ability of investors to short sell securities believed to be overvalued. In a very influential paper, Miller (1977) has argued that in the presence of investors with heterogeneous beliefs about an asset return and with short selling constraints, the asset will be overpriced because it cannot be sold (short) by investors with pessimistic beliefs, and as a consequence the marginal investor will be an optimist. In order to sell a security short, an investor must first borrow the security from a lender. Thus, the ability to borrow a particular security depends on the ease of locating a potential lender. This does not happen with certainty due to the over-the-counter nature of the market for borrowing and lending securities, which gives rise to short selling constraints that have a potential to influence asset prices. In this paper we study, both theoretically and empirically, the effect of short selling constraints on the cross section of asset prices. On the theoretical side, we propose a model that extends the standard CAPM by incorporating short selling constraints on the lending side of the market. On the empirical side, we use an extensive database on short interest and borrow rates covering over 95% of U.S. public equities to first demonstrate that standard models, such as the CAPM and multi-factor models, underprice short selling constrained stocks (i.e. stocks with high borrow fees) and overprice non-constrained stocks (i.e. stocks with borrow fees close to zero), and then to assess the performance of our model in explaining the cross section of asset prices. We show that our model, which incorporates short selling constraints, reduces the underpricing of short selling constrained stocks by at least 50 percent: unexplained residual negative returns are reduced in absolute value by about a half, and also cease to be statistically significant. In addition, the model corrects for the overpricing of non-constrained stock which constitute an overwhelming majority of all stocks, by reducing the unexplained residual positive returns by about 50 basis points per year.

Our model builds upon a standard CAPM with multiple assets and a continuum of investors by endowing investors with heterogeneous beliefs about future asset payoffs, which can create a demand for short selling by pessimistic investors. In order to short sell an asset, an investor must first borrow the asset from another investor who holds the asset
and has not lent it yet, at which point the borrower delivers a cash collateral to the lender in addition to any borrow fees.\textsuperscript{1} In the next period the borrower returns the asset and gets back the cash collateral. We impose short selling constraints on the security lending side of the market, meaning that an investor who goes long in a particular asset can only lend up to a fixed fraction of its holdings of the asset, where the fraction depends on and represents various types of short selling constraints specific to that asset at a given point in time. This assumption, for instance, represents restrictions on lending by some investors (i.e. to avoid downward pressure on stock prices, because of counterparty risk, etc). In addition, it could proxy for the decentralized nature of the market for borrowing and lending securities: the shares offered for lending by asset holders are not always located by potential borrowers, as matching between lenders and borrowers is not instantaneous. If in an equilibrium the number of shares lent by an investor who is long in an asset is smaller than the fraction representing the constraint, that means that short selling is unconstrained as borrowers are not borrowing the entire supply of the located lendable shares and consequently the lending is costless. If the number of shares lent is such that the constraint is binding, that means that all the located shares are lent, and borrow fees increase until the market is cleared (i.e. until the borrowing demand equals the supply of shares available for lending). When the model is taken to the data, short interest can be used to identify the magnitude of the constraint for any asset with positive cost of borrowing (net of brokerage fees).

The model gives rise to two equations that jointly determine equilibrium borrow fees and asset prices. It is possible to obtain one equation that links equilibrium borrow fees and asset prices free of the variable representing the amount of disagreement among investors, which is difficult to quantify and measure. The equation is testable as all of its components are quantifiable and can be constructed from the variables in our datasets. We express the equation in the CAPM style, linking asset specific excess returns with the market equity premium. Short selling constraints give rise to an asset specific alpha that can be decomposed into two parts: first, the market wide alpha which takes into account all borrow fees and short interests in the market, and is proportional to the asset specific beta;\textsuperscript{2} and second, the asset specific alpha which is negative and decreasing in the asset specific borrow fee and short interest.\textsuperscript{3} The intuition behind the ”part 2” component is as

\textsuperscript{1}The borrow fee need not be positive since the collateral itself is valuable as it can earn riskless return.
\textsuperscript{2}We refer to this component as ”part 1”.
\textsuperscript{3}We refer to this component as ”part 2”.
follows. First, higher fees reduce short selling, but also encourage going long in the asset because the potential income from securities lending is higher. Both of these forces put an upward pressure on the asset price. Second, for a constrained asset, higher short interest indicates that the asset is easier to lend (i.e. a higher fractions of shares owned can be lent), which also encourages going long in the asset, again causing a rise in the asset price. Interestingly, the interaction between borrow fee and short interest is important: the effect of the short interest on returns is increasing in the borrow fee, and similarly the effect of the borrow fee on returns is increasing in the short interest. To our best knowledge, the effect of this interaction between borrow fees and short interest on asset prices has not been identified in the literature. The “part 1” component is more subtle and overlooked in the literature: constrained short selling leads to overvaluation of the market portfolio, implying that unconstrained stocks (with positive beta) will produce higher risk-adjusted returns than what is predicted by the standard models.

We confirm the predictions of the model in the data. We begin by sorting stocks by borrow fees and grouping them into 10 portfolios. We show that standard models overprice the highest-borrow-fee portfolios, while underpricing the unconstrained portfolios relative to the market portfolio. Next, we estimate the model-implied alphas, and show that they account for most of the mispricing produced by the standard models. Finally, we decompose the alphas into “part 1” and “part 2” and show that each plays the role implied by the model: the “part 1” alpha corrects for the overpricing of the unconstrained portfolios, while “part 2” corrects for the underpricing of the constrained portfolios.

This paper is related to several strands of literature. On the theoretical side, we build upon the capital asset pricing model, introduced in Treynor (1961), Sharpe (1964), Lintner (1965a,b) and Mossin (1966). Short selling constraints are a part of a large literature on limits to arbitrage and various institutional constraints that prevent asset prices from reaching an efficient level, started by Shleifer and Vishny (1997), and recently surveyed in Gromb and Vayanos (2010). The analysis of short selling frictions in models with investors endowed with heterogeneous beliefs is explored in Harrison and Kreps (1978), Scheinkman and Xiong (2003), Hong and Stein (2003), Hong et al. (2006), Geanakoplos (2010), and Simsek (2013), among others. Duffie et al. (2002) study the effect of the over-the-counter nature of the market for borrowing and lending securities on borrow fees. We complement the theoretical literature by proposing a model that allows for a quantitative assessment
of the effect of the short selling frictions on the lending side of the market.

On the empirical side, a majority of studies use datasets obtained from an individual participating institution in the stock loan market, and are consequently limited in terms of the coverage of the market. An exception is Drechsler and Drechsler (2014), who use the same dataset with the sample running from 2004 to 2013, but focus on the frictions on the borrowing side of the market, in the sense that some market participants are limited from short selling stocks, which implies that the shorting premium is short sellers’ compensation for the concentrated risk they bear in shorting stocks that they find overpriced. While this story may play some role in explaining the overpricing due to short selling constraints, it does not explain what gives rise to borrow fees, which are taken as exogenous in their model. Among other relevant empirical papers, D’avolio (2002) describes the market for borrowing and lending U.S. equities, Asquith et al. (2005) use the data on short interest and institutional ownership to document underperformance of short selling constrained stocks, while Cohen et al. (2007) introduce a new identification strategy to isolate shifts in the supply and demand for shorting, and show that shorting demand is an important predictor of future stock returns. Our paper complements this literature by using a more comprehensive dataset to document underperformance of borrowing constrained stocks, but also overperformance of unconstrained stocks. In addition, we build a model that largely explains these anomalies, and outlines the role of the interaction between borrow fees and short interest in affecting asset prices, which has not been emphasized in the literature.

Layout. The rest of the paper is organized as follows. Section 2 describes the market for borrowing and lending U.S. equities. Section 3 lays out the theoretical model. Section 4 describes the data sources and provides the descriptive statistics. Section 5 presents the empirical strategy and main results. Section 6 concludes.

2 Institutional Background

The market for borrowing and lending U.S. equities is an over-the-counter market in which a typical securities lending transaction features an interaction between a would-be short seller, broker and natural lender. A borrower is typically a hedge fund or investment fund that wishes to short a particular stock due to informational reasons, to hedge a long
position in its portfolio, or to meet a stock delivery obligation. A natural lender is a mutual fund, index fund, pension fund, wealth fund, or an insurer who wishes to earn extra returns from borrow fees while keeping its long position, or in order to obtain cash financing from the cash collateral posted by the borrower. A broker participates on both sides of the market by locating and managing an inventory of shares available for lending, executing transactions and managing collateral.

A typical transaction is detailed in Figure 1. A would-be borrower first contacts its broker with a request to borrow a particular security. The broker might locate the stock in its own inventory, or in the accounts of the clients who hold the stock and have permitted the use of their securities for lending. If each of the two options fails, the broker could search for another potential lender in the market. Upon finding a lender and security, the cash collateral, typically 102% of the market value of the borrowed shares (for domestic securities), is transferred from the borrower to the lender, while the shares, with all the voting rights, are passes to the borrower. While the loan is active the borrower is obliged to pass all cash dividends and cash or stock distributions to the lender, in addition to a borrow fee, quoted as a percentage of the market value of the borrowed shares, and
usually paid on a monthly basis. Because the lender can earn a risk-free return on the cash collateral posted by the borrower, the borrow fee can actually be negative, in which case the lender rebates interest on the collateral at an agreed rate. The collateral is regularly adjusted to any changes in the market value of the borrowed shares. Upon the termination of the contract, which can happen at any time and can be requested by either the borrower or lender, the shares are returned to the lender, while the cash collateral goes back to the borrower.

In the absence of short selling frictions (which are due to the inability to always locate a potential lender) all the shares are available for lending, and because there is always excess supply of available shares (because the number of shares outstanding is positive) the borrowing is be costless. Because some investors are restricted from lending and because existing potential lenders are often difficult to locate, the demand for borrowing can exceed the supply of located shares, in which case borrow fees must rise in order to clear the market. The number of shares sold short at a point in time as a fraction of the total number of shares outstanding is referred to as short interest.

3 Model

In this section we study how short selling constraints affect asset borrow rates and prices in an equilibrium model with multiple assets. We endow investors with heterogeneous beliefs about future asset payoffs, which can create a demand for short selling by pessimistic investors. Short selling constraints are imposed on the security lending side of the market. This means that an investor who goes long in a particular asset can lend up to a fraction $\phi$ of its holdings of the asset, where $\phi$ represents various types of short selling constraints specific to that asset at a given point in time. The model implies a CAPM-like equation that links specific asset excess returns with the market equity premium. Short selling constraints give rise to asset specific alphas that depend on asset specific and market-wide short-sale constraints.
3.1 Setup

Consider an overlapping-generations (OLG) economy in discrete time with a unit mass of investors of type \( j \in \{A, B\} \) born at time \( t \) with wealth \( W_{j,t}^y \) (where \( y \) stands for "young") who live for 2 periods. Investors trade risky securities \( n = 1, \ldots, N \), assumed without loss of generality to be in the unit supply. Let \( \delta_t \) denote the dividend vector, and \( P_t \) denote the ex-post price vector for risky securities at time \( t \). Let \( E_t(P_{t+1} + \delta_{t+1}) = \mu_t \) and \( \text{Var}_t(P_{t+1} + \delta_{t+1}) = \Sigma_t \). Investors of type \( A \) believe that \( E_{A,t}(P_{t+1} + \delta_{t+1}) = \mu_t + \epsilon_t \) and investors of type \( B \) believe that \( E_{B,t}(P_{t+1} + \delta_{t+1}) = \mu_t - \epsilon_t \) here \( \epsilon_t \in \mathbb{R}^N \). Both types of investors agree on \( \text{Var}_t(P_{t+1} + \delta_{t+1}) \): \( \text{Var}_{A,t}(P_{t+1} + \delta_{t+1}) = \text{Var}_{B,t}(P_{t+1} + \delta_{t+1}) = \Sigma_t \). Thus, \( |\epsilon_t^{(n)}| \) denotes the amount of disagreement about asset \( n \) at time \( t \).

There is a market for borrowing and lending risky securities. Let \( f_t \) denote the vector of borrow rates at time \( t \), meaning that for each unit of asset \( n \) borrowed, the borrower pays fraction \( f_t^{(n)} \) of the price \( P_t^{(n)} \) to the lender at time \( t \). In addition, the borrower puts \( mP_t^{(n)} \) amount of cash as a collateral for every unit of stock \( n \) borrowed.\(^4\) We further assume that each investor who goes long in asset \( n \) at time \( t \) can lend to up to \( \phi_t^{(n)} \) units of asset \( n \) held, were \( \phi_t^{(n)} \) represents various types of short selling constraints specific to asset \( n \) at time \( t \). In addition to the risky securities, there is a one-period risk-free asset in infinite supply, a unit of which costs \( Q_t \) at time \( t \) and pays 1 unit of wealth at time \( t + 1 \).

A young investor of type \( j \in \{A, B\} \) at time \( t \) chooses a portfolio consisting of shares of risky assets, \( x_j \), risk-free bonds, \( b_j \), and shares to lend, \( z_j \), in order to maximize the mean-variance utility of its time \( t + 1 \) wealth \( W_{j,t+1} \), taking prices and fees as given. In particular,

\[
\max_{x_j, b_j, z_j} E_{j,t}[W_{j,t+1}] - \frac{\gamma}{2} \text{Var}_t[W_{j,t+1}] \tag{1}
\]

s.t.

\[
W_{j,t+1} = x_j'(P_{t+1} + \delta_{t+1}) + b_j - m \min (z_j, x_j)' P_t \tag{2}
\]

\[
W_{j,t+1}^y = x_j' P_t + b_j Q_t - \min (z_j, x_j)' \text{diag} [m + f_t^{(n)}] P_t \tag{3}
\]

\[
0 \leq z_j^{(n)} \leq \max \left( 0, \phi_t^{(n)} x_j^{(n)} \right), \text{ for } n = 1, \ldots, N \tag{4}
\]

Equation (2) gives the expression for the wealth of an old investor of type \( j \) at time \( t + 1 \) as a function of the investor’s portfolio holding, while Equation (3) is the budget constraint of

\(^4\)In practice, \( m = 1.02 \).
the same investor at time $t$. The first term on the right-hand side of each equation is due to
the holdings of risky assets, the second term is due to the position in risk-free bonds, while
the third term is due to the collateral. To understand the third term, first suppose that
a young investor of type $j$ at time $t$ shorts asset $n$, i.e. $x_{j}^{(n)} < 0$. Then the investor must
pay the total fee resulting from such a position in asset $n$ in the amount of $x_{j}^{(n)} f_{t}^{(n)} P_{t}^{(n)}$, in
addition to the collateral in the amount of $x_{j}^{(n)} m P_{t}^{(n)}$. Each of the two terms reduces
the amount of time $t$ wealth spent on the available assets, explaining the term in Equation
(3). Furthermore, the investor receives the collateral back at time $t + 1$, which explains
the term in Equation (2). Note that in that case $z_{j}^{(n)} = 0$ since the investor cannot lend
the asset that he does not own. If the investor on the other hand goes long in asset $n$, i.e.
$x_{j}^{(n)} > 0$, then he may lend some of the shares. In that case $z_{j}^{(n)} > 0$, and the cash flows
due to the collateral go in the opposite direction from the first case. Equation (4) is the
collateral constraint: an investor who goes long in asset $n$, i.e. $x_{j}^{(n)} > 0$, cannot lend more
than fraction $\phi_{t}^{(n)}$ of his holdings of asset $n$.

A competitive equilibrium consists of a sequence of prices $P_{t}$ and $Q_{t}$, fees $f_{t}$ and alloca-
tions $(x_{j,t}, b_{j,t}, z_{j,t})$ for $j \in \{A, B\}$ such that, given prices and fees, allocation $(x_{j,t}, b_{j,t}, z_{j,t})$
solves the problem of a young investor of type $j \in \{A, B\}$ at time $t$, and markets clear:

$$\frac{1}{2} x_{A,t} + \frac{1}{2} x_{B,t} = 1$$

(5)

$$\min(z_{A,t}, x_{A,t}) + \min(z_{B,t}, x_{B,t}) = 0$$

(6)

Equation (5) is the market clearing condition for the market for buying and selling risky as-
sets, while Equation (6) is an equivalent condition for the market for lending and borrowing
risky assets.

3.2 Results

Equilibrium Characterization

After substituting out $b_{j}$ and $W_{j,t+1}$, the problem of young investor of type $j$ at time $t$
simplifies to

\[
\max_{\mathbf{x}_j, z_j} \mathbf{x}_j' \left[ \mathbb{E}_{j,t} (\mathbf{P}_{t+1} + \delta_{t+1}) - R_{t+1}^f \mathbf{P}_t \right] - \frac{\gamma}{2} \mathbf{x}_j' \Sigma \mathbf{x}_j
\]

\[
+ R_{t+1}^f (z_j + \min (0, \mathbf{x}_j))' \begin{bmatrix} 1 - \frac{1}{R_{t+1}^f} \end{bmatrix} m + f_t^{(n)} \mathbf{P}_t
\]

s.t. \( 0 \leq z_j^{(n)} \leq \max \left( 0, \phi_t^{(n)} x_j^{(n)} \right) \)

where \( R_{t+1}^f \equiv \frac{1}{\hat{Q}_t} \) is the risk-free return from period \( t \) to \( t + 1 \). The first-order conditions with respect \( \mathbf{x}_j \) and \( z_j \) are:

\[
\mathbf{x}_j : \quad 0 = \mathbb{E}_{j,t} (\mathbf{P}_{t+1} + \delta_{t+1}) - R_{t+1}^f \mathbf{P}_t - \gamma \Sigma \mathbf{x}_j + \text{diag} \left[ \mathbf{1} \left( x_j^{(n)} \geq 0 \right) \phi_t^{(n)} \right] \lambda_{j,t}
\]

\[
+ R_{t+1}^f \text{diag} \left[ \mathbf{1} \left( x_j^{(n)} < 0 \right) \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t^{(n)} \right] \mathbf{P}_t
\]

\[
\mathbf{z}_j : \quad 0 = R_{t+1}^f \text{diag} \left[ \mathbf{1} \left( x_j^{(n)} \geq 0 \right) \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t^{(n)} \right] \mathbf{P}_t - \lambda_{j,t} + \eta_{j,t}
\]

where \( \lambda_{j,t} \) and \( \eta_{j,t} \) are vectors of Lagrange multipliers on the collateral and non-negativity lending constraints, respectively. Observe that if \( \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t^{(n)} = 0 \), then Equation (8) implies that \( \lambda_{j,t}^{(n)} = 0 \) and \( \eta_{j,t}^{(n)} = 0 \) and the last two terms on the right-hand side of the \( n \)th row of Equation (7) drop out. If \( \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t^{(n)} < 0 \), then Equation (8) implies that \( \lambda_{j,t}^{(n)} = 0 \) and \( \eta_{j,t}^{(n)} > 0 \), implying \( z_j^{(n)} = 0 \). Thus, in the equilibrium nobody can short asset \( n \), so the last two terms on the right-hand side of the \( n \)th row of Equation (7) drop out. Thus, the last two terms on the right-hand side of Equation (7) matter if and only if \( \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t^{(n)} > 0 \), i.e. \( f_t^{(n)} > -m \left( 1 - \frac{1}{R_{t+1}^f} \right) \). In that case, some agents must short asset \( n \) in the equilibrium because \( z_j^{(n)} > 0 \) whenever \( x_j^{(n)} > 0 \). Note that in that case we can link \( \phi_t^{(n)} \) to the short interest for asset \( n \). In particular, if \( f_t^{(n)} > -m \left( 1 - \frac{1}{R_{t+1}^f} \right) \), let \( j \) denote the type of investor who shorts asset \( n \) in the equilibrium. Then

\[
\phi_t^{(n)} = \frac{x_{t,j}^{(n)}}{1 - \frac{1}{2} x_{t,j}^{(n)}} = \frac{x_{t,j}^{(n)}}{1 - \frac{1}{2} x_{t,j}^{(n)}} = \frac{s_t^{(n)}}{1 + s_t^{(n)}}
\]

where \( s_t^{(n)} \equiv -\frac{1}{2} x_{t,j}^{(n)} \) is the short interest for asset \( n \).

Let \( \iota_t^{(n)} = 1 \) if \( f_t^{(n)} > -m \left( 1 - \frac{1}{R_{t+1}^f} \right) \) (i.e. if short selling constraint on asset \( n \) binds), and \( \iota_t^{(n)} = 0 \) otherwise. Using Equation (8) to substitute the multipliers out in Equation
we can without loss of generality assume that inequality \( \lambda_{j,t}^{(n)} > 0 \), then \( \eta_{j,t}^{(n)} = 0 \), we have that

\[
0 = \mathbb{E}_{j,t}(\mathbf{P}_{t+1} + \delta_{t+1}) - R_{t+1}^{f} \sum_{j} \mathbf{1} - \gamma \sum_{j} \mathbf{x}_j \\
+ R_{t+1}^{f} \text{diag} \left[ \left( 1 - \left( 1 - \phi_{t}^{(n)} \right) \mathbf{1} \left( x_{j}^{(n)} \geq 0 \right) \right) \left( 1 - \frac{1}{R_{t+1}^{f}} \right) m + f_{t}^{(n)} \right] \mathbf{P}_{t}
\]

Taking the difference and sum of Equation (9) between the two types of investors, we obtain the following equations, respectively:

\[
0 = \mathbb{E}_{t}(\mathbf{P}_{t+1} + \delta_{t+1}) - \gamma \sum_{j} \mathbf{1} - R_{t+1}^{f} \text{diag} \left[ \left( 1 - \phi_{t}^{(n)} \right) \left( 2 \cdot \mathbf{1} \left( x_{A}^{(n)} \geq 0 \right) - 1 \right) \left( 1 - \frac{1}{R_{t+1}^{f}} \right) m + f_{t}^{(n)} \right] \mathbf{P}_{t}
\]

\[
0 = \mathbb{E}_{t}(\mathbf{P}_{t+1} + \delta_{t+1}) - \gamma \sum_{j} \mathbf{1} - R_{t+1}^{f} \text{diag} \left[ \left( 1 + \phi_{t}^{(n)} \right) \left( 1 - \frac{1}{R_{t+1}^{f}} \right) m + f_{t}^{(n)} \right] \mathbf{P}_{t}
\]

where in the last equation we applied the equilibrium condition \( \frac{1}{2} (x_{A} + x_{B}) = 1 \). Observe also that if \( \xi_{t}^{(n)} > 0 \), then \( \frac{1}{2} (x_{A}^{(n)} - x_{B}^{(n)}) = \frac{1 + \phi_{t}^{(n)}}{1 - \phi_{t}^{(n)}} \left( 2 \cdot \mathbf{1} (x_{A}^{(n)} \geq 0) - 1 \right) \). If \( \xi_{t}^{(n)} = 0 \), then we can without loss of generality assume that inequality \( f_{t}^{(n)} \leq - \left( 1 - \frac{1}{R_{t+1}^{f}} \right) m \) in fact holds with equality (since the last term in the two equations above drops out anyways).

Defining \( \hat{x}_{t} \equiv \frac{1}{2} (x_{A,t} - x_{B,t}) \), we summarize the equilibrium conditions for the prices \( \mathbf{P}_{t} \) and borrow rates \( f_{t} \) in the following theorem:

**Theorem 1.** The equilibrium prices \( \mathbf{P}_{t} \) and borrow rates \( f_{t} \) satisfy the following set of conditions:

\[
\mathbb{E}_{t}(\mathbf{P}_{t+1} + \delta_{t+1}) - R_{t+1}^{f} \mathbf{P}_{t} = \gamma \sum_{j} \mathbf{1} - \frac{1}{2} R_{t+1}^{f} \text{diag} \left[ \left( 1 + \phi_{t}^{(n)} \right) \left( 1 - \frac{1}{R_{t+1}^{f}} \right) m + f_{t}^{(n)} \right] \mathbf{P}_{t} \quad (10)
\]

\[
\epsilon_{t} = \gamma \sum_{j} \hat{x}_{t} \quad (11)
\]

\[
f_{t}^{(n)} \geq - \left( 1 - \frac{1}{R_{t+1}^{f}} \right) m, \quad \text{for } n = 1, ..., N \quad (12)
\]

\[
- \frac{1 + \phi_{t}^{(n)}}{1 - \phi_{t}^{(n)}} \leq \hat{x}_{t}^{(n)} \leq \frac{1 + \phi_{t}^{(n)}}{1 - \phi_{t}^{(n)}}, \quad \text{for } n = 1, ..., N \quad (13)
\]

\[
\hat{x}_{t}^{(n)} = \frac{1 + \phi_{t}^{(n)}}{1 - \phi_{t}^{(n)}} \left( 2 \cdot \mathbf{1} (y_{A}^{(n)} \geq 0) - 1 \right), \quad \text{if } f_{t}^{(n)} > - \left( 1 - \frac{1}{R_{t+1}^{f}} \right) m, \quad \text{for } n = 1, ..., N \quad (14)
\]

Note that for an asset \( n \) for which short selling constraint does not bind, i.e. \( f_{t}^{(n)} = \)
\[
-E_t \left( P_{t+1}^{(n)} + \delta_{t+1}^{(n)} \right) - R_{t+1}^f P_t^{(n)} = \gamma \Sigma_t^{(n)} 1
\]

where \( \Sigma_t^{(n)} \) is the \( n^{th} \) row of the covariance matrix \( \Sigma_t \).

Modified CAPM

Equation (10) in Theorem 1 provides a relationship between asset prices and borrow rates. Importantly, the disagreement vector \( \epsilon_t \) does not enter Equation (10), but only Equation (11). Consequently, all the variables in Equation (10) are observable in the data, which allows for empirical testing. Focusing on asset \( n \) in Equation (10), we have that

\[
E_t \left( P_{t+1}^{(n)} + \delta_{t+1}^{(n)} \right) - R_{t+1}^f P_t^{(n)} = \gamma \text{Cov}_t \left( P_{t+1}^{(n)} + \delta_{t+1}^{(n)} \right) - \frac{1}{2} R_{t+1}^f \left( 1 + \phi_t^{(n)} \right) \left( 1 - \frac{1}{R_{t+1}^m} \right) m + f_t^{(n)} P_t^{(n)}
\]

and dividing by \( P_t^{(n)} \), we get

\[
E_t \left( R_{t+1}^{(n)} \right) - R_{t+1}^f = \gamma \text{Cov}_t \left( R_{t+1}^{(n)}, R_{t+1}^m \right) P_t^1 - \frac{1}{2} R_{t+1}^f \left( 1 + \phi_t^{(n)} \right) \left( 1 - \frac{1}{R_{t+1}^m} \right) m + f_t^{(n)}
\]

where \( R_{t+1}^{(n)} \equiv \frac{P_{t+1}^{(n)} + \delta_{t+1}^{(n)}}{P_t^{(n)}} \) and \( R_{t+1}^m \equiv \frac{(P_{t+1} + \delta_{t+1})^1}{P_t^1} \) are returns on asset \( n \) and on the market portfolio of risky assets. Multiplying Equation (15) by market weight \( w_t^{(n)} \equiv \frac{P_t^{(n)}}{P_t^m} \), and summing over the equations for all \( N \) risky assets, we obtain the following equation

\[
E_t (R_{t+1}^m) - R_{t+1}^f = \gamma \text{Var}_t (R_{t+1}^m) P_t^1 - \frac{1}{2} R_{t+1}^f \left[ 1 + \phi_t^{(n)} \right] \left[ 1 - \frac{1}{R_{t+1}^m} \right] m + f_t^m
\]

Combining Equations (15) and (16), and defining \( r_{t+1}^{(n)} \equiv R_{t+1}^{(n)} - 1 \), \( r_{t+1}^m \equiv R_{t+1}^m - 1 \), and \( r_{t+1}^f \equiv R_{t+1}^f - 1 \), we obtain the following result:

**Theorem 2.** In any equilibrium, we have that

\[
E_t (r_{t+1}^{(n)}) - r_{t+1}^f = \alpha_t^{(n)} + \beta_t^{(n)} \left( E (r_{t+1}^m) - r_{t+1}^f \right)
\]
where

\[
\alpha_t^{(n)} \equiv \frac{1}{2} \left( 1 + r_{t+1}^f \right) \left[ \beta_t^{(n)} \mathbf{w}_t^f \text{diag} \left[ 1 + \phi_t^{(n)} \right] \left( 1 - \frac{1}{1 + r_{t+1}^f} \right) m + f_t \right] - \left( 1 + \phi_t^{(n)} \right) \left( \left( 1 - \frac{1}{1 + r_{t+1}^f} \right) m + f_t^{(n)} \right)
\]

\[
\beta_t^{(n)} \equiv \frac{\text{Cov}_t \left( r_{t+1}^{(n)}, r_{t+1}^m \right)}{\text{Var}_t \left( r_{t+1}^m \right)}
\]

If short selling in unconstrained for all the assets at time \( t \), \( \alpha_t \) will be \( 0 \), and we get the standard CAPM. Otherwise, \( \alpha_t \neq 0 \) in general. Note that \( \alpha_t^{(n)} \) for asset \( n \) depends on the short interest and borrow rates of other assets. \( \alpha_t \) is the sum of two components. We first explain the intuition behind the second component, \( \alpha_{2,t}^{(n)} \equiv -\frac{R_{t+1}^f}{2} \left( 1 + \phi_t^{(n)} \right) \left( \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t^{(n)} \right), \) referred to as ”part 1”. It represents the direct effect of short sale constraints on asset \( n \) and has a negative effect on \( \alpha_t^{(n)} \). All else equal, the higher the borrow rate \( f_t^{(n)} \), the more negative this term is. This is intuitive: the higher the borrow rate, the more attractive it is to buy the asset, and less attractive it is to sell, which pushes up the price and reduces the expected return. More interestingly, for any borrow rate \( f_t^{(n)} \) in excess of \(-\left( 1 - \frac{1}{R_{t+1}^f} \right) m, \) this term decreases when the short selling constraint gets relaxed (i.e. when \( \phi_t^{(n)} \) goes up). Intuitively, when \( \phi_t^{(n)} \) goes up and short selling for asset \( n \) is constrained, long investors find the asset more attractive because they can earn more by lending it. If the borrow rate \( f_t^{(n)} \) is not allowed to change, short investors have no desire to change their short position. To keep the market in the equilibrium, price \( P_t^{(n)} \) needs to increase to discourage long investors from demanding and lending too much, and to encourage short investors to absorb the excess lending supply. In the equilibrium, \( f_t^{(n)} \) will of course respond as well, and the price \( P_t^{(n)} \) could move in either direction. To better understand the effect of short and long investors on the return, observe that this component can be written as the (negative of the) sum of two parts: \( \frac{R_{t+1}^f}{2} \left( \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t \right) \) and \( \frac{R_{t+1}^f}{2} \phi_t^{(n)} \left( \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t^{(n)} \right) \). The first part is the cost of short selling a dollar worth of asset \( n \), weighted by the mass of short sellers in that asset, \( \frac{1}{2} \), and the second part is the gain of lending the maximum fraction of a dollar worth of asset \( n \) held, weighted by the mass of long investors in that asset, \( \frac{1}{2} \).

The first component in \( \alpha_t^{(n)} \), \( \alpha_{1,t}^{(n)} \equiv \frac{1}{2} R_{t+1}^f \beta_t^{(n)} \mathbf{w}_t^f \text{diag} \left[ 1 + \phi_t^{(n)} \right] \left( 1 - \frac{1}{R_{t+1}^f} \right) m + f_t \right), \) or ”part 1”, comes from the fact that the market portfolio is ”overvalued” due to short selling constraints. Note that this term depends on a specific asset only through the asset beta. If \( \beta_t^{(n)} > 0 \), then this term has a positive effect on \( \alpha_t^{(n)} \): returns for asset \( n \) and
the market correlate positively, so if the expected market return goes down (due to short selling constraints), then the expected return for asset \( n \) goes up relative to the market. The effect is opposite if \( \beta_t^{(n)} < 0 \).

4 Data

4.1 Data Sources

We obtain data related to short selling-borrow fees and short interest from Markit Securities Finance. Markit collects self-reported daily data on stock loan availability and stock lending transactions from prime brokers, custodians, asset managers and hedge funds, thus covering both sides of the market. Their proprietary database is the most comprehensive database on short selling. We believe the accuracy of the data is not a reason for concern, since players have limited incentives to report false information due to Markit’s wide coverage and the anonymity of the data. Markit offers the borrowing and lending samples separately to different types of users, thus further protecting the anonymity of the data. We have access to the borrowing sample, so borrow fee data (based on transactions reported by borrowers) includes the brokers’ spread. Markit reports the number of shares borrowed both by the borrower sample (that is tied to the reported borrow fees) and by the entire market. We assume the fee data is representative of the entire market and use the latter short interest measure to calculate short interest ratio as the number of shares sold short as a percentage of shares outstanding.

We merge the short selling data with the CRSP database on security prices and returns (from Wharton Research Data Services, WRDS) by CUSIP. We keep only stocks classified as ”common stock” (CRSP codes 10 and 11) trading on the main US exchanges (NYSE, NASDAQ, and AMEX, exchange codes 11, 14, and 12, respectively), resulting in a sample of over 4,000 stocks. We calculate total returns and number of shares outstanding for each stock using dividend and stock split adjusted prices and drop any weeks (for a given stock) where the stock did not have any trades in three or more regular trading days.

Finally, we obtain two-year CAPM, FF3, and FFC4 weekly beta estimates for each stock (which will be inputs for our alpha calculations) from Beta Suite by WRDS, which we merge to CRSP by PERMNO, and the relevant time series of risk-free rates, market
returns, and other FF3 and FFC4 factors from Kenneth French’s website. Although all of
our data is available on a daily frequency, we choose a weekly frequency instead in order to
reduce some of the noise in stock prices. Our sample runs from January 4, 2010 to August
31, 2015.

### 4.2 Summary Statistics

High borrow fees are relatively uncommon: on an average day, most stocks have a borrow
fee of zero plus a small brokers’ spread of 0.25-0.375% per year, and 80% of stocks have a
borrow fee (including spread) of less than 1%, also known as ”general collateral”. Only 5% of
stocks have a borrow fee greater than 10%, and only 1% have a fee greater than 30%.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Min fee (% yr)</th>
<th>Avg fee (% yr)</th>
<th>Max fee (% yr)</th>
<th>SIR (%)</th>
<th>Return (med % wk)</th>
<th>Market cap ($ in million)</th>
<th>n stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>1.5</td>
<td>.33</td>
<td>6560</td>
<td>2470</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>0.6</td>
<td>1.0</td>
<td>5.9</td>
<td>.69</td>
<td>1467</td>
<td>329</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>1.4</td>
<td>2.0</td>
<td>5.8</td>
<td>.37</td>
<td>894</td>
<td>155</td>
</tr>
<tr>
<td>4</td>
<td>2.0</td>
<td>2.5</td>
<td>3.0</td>
<td>6.8</td>
<td>.37</td>
<td>749</td>
<td>63</td>
</tr>
<tr>
<td>5</td>
<td>3.0</td>
<td>3.8</td>
<td>5.0</td>
<td>7.5</td>
<td>.47</td>
<td>663</td>
<td>92</td>
</tr>
<tr>
<td>6</td>
<td>5.0</td>
<td>5.9</td>
<td>7.0</td>
<td>8.1</td>
<td>.16</td>
<td>536</td>
<td>44</td>
</tr>
<tr>
<td>7</td>
<td>7.0</td>
<td>8.3</td>
<td>10.0</td>
<td>9.0</td>
<td>.27</td>
<td>458</td>
<td>65</td>
</tr>
<tr>
<td>8</td>
<td>10.0</td>
<td>13.6</td>
<td>20.0</td>
<td>10.9</td>
<td>-.04</td>
<td>498</td>
<td>65</td>
</tr>
<tr>
<td>9</td>
<td>20.0</td>
<td>24.8</td>
<td>30.0</td>
<td>10.1</td>
<td>-.15</td>
<td>474</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>30.0</td>
<td>46.2</td>
<td>216.7</td>
<td>8.9</td>
<td>-.01</td>
<td>406</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics

Notes. The sample runs from January 4, 2010 to August 31, 2015. Borrow fees are expressed in terms of
percent of the asset price, charged on an annual basis. SIR stands for short interest, which is the number
of shares sold short as a percentage of the number of shares outstanding. Return is the median weekly
return expressed in terms of percent. Market cap is the average market capitalization of a company in the
basket. n stocks stands for the average number of companies included in the basket.

Table 1 presents summary information on the short interest ratio, median return, and
average market cap by borrow fee bin, as well as the average number of stocks in each
bin through time. We chose the bin cutoffs arbitrarily using different bin widths with
the goal of grouping stocks with comparable short selling characteristics at the expense of
equal bin sizes. For example, a pair of stocks with fees of 21% and 29% are viewed by
market participants as much more similar than stocks with fees of 1% and 9%. If we had
instead separated our sample into deciles, all relatively constrained stocks would have been
grouped in the same bin, which would have not allowed us to properly test our model. Our results are not sensitive to different choices of bin cutoffs.

The average market cap of high borrow fee stocks is an order of magnitude lower than that of low borrow fee stocks. This is expected, as large cap stocks have a large analyst following and represent the bulk of institutional investments, so there is a large number of research articles published about them, helping the market to come close to having a consensus view on the firms’ prospects and thus reducing disagreement and short selling demand, while also making it easier for brokers to find lendable shares, thus increasing short selling supply; on the other hand, small cap stocks have a lower following and are thus more susceptible to differences of opinion among investors as well as both intentional and unintentional inaccuracies in their financial statements, which increases short selling demand, while their more fragmented and lower institutional ownership makes it harder for brokers to find lendable shares, thus decreasing short selling supply. Stocks with high borrow fees also have negative median gross (i.e. before lending fees) weekly returns in our sample. The short interest ratio is increasing in the borrow fees for stocks with borrow fees below 20%; after that, it is decreasing. This is intuitive, as short interest is an equilibrium variable: the most highly constrained stocks are the ones with the highest fees, but one of the reasons they are constrained is their limited availability of lending supply, which mechanically results in a lower short interest ratio.

5 Empirical Analysis

5.1 Methodology

We follow the extensive anomalies literature in finance in testing an asset pricing model’s performance with respect to a specific variable by first sorting stocks according to the variable and grouping them into portfolios and then fitting the model to historical data on stock returns and thus testing whether the model can explain the portfolios’ returns or there is an unexplained residual alpha (constant term in a regression framework) that is statistically significantly different from zero. The baseline models are CAPM and the Fama-French three-factor (FF3) and Fama-French-Carhart four-factor (FFC4) models. Thus, if an asset
pricing model proposes $E(r^{(n)}) = f(X)$ for an asset $n$.\(^5\) After sorting and assigning stocks to portfolios, we test whether the residual $\alpha^{(n)}_{\text{res}}$ is equal to 0 in the following regressions:

$$E(r^{(n)}) - f(X) = \alpha^{(n)}_{\text{res}} + \epsilon^{(n)} \quad (18)$$

If some $\alpha^{(n)}_{\text{res}} \neq 0$ (generally, but not necessarily, for one or more of the extreme portfolios), the model is considered incapable of explaining asset prices for stocks with a particular characteristic, which is then considered a market anomaly. For instance, in the case of a CAMP test, $X = (r^f, r^m, \beta^{(n)})$, and $f(X) = r^f + \beta^{(n)}(r^m - r^f)$, where $\beta^{(n)} = \frac{\text{Cov}(r^{(n)}, r^m)}{\text{Var}(r^m)}$, as stated in Theorem 2.

We start by sorting stocks by borrow fees and grouping them into the portfolios detailed in Table 1, where portfolio 1 consists of stocks that are easy (and essentially free) to short, with borrow fees below 0.5% per year, and portfolio 10 consists of stocks that are hard and costly to short, with borrow fees above 30% per year. As documented before,\(^6\) standard models such as CAPM, FF3, and FFC4 fail to explain returns for high borrow fee stocks, with negative residual alphas that are different from zero at standard significance levels. In addition, we find that low borrow fee stocks have positive residual alphas that are small but still economically significant.

Next, we test whether the proposed model that incorporates short selling frictions is able to empirically explain asset prices where standard models fail. We first calculate the proposed model’s ”short alpha”, $\alpha^{(n)}$ for asset $n$, as defined in Theorem 2 (i.e. the alpha predicted by the model as a result of short selling constraints), using the data from Markit and different beta estimates. Next, we perform the test outlined in Equation (18), with $X = (r^f, r^m, \alpha^{(n)}, \beta^{(n)})$, and $f(X) = \alpha^{(n)} + r^f + \beta^{(n)}(r^m - r^f)$. Finally, we compare the residual alphas resulting from this estimation to the residual alphas from the baseline model (i.e. CAPM).

Finally, we separately study the effect of each of the components of $\alpha$, $\alpha_1$ (i.e. part 1) and $\alpha_2$ (i.e. part 2). We do this because the two components play different roles in the model. In particular, part 1 results from the market portfolio being overvalued, which implies that the unconstrained stocks will produce higher\(^7\) risk-adjusted returns in the

\(^{5}\)where $X$ is the vector of variables specific to a particular test

\(^{6}\)See, for example, Drechsler and Drechsler (2014).

\(^{7}\)for any positive beta stock
future relative to the market portfolio (i.e. implying that standard models overprice individual stocks on average). Part 2, the negative and much larger (in absolute value) term, comes from the fact that constrained stocks underperform their CAPM (or FF3/FFC4) risk adjusted expected returns. Thus, part 1 should correct for the overpricing of all stocks, whereas part 2 should correct for the underpricing of high borrow fee stocks. To perform the test for part i, for \( i \in \{1, 2\} \), in Equation (18) we take \( X = (r^f, r^m, \alpha^{(n)}_i, \beta^{(n)}_i) \) and \( f(X) = \alpha^{(n)}_i + r^f + \beta^{(n)}_i(r^m - r^f) \).

5.2 Results

Table 2 shows residual alphas for various models for portfolios sorted by borrow fees, as detailed in Table 1. Columns (1)-(3) provide residual alphas from the capital asset pricing model (CAPM), the Fama-French three-factor (FF3) and Fama-French-Carhart four-factor (FFC4) models, respectively. Column (4) shows the residual alpha from the model proposed in the paper, referred to as "modified CAPM". Columns (5) and (6) provide residual alphas from equivalent modifications of FF3 and FFC4 models. Since we looks at returns at weekly frequencies, residual alphas are computed with respect to weekly returns and expressed in percentage terms.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>CAPM (1)</th>
<th>FF3 (2)</th>
<th>FFC4 (3)</th>
<th>Mod CAPM (4)</th>
<th>Mod FF3 (5)</th>
<th>Mod FFC4 (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.08</td>
<td>0.10</td>
<td>0.11*</td>
<td>0.08</td>
<td>0.12*</td>
<td>0.12*</td>
</tr>
<tr>
<td>3</td>
<td>0.04</td>
<td>0.09</td>
<td>0.11</td>
<td>0.05</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>4</td>
<td>-0.07</td>
<td>-0.01</td>
<td>-0.00</td>
<td>-0.05</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>5</td>
<td>-0.21**</td>
<td>-0.18*</td>
<td>-0.16*</td>
<td>-0.17*</td>
<td>-0.13</td>
<td>-0.12</td>
</tr>
<tr>
<td>6</td>
<td>-0.24*</td>
<td>-0.21</td>
<td>-0.17</td>
<td>-0.19</td>
<td>-0.14</td>
<td>-0.14</td>
</tr>
<tr>
<td>7</td>
<td>-0.22*</td>
<td>-0.19*</td>
<td>-0.18</td>
<td>-0.14</td>
<td>-0.10</td>
<td>-0.09</td>
</tr>
<tr>
<td>8</td>
<td>-0.26**</td>
<td>-0.23*</td>
<td>-0.22*</td>
<td>-0.12</td>
<td>-0.08</td>
<td>-0.08</td>
</tr>
<tr>
<td>9</td>
<td>-0.60***</td>
<td>-0.58***</td>
<td>-0.53**</td>
<td>-0.35</td>
<td>-0.29</td>
<td>-0.28</td>
</tr>
<tr>
<td>10</td>
<td>-0.26</td>
<td>-0.18</td>
<td>-0.15</td>
<td>0.21</td>
<td>0.30</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 2: Residual Alphas

Notes. Residual alphas are computed with respect to weekly returns and expressed in percentage terms. Columns (1)-(3) provide residual alphas for CAPM, the Fama-French three-factor (FF3) and Fama-French-Carhart four-factor (FFC4) models, respectively. Column (4) shows the residual alpha from the model proposed in the paper, referred to as "modified CAPM". Columns (5) and (6) provide residual alphas from equivalent modifications of FF3 and FFC4 models. ***, **, * indicate significance at 1, 5, and 10%.
It can be seen from columns (1)-(3) that standard models overpredict the returns of high fee portfolios (i.e. portfolios 5-10), with statistically significant and economically large estimates. In addition, the standard models underpredict the performance of low fee portfolios, although the magnitudes are substantially smaller than in the case of large fee portfolios. Including the model-implied alphas substantially reduces underpricing of high fee portfolios: the estimates cease to be statistically significant, and the magnitudes fall in absolute value by about 50% (i.e. for portfolios 7, 8 and 9). The modified CAPM does not perform substantially differently from the modified FF3 and FFC4 models.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>CAPM</th>
<th>CAPM+alpha1</th>
<th>CAPM+alpha2</th>
<th>Mod CAPM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>1</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02*</td>
<td>0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.08</td>
<td>0.07</td>
<td>0.09</td>
<td>0.08</td>
</tr>
<tr>
<td>3</td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>-0.07</td>
<td>-0.07</td>
<td>-0.04</td>
<td>-0.05</td>
</tr>
<tr>
<td>5</td>
<td>-0.21**</td>
<td>-0.21**</td>
<td>-0.17*</td>
<td>-0.17*</td>
</tr>
<tr>
<td>6</td>
<td>-0.24*</td>
<td>-0.25*</td>
<td>-0.18</td>
<td>-0.19</td>
</tr>
<tr>
<td>7</td>
<td>-0.22*</td>
<td>-0.23**</td>
<td>-0.14</td>
<td>-0.14</td>
</tr>
<tr>
<td>8</td>
<td>-0.26**</td>
<td>-0.27**</td>
<td>-0.12</td>
<td>-0.12</td>
</tr>
<tr>
<td>9</td>
<td>-0.60***</td>
<td>-0.61***</td>
<td>-0.34</td>
<td>-0.35</td>
</tr>
<tr>
<td>10</td>
<td>-0.26</td>
<td>-0.27</td>
<td>0.22</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Table 3: Residual Alphas due to Decomposition of Model-Implied Alpha

Notes. Residual alphas are computed with respect to weekly returns and expressed in percentage terms. Column (1) provides residual alphas for CAPM. Columns (2) and (3) modify the CAPM by including part 1 and part 2 of the model-implied alpha, respectively. Column (4) shows the residual alpha from the model proposed in the paper. ***, **, * indicate significance at 1, 5, and 10%.

Table 3 shows residual alphas after applying separately each of the components of the model-implied alpha, $\alpha_1$ and $\alpha_2$ (i.e. columns (2) and (3), respectively), to the CAPM. For comparison, included are also residual alphas from the CAPM (i.e. column (1)) and from the modified CAPM which includes both components of $\alpha$ (i.e. column (4)). It can be seen that $\alpha_2$, and not $\alpha_1$, corrects for the underpricing of high fee portfolios. On the other hand, $\alpha_1$ reduces the residual alpha for all the portfolios by about 1 basis point (i.e. by comparing columns (3) and (4)). Since we are working with weekly returns, this amounts to about 0.5% reduction on an annual basis, which is economically significant. This effect is intuitive, since $\alpha_1$ results from the market portfolio being overvalued, which implies that stocks will produce higher risk-adjusted returns in the future relative to

\* for any positive beta stock
the market portfolio. Because low fee portfolios are overpriced by standard models, and most stocks fall into this category (i.e. see Table 1), stocks are on average overpriced by standard models, and thus the inclusion of $\alpha_1$ works in the desired direction by reducing the overpricing across the board. Even through $\alpha_1$ appears to have a smaller effect than $\alpha_2$, note that $\alpha_1$ affects all stocks, while $\alpha_2$ only affect the high borrow fee stocks, which constitute a small fraction of the stock market (i.e. see Table 1).

6 Conclusion

In this paper we study the effect of short selling constraints on the cross section of asset prices. In a theoretical framework, we show that standard models underprice short selling constrained stocks and overprice unconstrained stocks relative to the market portfolio. We confirm this prediction empirically, using an extensive database on short interest and borrow rates covering most of U.S. public equities, and further show that our model accounts for most of the mispricing.

There are several directions for future research. First, we have focused in our model on short selling frictions on the lending side of the market. We have done so because the data allows us to identify the size of this friction in the context of the model. An interesting and important topic is the role of short selling constraints on the demand-side, i.e. restrictions on short selling placed on some investors such as various types of mutual funds. How important is this friction, and how does it interact with frictions on the supply-side? While some purely empirical work has been done in relation to this question, theoretical and quantitative work is still in its infancy. Second, in this paper we have tested a part of the model that outlines the relationship between borrow fees and asset prices, while leaving out the part that links beliefs/disagreement to these two variables. Properly measuring and quantifying beliefs and disagreement and studying its effect on asset prices through short selling constraints in a quantitative framework is an exciting avenue for future research.
References


