FUNDAÇÃO GETULIO VARGAS
ESCOLA de PÓS-GRADUAÇÃO em ECONOMIA

Lira Rocha da Mota

Stock Lending Market, Short-Selling Restrictions, and the Cross-Section of Returns.

Rio de Janeiro
2017

Lira Rocha da Mota

Stock Lending Market, Short-Selling Restrictions, and the Cross-Section of Returns.

Dissertação para obtenção do grau de doutor apresentada à Escola de Pós-Grauação em Economia

Área de concentração: Finanças

Orientador: Carlos Eugênio Ellery da Costa

Co-Orientador: Marco Antonio Bonomo

Rio de Janeiro
2017

Ficha catalográfica elaborada pela Biblioteca Mario Henrique Simonsen/FGV

Mota, Lira Rocha da
Stock lending market, short-selling restrictions, and the cross-section of returns / Lira Rocha da Mota. – 2017.
112 f.

Tese (doutorado) - Fundação Getulio Vargas, Escola de Pós-Graduação em Economia.
Orientador: Carlos Eugênio Ellery Lustosa da Costa.
Coorientador: Marco Antonio Bonomo.
Inclui bibliografia.

1. Mercado financeiro. 2. Investimentos. 3. Empréstimo de títulos. 4. Venda a descoberto (Finanças). I. Costa, Carlos Eugênio da. II. Bonomo, Marco Antônio Cesar. III. Fundação Getulio Vargas. Escola de Pós-Graduação em Economia. IV. Título.

CDD – 332.6

LIRA ROCHA DA MOTA

"STOCK LENDING MARKET, SHORT-SELLING RESTRICTIONS, AND THE CROSS-SECTION OF RETURNS "

Tese apresentada ao Curso de Doutorado em Economia da Escola de Pós-Graduação em Economia para obtenção do grau de Doutor em Economia.

Data da defesa: 19/09/2016

Aprovada em:

ASSINATURA DOS MEMBROS DA BANCA EXAMINADORA

Carlos Eugenio Ellery Lustosa da Costa
Orientador (a)

Felipe Saraiva Iachan

Luis Henrique Bertolino Braido

Marco Antonio Cesar Bonomo

Bruno Cara Giovannetti

Ruy Monteiro Ribeiro

Agradecimentos

Agradeço primeiramente à minha família. Agradeço ao meu pai, Jesus, e a minha mãe, Luzanira, que sempre me apoiaram e pacientemente sempre me ajudaram e me acalmaram em momentos difíceis. Agradeço ao Luca, meu fiel companheiro, que esteve ao meu lado desde o começo desse projeto, ajudando tanto moralmente como também colocando a mão na massa muitas vezes.

Agradeço ao meu orientador, Marco Bonomo, que me guiou durante esses anos, sempre disponível para discutir ideias e me aconselhar. Agradeço ao meu orientador, Carlos Eugênio, que sempre manteve as portas de seu escritório aberas para mim, me inspirando com sua paixão por economia e clareza de pensamentos. Agradeço ao João Manuel de Mello que muito me ensinou sobre econometria. Agradeço Kent Daniel e Tano Santos por todo o conhecimento que compartilharam comigo e por terem aberto as portas para que pudesse trabalhar em projetos de pesquisa, que mais tarde viraria o terceiro capítulo dessa tese. Agradeço também aos meus professores da EPGE que foram fundamentais na minha formação.

Agradeço aos membros da minha banca: Bruno Giovanetti, Felipe Iachan, Luis Braido e Ruy Monteiro pelos comentários muito úteis à finalização dessa tese.

Agradeço aos meus amigos do doutorado, principalmente Fernando Barbosa e Simon Rotkke pelas longas horas de trabalho conjunto.

Finalmente, agradeço a Capes pela bolsa de doutorado e a Anbima pelo suporte financeiro fornecido via o Prêmio de Projeto de Doutorado.

Lista de Tabelas

1.1 **Summary Statistics: Brazilian Equity Loan Market** Number of stocks is the number of different shares traded in stock loan market in a certain year. Volume is the financial volume, price times number of shares, of all stocks lent. Average fee is the mean of stocks daily loan fee, while daily loan fee is a value weighted average loan fee each day for each stock that accounts for the lenders fee and commissions fees. Short-interest is number of shares held in loan contracts normalized by the total number of shares outstanding, this measure is calculated daily for each stock and the average presented is the simple mean. 13

1.2 **Summary Statistics for Arbitrage Contracts** Return is defined as stock return minus IBRX50 return cumulated for 8 days beging at two days before the record date. Arbitrage loan contracts are the ones that generated tax benefits - i.e, borrowers are mutual funds and lenders are retail investors or foreign investors. Arbitrage Short-Interest is the total shares involved in arbitrage loan contracts divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the IoNE dividend value as a percentage of the ex-date price. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. 13

1.3 **Reduced Form Results** Arbitrage Short-Interest is the total shares involved in arbitrage loan contracts divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. Robust standard errors corrected for clustering at the ticker level. 14

1.4 **First Stage.** Arbitrage Short-Interest is the total shares involved in arbitrage loan contracts divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. Robust standard errors clustered at the ticker level in parentheses. 15

1.5 **Unweighted Second Stage Regressions.** This table shows the estimates of equation 1.4. The dependent variable is the cumulative returns.The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. Robust standard errors are clustered at the ticker level. . . 16

1.6 **Instrument** IoNE is the IoNE dividend value as a percentage of the ex-date price. Log MC is the log of Market Cap. Illiiquidity is defined in Equation 1.3. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median SI is the median short-interest for the whole sample. Herfindahl Index (HI) is the sum of squares of each broker market share. Increases in the index indicate a decrease in competition. SI Before is the average short interest from 21 to 17 days before the event. Robust standard errors corrected for clustering at the ticker level. 17

1.7 **Reduced Form: Analyses of Persistence.** The dependent variable is the cumulative in different periods.$Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample.. Robust standard errors corrected for clustering at the ticker level. 18

1.8 **Second Stage Regressions: Analyses of Persistence** This table shows the estimates of equation 1.4. The dependent variable is the cumulative returns in different periods. Fee on Non-Arbitrage (Spec) Transactions is instrumented by the delta fee in arbitrage transactions. Robust standard errors are clustered at the ticker level.The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. 19

1.9 **Placebo test using normal dividends payout - Summary Statistics for "Arbitrage Contracts"** Return is defined as stock return minus IBRX50 return cumulated for 5 days beging at two days before the record date. Arbitrage loan contracts are the ones that generated tax benefits - i.e, borrowers are mutual funds and lenders are retail investors or foreign investors. Arbitrage Short-Interest is the total shares involved in arbitrage loan contracts divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the IoNE dividend value as a percentage of the ex-date price. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. 20

1.10 **Placebo test using normal dividends payout - Reduced Form Results** "Arbitrage Short-Interest"are loan contracts among arbitrage groups for the IoNE dividends divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts in the case of IoNE dividends. The method for calculating $Spec\Delta Fee$ is analogous. Normal Dividend is the payout value as a percentage of the ex-date price. Turnover is the average daily turnover in the 21 days before the IoNE dividend event. Average daily loan fee is the value weighted average loan fee in a day. Median loan fee is the median of average daily loan fee in the whole sample.Liquidity is defined in Equation 1.3. Robust standard errors corrected for clustering at the ticker level. Median SI is the median short-interest for the whole sample. 21

2.1 **Summary Statistics** This table reports summary statistics for the main variables in our study. Short interest is the total borrowed shares over total shares outstanding. Each month, SD Fee is calculated as the monthly average of daily standard deviation of fees. Average Fee is the average value weighted loan fee. Short interest is the total of borrowed shares over total shares outstanding. Market Cap is represented in Millions. Book-to-market ratio is calculated in the end of December of each year. The cases with negative book value are deleted. Herfindahl Index (HI) is the sum of squares of the market shares of the brokers. Increases in the index indicate a decrease in competition and an increase of market power . 38

2.2 **Panel Regression** Panel Regression with individual and time fixed effects. Log MC is the log of market cap. BM stands for Book to Market. Turnover is calculated by dividing the total number of shares traded in a year by the number of shares outstanding. Herfindahl Index (HI) is the sum of squares of each broker market share. Increases in the index indicate a decrease in competition. Degree is simply the number of different brokers lending (borrowing) that stock in the last month. 39

2.3 **Average Statistics per Year** This table reports Number of Observations and the average values of our main variables for each year. N Obs is the number of observations,i.e. datapoints. N Stocks is the number of stocks composing the dataset that year. BM stands for Book to Market. Short interest is the total borrowed shares over total shares outstanding. Each month, SD Fee is calculated as the monthly average of daily standard deviation of fees. All values presented as year average 40

2.4 **Correlation Matrix** This table shows the correlation between the variables used to sort the long short portfolios. SD Fee is the monthly average of daily standard deviation of fees; Avg Fee is the monthly average of daily average fees. Short interest is the total borrowed shares over total shares outstanding. DTC represents Days to Cover, i.e. the ratio between short interest and turnover. Each data-point here represent a monthly measure for a specific stock. 40

2.5 **Summary of Portfolio characteristics** For each month we calculate the monthly (market cap weighted) average of each variable. This table displays the time series average for each variable. 40

2.6 **Returns to Long Short Strategies** Our main portfolio is based on loan fee dispersion, SdMLd, Small dispersion Minus Large dispersion. CMEw stands for Cheap Minus Expensive and relates to average loan fee. The remaining small minus large portfolios, SsiMLsi and SdtcMLdtc, are based on Short Interest and Days to Cover. Sharpe index based on annualized returns. 41

2.7 **Market Factor Returns** This table presents average returns, standard deviation of returns and sharpe ratio for the main market factors in Brazil. HML stands for High Minus Low ; SMB represents Small minus Big; Winners minus Losers is abbreviated as WML and finally IML relates to liquidity, Illiquid Minus Liquid. Those factors relate to Book to Market, Market Cap, Momentum and Liquidity, respectively 41

2.8 **Returns to portfolio strategies based on Loan Fee Dispersion** This table provides portfolio alphas and loading, sorted on Loan Fee Dispersion. At the end of each month, all the stocks are sorted into three quantile: small (bottom 30%), Middle (40%), large (top 30%) deciles based on their loan fee dispersion at the end of each month. Portfolio returns are computed over the next month minus the monthly risk free rate - 30 day DI Swap. For the analysis we consider 5 factors: Market, Small Minus Big Factor (SMB), High Minus Low (HML), Winners Minus Losers (WML), Illiquid Minus Liquid (IML). The factors considered were extracted from NEFIN. Data runs from January 2007 to June 2013. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively. 42

2.9 **Returns to portfolio strategies based on Average Loan Fee** This table provides portfolio alphas and loading, sorted on Average Loan Fee. At the end of each month, all the stocks are sorted into three quantile: small (bottom 30%), Middle (40%), large (top 30%) deciles based on their average loan fee at the end of each month. Portfolio returns are computed over the next month minus the monthly risk free rate - 30 day DI Swap. For the analysis we consider 5 factors: Market, Small Minus Big Factor (SMB), High Minus Low (HML), Winners Minus Losers (WML), Illiquid Minus Liquid (IML). The factors considered were extracted from NEFIN. Data runs from January 2007 to June 2013. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively. 44

2.10 **Returns to portfolio strategies based on Short Interest** This table provides portfolio alphas and loading, sorted on Short Interest. At the end of each month, all the stocks are sorted into three quantile: small (bottom 30%), Middle (40%), large (top 30%) deciles based on their short interest at the end of each month. Portfolio returns are computed over the next month minus the monthly risk free rate - 30 day DI Swap. For the analysis we consider 5 factors: Market, Small Minus Big Factor (SMB), High Minus Low (HML), Winners Minus Losers (WML), Illiquid Minus Liquid (IML). The factors considered were extracted from NEFIN. Data runs from January 2007 to June 2013. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively. 46

2.11 **Returns to portfolio strategies based on Days to Cover** This table provides portfolio alphas and loading, sorted on Days to Cover. At the end of each month, all the stocks are sorted into three quantile: small (bottom 30%), Middle (40%), large (top 30%) deciles based on their days to cover at the end of each month. Portfolio returns are computed over the next month minus the monthly risk free rate - 30 day DI Swap. For the analysis we consider 5 factors: Market, Small Minus Big Factor (SMB), High Minus Low (HML), Winners Minus Losers (WML), Illiquid Minus Liquid (IML). The factors considered were extracted from NEFIN. Data runs from January 2007 to June 2013. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively. 48

2.12 **: Fama-MacBeth Regressions** This table reports results from Fama-Macbeth (1973) regression of monthly stock returns on Fee Dispersion, Average Fee, Short-Interest and Days to Cover. Fee Dispersion is the average daily standard deviation. Average Fee is the average value weighted loan fee. Short interest is the total of borrowed shares over total shares outstanding. Days-to-cover is short interest ratio over daily turnover. Size (log(MC)) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (log(BM)) is the natural log of book-to-market ratio calculated as end of December of each year. The cases with negative book value are deleted. The short term reversal measure (Reversal) is the lagged monthly return. All the t-statistics are Newey and West (1987) adjusted to control for heteroskedasticity and autocorrelation. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively. 50

3.1 **Average monthly excess returns for the test portfolios.** The sample period is July 1963 to December 2014. Stocks are sorted into 3 portfolios based on the respective characteristic - book-to-market (BEME), operating profitability (OP) or investment (INV) and independently into 3 size (ME) groups. These are depicted row-wise and indicated in the first two columns. Last, portfolios are sorted into 3 further porfolios based on the loadings forecast, conditional on the first two sorts. These portfolios are displayed column-wise. The last column shows average returns of all 9 respective characteristic portfolios. The last row shows averages of all 9 respective loadings portfolios. In the top panels we use the low power and in the bottom panels we use the high power methodology. 74

3.2 **Average monthly characteristics for the test portfolios.** Stocks are sorted into 3 portfolios based on the respective characteristic - book-to-market (BEME), operating profitability (OP) or investment (INV) and independently into 3 size (ME) groups. These are depicted row-wise and indicated in the first two columns. Last, portfolios are sorted into 3 further porfolios based on the loadings forecast, conditional on the first two sorts. These portfolios are displayed column-wise. At each yearly formation date, the average respective characteristic (BEME, OP, or INV) for each portfolio is calculated, using value weighting. At each point, the characteristic is divided by the NYSE median at that point in time. The time series from 1963 to 2014 is then averaged to get the numbers that are presented in the table below. Note that, while sorts in the high power panels are based on industry-adjusted characteristics, the averages and medians are calculated based on the unadjusted characteristics. The last column shows average characteristics of all 9 respective characteristic portfolios. The last row shows averages of all 9 respective loadings portfolios. In the top panels we use the low power and in the bottom panels we use the high power methodology. 75

3.3 **Sorting-factor exposures and five-factor alphas.** The last column shows the return of long low-loading short high-loading hedge-portfolios. The last row shows averages of all 9 loadings portfolios. In the top panels we use the low power and in the bottom panels we use the high power methodology. Alphas and ex-post loadings on the relevant factor are obtained from a regression of monthly excess returns of the test-portfolios on the 5 Fama and French factors from July 1963 to December 2014. 76

3.0 **Results of time-series regressions on characteristics-balanced hedge-portfolios.** Stocks are first sorted based on size and one of book-to-market, profitability or investment into 3x3 portfolios. Conditional on those sorts, they are subsequently sorted into 3 portfolios based on the respective loading, i.e., on HML, RMW or CMA. For Mkt-RF and SMB we use the prior sort on size and book-to-market. The "hedge-portfolio"then goes long the low loading and short the high loading portfolios. On the bottom, we form combination-portfolios that put equal weight on three (HML, RMW, CMA), four (HML, RMW, CMA, Mkt-RF) or five (HML, RMW, CMA, Mkt-RF, SMB) hedge-portfolios portfolios. Monthly returns of these portfolios are then regressed on the 5 Fama and French factors in the sample period from July 1963 to December 2014. In Panel A we use the low power and in Panel B we use the high power methodology. 81

3.1 **Results of time-series regressions on characteristics-balanced hedge-portfolios.** Stocks are first sorted based on size and one of book-to-market, profitability or investment into 3x3 portfolios. Conditional on those sorts, they are subsequently sorted into 3 portfolios based on the respective loading, i.e., on HML, RMW or CMA. For Mkt-RF and SMB we use the prior sort on size and book-to-market. The "hedge-portfolio"then goes long the low loading and short the high loading portfolios. On the bottom, we form combination-portfolios that put equal weight on three (HML, RMW, CMA), four (HML, RMW, CMA, Mkt-RF) or five (HML, RMW, CMA, Mkt-RF, SMB) hedge-portfolios portfolios. Monthly returns of these portfolios are then regressed on the 5 Fama and French factors in the sample period from July 1963 to December 2014. In Panel A we use industry-adjusted characteristics and 36 monthly observations for beta forecasts. In Panel B we use un-adjusted characteristics, but separate estimation of correlations and variances (with a 5 and 1 year window respectively) based on daily data and an additional intercept for the rank-year when estimating the loadings forecasts. . . . 82

Lista de Figuras

1.1 Typical dates flowchart for the IoNE dividend event. 6

1.2 **Distribution of Loan Fees By Firms by Year** Figure shows the distribution of the average daily value-weighted loan fees in percentage points per year between January 2007 and June 2013. The vertical axis shows the frequency of firms with loan fees in the interval reported on the horizontal axis. 22

1.3 Payout Announcement 22

1.4 **Loan Fee around Record Date of an IoNE dividend event** Figure shows the average loan fee around the record date of the IoNE dividend event. The daily loan fee for each share is the value weighed fee among all contracts in a certain day. The average loan fee is the common mean among shares. We consider 487 IoNE dividend events from January 2010 until June 2013. 23

1.5 **Loan Fee around Record Date of an IoNE dividend event by Groups** Arbitrage loan contracts are the ones that have tax benefits - i.e, borrowers are mutual funds and lenders are retail investors or foreign investors that took place before the record date and were liquidated after it. Speculative contracts are the other contracts. We define the daily loan fee as the value weighed average loan fee for each stock and each day around the IoNE date. The figure shows the average loan fee among all stocks for each day around the IoNE record date. 24

1.6 **Short-Interest around Record Date of an IoNE Event** Figure shows the average short-interest around the record date of the IoNE dividend event. The average short-interest is the common mean among shares. 25

1.7 **Short Interest around Record Date of an IoNE dividend event by Groups** Arbitrage loan contracts are the ones that have tax benefits - i.e, borrowers are mutual funds and lenders are retail investors or foreign investors that took place before the record date and were liquidated after it. Speculative contracts are the other contracts. We define the daily loan fee as the value weighed average loan fee for each stock and each day around the IoNE date. The figure shows average short interest among stocks for each day around the IoNE dividend record date. . 26

1.8 **Abnormal Returns around the IoNE Dividend Date** Figure show the cumulative abnormal returns starting from 10 days previous the record date. Abnormal return is calculated as the stock return minus the stock loading on the market portfolio IBX50, that accounts for the first 50 biggest stocks in market capitalization. The dotted line is the 95% percent confidence interval. 27

1.9 **Reduced Form Coefficient for Different Return Windows** Figure shows the estimated coefficient for different periods of 8-days cumulative returns. X-axis indicates the final date of 8-days cumulative returns, y-axis is the coefficient of $Arb\Delta Fee$ in reduced form regressions when considering as independent variables $Arb\Delta Fee$ and controls. The dotted line is the 95% percent confidence interval. . 27

1.10 **Second Stage Coefficient for Different Return Windows** Figure shows the estimated coefficient for different periods of 8-days cumulative returns. X-axis indicates the final date of 8-days cumulative returns, y-axis is the coefficient of $Spec\Delta Fee$ in structural form regressions when considering just one instrument: $Arb\Delta Fee$ and controls. The dotted line is the 95% percent confidence interval. . 28
1.11 **Loan Fee around Record Date of an normal dividend event** Figure shows the average loan fee around the record date of the normal dividend event. The daily loan fee for each share is the value weighed fee among all contracts in a certain day. The average loan fee is the common mean among shares. We consider 487 IoNE dividend events from January 2010 until June 2013. 28
1.12 **Short-Interest around Record Date of an IoNE Event** Figure shows the average short-interest around the record date of the IoNE dividend event. The average short-interest is the common mean among shares. 29
1.13 **Abnormal Returns around the Normal Dividend Date** Figure show the cumulative abnormal returns starting from 10 days previous the record date. Abnormal return is calculated as the stock return minus the stock loading on the market portfolio IBX50, that accounts for the first 50 biggest stocks in market capitalization. The dotted line is the 95% percent confidence interval. 29

2.1 **Petrobras Loan Fee** Histogram of fees paid by investors renting PETR4 (Petrobras) on 30th April 2008 and April first 2009. The y-axis represents the number of contracts, x-axis shows the fee. 51
2.2 **Mutual Funds Fee Relative Spread: Lender.** Evolution of relative spread of loan fees received by individual person and foreign investors when compared to mutual funds 52
2.3 **Mutual Funds Fee Relative Spread: Borrower.** Evolution of relative spread of loan fees paid by individual person and foreign investors when compared to mutual funds 52
2.4 **Loan fee Measures Over Time.** Time series of aggregate measures. Each company daily average loan fee is calculated as the mean fee for contracts opened that day. The equal weighted cross sectional average composes our aggregated measure. The other aggregated measures are analogous. 53
2.5 **Number of stocks for each portfolio.** 53
2.6 **Cumulative Returns for each long-short portfolio.** 54
2.7 **Transition Matrix** The figure plots the transition probabilities of stocks across the three sd-fee sorted buckets over the period of one month. For example, the first column shows the percentage of stocks in the first bucket that end up in each of the three buckets after one month. 54

3.1 **Rolling regression R^2s – HML returns on industry returns** This figure plots the adjusted R^2 from 126-day rolling regressions of daily HML returns on the twelve daily industry excess returns. The time period is January 1981-December 2015. 68
3.2 **HML loadings on industry factors.** The upper panel of this figure plots the betas from rolling 126-day regressions of the daily returns to the HML-factor portfolio on the twelve daily industry excess returns over the January 1981-December 2015 time period. The lower panel plots only the betas for the Money and Business Equipment industry portfolios, and excludes the other 10 industry factors. 69
3.3 **Volatility of the money and business equipment factors.** This figure plots 126-day volatility of the daily returns to the Money and the Business Equipment factors over the January 1981-December 2015 time period. 70

3.4 **Rolling regression R^2s – HML returns on *Money* industry returns.** This figure plots the adjusted R^2 from 126-day rolling regressions of daily HML returns on the daily *Money* industry returns from the 12 industry returns. The time period is January 1981-December 2015. 70

3.5 **Ex-post loading vs. characteristic.** This figure shows the time series average of post-formation factor-loading on the x-axis and the time series average of the respective characteristic on the y-axis of each of the 27 portfolios formed on size, characteristic and factor-loading. Panels A uses the low power methodology and B uses the high power methodology. The first row uses sorts on book-to-market and HML-loading, the second one operating profitability and RMW-loading and the last one investment and CMA-loading. 71

3.6 **Portfolio Cumulative Returns.** This figure plots the cumulative returns of of the five FF(2015) portfolios, and the residual portfolio. The residual portfolio is the equal-weighted combination of the HML, RMW, and CMA hedge portfolios, orthogonalized to the five-factors. Each portfolio assumes an investment of \$1 at close on the last trading day of June 1963, and earns a return of $(1 + r_{LS,t} + r_{f,t})$ in each month t, where $r_{LS,t}$ is the long-short portfolio return, and $r_{f,t}$ is the one month risk free rate. 72

Sumário

1 Short-Selling Restrictions and Returns: a Natural Experiment **1**
1.1 Introduction . . . 1
1.2 Stock Loan Market in Brazil . . . 4
1.3 The Tax Arbitrage . . . 5
1.4 Data . . . 6
1.5 Empirical Strategy: Identification and Specification . . . 7
1.5.1 Specification . . . 9
1.6 Results . . . 9
1.6.1 An Event Study . . . 9
1.6.2 Reduced Form . . . 9
1.6.3 First Stage . . . 10
1.6.4 Second Stage . . . 10
1.6.5 Determinants of Instrument Variation . . . 10
1.7 Falsification Tests . . . 11
1.7.1 Persistence Test and Reversal . . . 11
1.7.2 Placebo Test . . . 11
1.8 Conclusions . . . 12
1.9 Tables . . . 13
1.10 Figures . . . 22

2 Loan Fee Dispersion and the Cross-Section of Returns **30**
2.1 Data, Variables and Summary Statistics . . . 33
2.2 Loan Fee Dispersion . . . 34
2.3 Loan Fee Dispersion and Returns . . . 36
2.3.1 Sorts on Loan Fee Dispersion . . . 36
2.3.2 Fama and MacBeth Regressions . . . 37
2.4 Conclusion . . . 37
2.5 Tables . . . 38
2.6 Figures . . . 51

3 The Cross-Section of Risk and Return **55**
3.1 Introduction . . . 55
3.2 Theory . . . 58
3.2.1 Characteristic-Based Return Factors . . . 59
3.2.2 Relation between the characteristic-weighted and MVE portfolio . . . 59
3.2.3 An optimized characteristic-based portfolio . . . 61
3.3 Industry Factor Loadings . . . 61
3.4 Low- and High-Power Test Based Hedge Portfolios . . . 63
3.4.1 Construction of the factors and the set of test portfolios . . . 63
3.4.2 Postformation loadings . . . 65
3.4.3 The pricing of the test portfolios . . . 65

3.4.4 The main result 66
3.4.5 Industry adjustments 66
3.5 Conclusions 67
3.6 Figures 68
3.7 Tables 73

Capítulo 1

Short-Selling Restrictions and Returns: a Natural Experiment[1]

ABSTRACT

We estimate the causal impact of short-selling restrictions on returns. We take advantage of a unique dataset and exploit a source of exogenous variation in loan fees provided by a tax-arbitrage opportunity that existed in Brazil from 1995-2015. The tax-arbitrage opportunity stems from the fact that domestic mutual funds were exempted from income taxes on dividends received by stocks they borrowed, whereas the original owner would be taxed if she did not lend out the stock. Because we observe all equity loan transactions and transacting parts investor identity, we can distinguish equity lending motivated by tax-arbitrage from speculative transactions according to the borrower-lender match. We show that the loan fee on tax-motivated transactions is a source of exogenous variation to estimate the causal impact of the (endogenous) loan fee on stock prices. We find that increases in stock loan fees have strong impact on stock prices.

1.1 Introduction

Since the seminal article of Miller (1977), the impact of short-selling constraints on financial markets have been subject of numerous theoretical[2] and empirical studies.[3] A well known predicted effect is that the interaction between heterogeneous valuations and short-sale constraint would lead to prices superior to the average valuation. Thus, both an increase in heterogeneity or in short-sale constraints would pressure prices upwards.[4] While several empirical papers have convincingly illustrated the effect of the (cross-sectional and temporal) variation in heterogeneity

[1] This is a joint work with Marco Bonomo and João Manuel de Mello.

[2] e.g. Harrison and Kreps (1978), Diamond and Verrecchia (1987), Duffie, Garleanu, and Pedersen (2002), Scheinkman and Xiong (2003), Hong and Stein (2003), Hong, Scheinkman, and Xiong (2006), Blocher, Reed, and Wesep (2013)

[3] e.g. Chen, Hong, and Stein (2002a), Asquith, Pathak, and Ritter (2005a), Cohen, Diether, and Malloy (2007), Saffi and Sigurdsson (2011), Beber and Pagano (2013), Boehmer, Jones, and Zhang (2013), De-Losso, Genaro, and Giovannetti (2013), Kaplan, Moskowitz, and Sensoy (2013), Prado, Saffi, and Sturgess (2014)

[4] Boehme, Danielsen, and Sorescu (2006), Desai, Ramesh, Thiagarajan, and Balachandran (2002), Nagel (2005)

and loan demand on stock prices,[5] documenting the effect of variation of stock loans supply poses a greater challenge.

Identification of the causal impact of short-selling restrictions on returns has been elusive for three reasons: data availability, lack of good measures of short-sale restriction, and, more importantly, the endogeneity of the decision to sell short. Several studies examined the effect of short-selling restrictions, but limited data availability has been always an issue (e.g., Cohen et al. (2007), Saffi and Sigurdsson (2011)). In most markets, lending services are provided over-the-counter by big custodian banks. Data are rarely available for any significant fraction of the market. For this reason, researchers have relied on proxies for short-sale constraints, such as institutional ownership for short supply and the Markit loan fee index for loan fee. However, these proxies tend to show little variation in the short-run.

The Brazilian security lending market provides a unique opportunity to circumvent the difficulties with data. In contrast to other countries, all lending transactions are registered at a centralized platform - the BM&FBovespa. Because of the centralized nature of the market, we have access to a unique dataset containing all lending transactions, and their characteristics, such as type of borrower and lender (individual investors, mutual funds, etc.), contract length, borrowing fees, number of securities lent, among others.

Data availability is just one obstacle to the proper measure of the impact of short-selling restrictions on stock returns. Even more challenging is the fact that the decision to supply or demand equity loans is not random, and probably driven by investors' (unobserved) expectations of future returns. One cannot infer causality by associating lending fees - as a measure of short selling constraints - with returns because fees are determined in equilibrium.

We explore an exogenous variation in the supply of loanable stocks provided by a tax arbitrage opportunity. Brazilian firms distribute two types of dividends: standard dividends and IoNE dividends (literally Interest on Net Equity or IoNE). For tax reasons, firms prefer IoNE dividends, which are treated as interest expenses. Shareholders pay a 15% tax rate on IoNE dividends, as opposed to an average of 25% that the firm pays on profits (equity holders are exempt from taxes on standard dividends).

Domestic mutual funds were exempt from the 15% tax on IoNE, and the tax status is determined by the holder of the stock on the record date (not by the actual owner of the stock). Thus, transactions whose lender *is not* a domestic mutual fund and the borrower *is* a domestic mutual fund economize the 15% IoNe dividend tax. We show that the tax-arbitrage opportunity puts pressure on the lending market during IoNE dividend events, causing lending fees to rise sharply when the borrower, but not the lender, is a mutual fund. To a lesser but still very significant degree, fees also rise in transactions when the borrower is not a domestic mutual fund, which do not save the 15% IoNe dividend tax. We argue that the increase in lending activity from non-domestic mutual funds to domestic mutual funds - which we call the tax-arbitrage segment - is not related to the fundamental value of the firm. Thus, during IoNE events, the tax-arbitrage opportunity exogenously reduces the supply of loanable funds for short-selling purposes, which we call the speculative segment. In short, we have a quasi-natural experiment in the stock lending market.

We also deal with the incomplete price drop on the *ex-dividend* date, which challenges identification strategies that explore dividend distribution events. In a nutshell, we exploit the fact that, while spot market transactions take three business days to settle, lending transactions are settled on the same business day. This allows us to isolate the incomplete price drop on the *ex-dividend*.

We have three findings. First, tightness in the tax-arbitrage segment causes abnormal returns in stock prices (the *Reduced Form*). More specifically, the larger the increase in lending fee in the tax-arbitrage segment, the higher the abnormal returns in stock prices during the days

[5] e.g. Chen et al. (2002a), Asquith et al. (2005a), Cohen et al. (2007), De-Losso et al. (2013), Prado et al. (2014)

surrounding the IoNE dividend event. During a typical event, stock prices have a 0.44% return in excess of the market accumulated over 8 trading sessions.

Second, tightness in the tax-arbitrage segments spills over to the speculative segment. The larger the increase of the lending fees in the tax-arbitrage segment, the larger the increase of lending fees in the speculative segment. This is the *mechanism* through which the exogenous shock propagates to share prices.

Finally, we use the increase in lending fees in the tax-arbitrage segment as a source of exogenous variation to estimate the causal effect of lending fees on stock prices (the *Structural Form*). During IoNE dividend distribution events, the average increase in the lending fees in the speculative market causes a 0.385% abnormal return in the stock price. Causal interpretation demands that increases in tightness in the tax-arbitrage segment have no direct impact on share prices (only through the lending fee for short-selling transactions). We provide evidence that this is a reasonable assumption in our setting.

The empirical literature has dealt with the problem of the endogeneity of short-selling restrictions with varying degrees of success. Saffi and Sigurdsson (2011) explore a panel of stocks and associate lending supply and fees to several measures of price efficiency, such as bid-ask spreads. Their panel strategy with fixed-effects and a long time-series accounts for a large fraction of variation. However, it is not possible to guarantee that unobserved (to the econometrician) shocks are not driving simultaneously both returns (or bid-ask spreads) and fees or lending supply. In fact, earlier papers showed that short-interest predicted negative future abnormal returns, which suggests that short-interest has informational content (Cohen et al. (2007); Figlewski (1981).

Cohen et al. (2007) choose a different strategy. They postulate that higher fees associated with increases in lending indicate a (net) increase in the demand for shorting. Increases in fees coupled with reductions in lending represent a supply shock. They found that no significant effect of their identified supply shocks on returns. Their identification strategy implicitly assumes that supply shocks do not themselves contain any information about the returns, which in many settings can be a strong assumption. De-Losso et al. (2013), exploring a Brazilian sample as we do, define a supply shock as in Cohen et al. (2007), with the advantage of observing the size of the supply shock for each stock for a set of small stocks for which lending offers are public information. They found a significant negative relation between supply of loans and returns. Similarly to Cohen et al. (2007), the identification strategy works only if the decision to change the supply is driven solely by factors other than opinions on the stock value.

According to Prado et al. (2014), stocks with higher concentration of institutional ownership tend to have lower lending supply. They found that higher concentration of institutional ownership is associated with higher idiosyncratic volatility of returns, limiting arbitrage. They show that demand shocks, as identified in Cohen et al. (2007), generate substantially lower returns in stocks with higher concentration of institutional ownership. Chuprinin and Massa (2012) also use characteristics of institutional investors as proxies for short-sale constraints. When they instrument short-interest by those characteristics they obtain that higher short-interest is associated with higher return, corroborating the Miller hypothesis.

Boehmer et al. (2013) explores the short-sale ban of 2008 of more than 1,000 stocks, and find no price bump for these stocks. Their identification strategy is to match banned stocks with similar stocks that have never suffered the ban. Matching is an adequate procedure to account for observed heterogeneity, but the decision to ban short-selling of some particular stocks may be driven than factors unobserved to the econometrician (which may be the point of short-selling, namely, superior information).

Kaplan et al. (2013) perform a field experiment, randomly increasing the supply of stocks available for lending by a particular money manager. They find almost no impact on outcome variables such as returns and bid-ask spreads. Experimental data provides convincing exogenous variation. However, the supply shock may be too small to produce a quantitatively relevant impact.

From a methodological viewpoint, our work is closest to Thornock (2013) and Blocher et al. (2013). Both analyze the effect of a restriction in the supply of shares around dividend events due to differential tax treatment between actual and replacement dividends.

Thornock (2013) finds that, around dividend events, the equity loan market gets tighter, with higher loan fees and lower short-interest. Similarly to our experiment, the differential tax treatment produces a negative shift supply of loanable shares. Thornock (2013) associates the negative shift in supply to market quality measures. One of them is the occurrence of average abnormal returns on ex-dividend dates. He finds that the dividend distribution events are associated with abnormal returns for high-dividend yield stocks and for stocks whose equity lending market was tight prior to the ex-dividend date.

Blocher et al. (2013) propose a framework for analyzing the impact of the equity loan market's impact on share prices. In their model, short-selling restrictions have an impact on share prices only for hard-to-borrow stocks. Following Thornock (2013), they use the increase in loan market tightness around dividend events as one of their strategies to test the model. They find that dividend distribution events are associated with abnormal returns around the ex-dividend date for stocks whose equity lending market is tight. Differently from Thornock (2013), Blocher et al. (2013) measure tightness through lending rates, not short volume.

Both Thornock (2013) and Blocher et al. (2013) measure the stock market return around *ex-dividend* dates. Hence, their results are subject to confounding effect of the incomplete price-drop phenomenon at ex-dividend date. In order to exclude this confounding effect, the *ex-dividend* data is the starting point of our cumulative return window. The only thing left is the increased tightness of the stock loan market around the record date.

Relative to the literature, we also move forward by documenting a clear mechanism behind the abnormal returns during dividend distribution events. Thornock (2013) and Blocher et al. (2013) results are based on measuring abnormal returns during events when the equity lending market is arguably tight, a reduced-form object. Our data enables us to disentangle two equity lending markets: one affected (the tax arbitrage market) and one not affected by tax arbitrage opportunity (the speculative market). We have a clear mechanism: the tax-arbitrage opportunity, an exogenous shock, causes tightness in tax-arbitrage segment; tightness in the tax-arbitrage segment causes tightness in the speculative market, which, in turn, causes abnormal returns. Differently from received literature, we can directly link *exogenous* variation in lending fees to abnormal returns.

The paper is organized as follows. Section 1.2 gives an overview of the market for stock loans in Brazil. In Section 1.3 we explain the tax arbitrage opportunity that impacts the loan market in the period surrounding the IoNE payment dates. In Section 1.4 we describe the data used for the paper. In Section 1.5 we present our identification strategy. The results are analyzed in section 1.6. Section 1.7 we present falsification tests. First we verify by changing the return window around the dividend record date, that the results we obtained cannot we reproduced for windows that are not consistent with the proposed mechanism. Then, we perform a placebo test by reproducing our empirical strategy around ordinary dividend dates. The last Section concludes.

1.2 Stock Loan Market in Brazil

Similarly to other countries, trading in equity loan market in Brazil is mostly over-the-counter (OTC). Differently from other countries, all loan contracts must be registered in the Brazilian stock exchange, BM&FBovespa. BM&FBovespa acts as a clearing platform, and as a central counterpart. It guarantees all loan contracts, and keeps track of the contract collateral.

A typical lending operation involves the exchange and four different participants: the lender, the borrower, the lender's broker and the borrower's broker. The cost to borrow a stock is the loan fee (which includes both brokers commission fees), plus an exchange fee of 0.25% annually.

Table 1.1 contains summary statistics about the Brazilian equity loan market from 2007 through 2013. In this period, the number of stocks lent became approximately constant at 300. However, other statistics indicated that the stock loan market gained importance. The average annual growth in the volume of equity lending was 23.5%. The average short-interest doubled, reaching 1.236% in 2013. Average loan fees dropped by an average 43 basis points (bps) per year, from 5% to 2.148% rate. On the lending side, retail investors, domestic mutual funds and foreign investors represent approximately 90% of the market. Borrowers are mostly domestic mutual funds (56.7%) and foreign investors (27.8%)

Figure 1.2 depicts the distribution of the average loan fees for each stock from January 2007 through June 2013. Fees are high in Brazil, relative to more mature markets. 20% of the fees charged were higher than 6% per year. Nevertheless, the performance of the lending market is improving rapidly, according to this metric, with the whole distribution shifting to the left (Figure 1.2).

1.3 The Tax Arbitrage

Until a change in law in August 2014, a loophole in the tax law provided a tax-arbitrage opportunity involving equity lending.

Brazilian firms distribute profits through two types of dividend, standard and IoNE dividends (literally Interest on Net Equity). The difference is the tax treatment. The IoNE dividend was created in 1995 with the stated goal of incentivizing firms' capitalization. Corporate taxable income is reduced by the distribution of IoNE dividends, but shareholders pay a 15% tax rate (the same rate as interest on long-term debt, thus the name Interest on Net Equity). Standard dividends are not deducted from the corporate taxable income, but their recipients are exempt from taxation. In Brazil the law requires the distribution of a minimum of 25% of profits.[6] Firms prefer to distribute IoNE dividends because the corporate tax rate is 25%, and thus IoNE dividends economize on taxes. In fact, there is an upper limit to the amount of IoNE dividends distributed, but not to standard dividends. Firms can distribute IoNE dividends up to the smallest of three following numbers: the net worth times the Long-Term Interest Rate (a prime rate determined by the federal government), 50% of the current period earnings before corporate taxes, and 50% of the accumulated earnings and reserves in previous periods.

In order to avoid double taxation, mutual fund and investment clubs are exempt from security earning taxes, since funds shareholders pay taxes on the NAVs appreciation. The difference in tax treatment between types of investors generated a tax arbitrage until the August 2014 law closed the loophole. Since the borrower was considered the owner of the stock for tax purposes, funds could borrow a stock from a taxable investor during the period of its IoNE dividend distribution and receive the full dividend amount. They had to pay the original stockholder only the dividend net of the 15% of taxes, retaining the difference. Taxable stock lenders (e.g. retail and foreign investors) and brokers could appropriate part of the funds gains by increasing their lending and brokerage fees, respectively.

Relevant dates for dividend distribution in Brazil are similar to those in the United States. Figure 1.1 depicts the flowchart of important dates for our empirical exercise in a typical IoNE distribution affect. The first date is the *earnings' announcement day*, an official public statement of a company's information for a specific period, typically a quarter. On *earnings' announcement date*, investors get to know the company's distribution policy, including the amount of payouts in the form of standard and IoNE dividends. Earnings' announcements usually cluster around the

[6] Almost all companies in our sample operate under the so-called Real Income regime. The Real Income regime is progressive. Companies pay a 15% tax rate on annual incomes up to BRL 240,000.00, approximately US$89,000.00; above this threshold companies pay an additional 15% surcharge. Almost all companies in our sample have annual income well above the threshold. The tax rate is then about 24% for those companies. In contrast, the tax rate on IoNE dividends is 15%. Thus, when corporations choose to distribute, they almost always distribute through IoNE up to the legal. Firms often choose to distribute the minimum of 25% of profits.

earnings' season. The second date is the *payout announcement day* (also called the *declaration day*), when the next IoNE dividend payment is announced, as well as the ex-dividend, record, and payment dates. The payout announcement usually happens a month after the earnings' announcement. The next is the ex-date day. Stocks bought in the spot market on or after the ex-date do not receive the announced IoNE. The ex-date is at least one business day after the*payout announcement*. The *record date day* determines eligibility to the IoNE dividend. Holders of the stock as shown in the company's books at the record date are entitled to the IoNE dividend. Because it takes three trading days for a stock transaction in the spot market to be settled, the *ex-date* is two trading days before the *record date day* and the last *cum day* is three business days before the *record date day*. Since stock loan transactions are registered immediately, when some investor borrows a stock before the register day, she is entitled to receive its dividend payment. Therefore, the borrower has two days more than the buyer to do the transaction in time to have the right to receive the dividend. We will exploit this lag for identification purposes. Finally, the payment date is when the IoNE dividend payment is made. The payment date varies considerably.

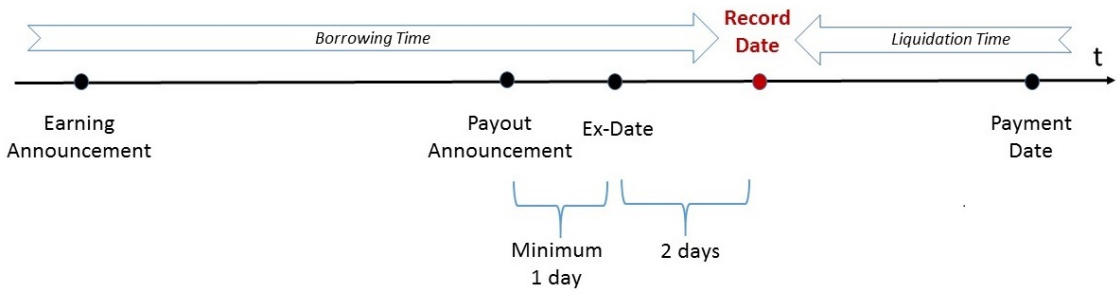


Figura 1.1: Typical dates flowchart for the IoNE dividend event.

1.4 Data

BM&FBovespa keeps detailed records of all lending transaction. We observe the whole process: offers, contracts, and liquidation. We also observe information about investors, brokers, and maturity.

The data set runs from January 2010 to June 2013. We use the following data: contract date, liquidation date, loan fee, quantity of shares and investor type. We merge the BM&FBovespa data with the Radar dataset, which contains information about payout events. Merging the two datasets allows us to identify loan transactions that are eligible for tax arbitrage, the *tax-arbitrage* contracts, and non-eligible transactions, the *speculative contracts*.

We define a loan contract to be *tax-arbitrage driven* if and only if it has three characteristics: 1) the lender is a taxable investor and borrower can receive IoNE dividends exempt from taxes (mutual funds and investment clubs); 2) it takes place the day or before the record date; 3) it is liquidated after the record date. All other contracts are considered *speculative driven*.

Information on stocks returns and traded volume are from Economática. We use abnormal returns, the stock return minus the market portfolio, measured by the IBX50 [7]. For each IoNE dividend event, we specify a window which includes 21 days before and after the event. Returns are defined as:

$$r_{t,t+1} = log\left(\frac{P_{t+1} + ND_{t+1}}{P_t}\right)$$

[7] The index is value weighted and accounts for the 50 stocks with the highest market capitalization

where ND stands for net dividends, i.e., gross dividends discounted with the highest tax rate. We isolate the imperfect *ex-dividend date* price drop effect by considering returns after the *ex-dividend date*. In short, dividend adjustment in returns is not a concern.

The loan fee of a given stock in a given day is the value-weighted average fee of loan contracts that took place that day. ΔFee is the percent change in the loan fee inside and outside of the event. Inside the event means within 5 trading days of the record date. Outside means belonging to the interval $[-21, -17] \cup [17, 21]$, where 0 is the IoNE record date[8].

$$\Delta Fee = \frac{Fee_{in} - Fee_{out}}{Fee_{out}} \tag{1.1}$$

We also calculate an illiquidity measure following Amihud (2002). We measure stock i illiquidity at day t as the average ratio of the daily absolute log return to the trading volume on that day, as described in equation 1.2. For every event, we calculate the average illiquidity for that stock in the last 252 business days, as described in equation 1.3.

$$Illiquidity_{it} = \frac{|ln(r_{it})|}{Volume_{it}} \tag{1.2}$$

$$Illiquidity_{event} = \frac{1}{N_{event}} \sum_{t=1}^{N_{event}} \frac{|ln(r_{it})|}{Volume_{it}} \tag{1.3}$$

The IoNE dividend event has to satisfy certain criteria to be in the sample. There must be arbitrage contracts starting in the five days before the record date of the event. Stocks must have had been traded during all 21 trading days before and after the record date. We demand that we have data on turnover during all 21 trading days before the record date. We trim outliers by discarding the 1% highest returns in the sample [9]. Events in sample include stocks that represent 70% of BM&FBovespa market capitalization. Table 1.2 contains summary statistics for those events.

Summary statistics hint at our main results. Stocks show an average 0.48% return over the IBRX50 index in a 8-trading-day period around the IoNE dividend event, but dispersion is large. Total short-interest increases drastically. Fees on arbitrage contracts increase fivefold, reflecting the fact that lenders appropriate part of the tax arbitrage profit. Fees on speculative contracts also increase significantly, 73%. The amount of IoNE dividend distributed is roughly 0.23% of the companies' market cap.

1.5 Empirical Strategy: Identification and Specification

The IoNE dividend distribution events provide exogenous variation to estimate the causal impact of short-selling restrictions on returns. We deal with two elusive problems: the endogeneity of the changes in lending fees, and the contamination of the incomplete *ex-dividend* price drop.

The tax arbitrage opportunity described in 1.3 provides quasi-natural experimental variation in lending fees on speculative lending transaction (i.e., borrowing to short sell).

During events of IoNE dividend distribution, the lending market split into two segments. In the *tax-arbitrage* segment, lenders that are subject to tax withholding (retail and foreign investors, banks, and firms) lend to borrower that are tax exempt (domestic mutual funds). The *speculative* segment contains all transactions conducted between all other combinations of lender-borrower pairs (e.g., domestic mutual funds lending to domestic mutual funds). Because

[8] For robustness, we tried different intervals for the ΔFee calculation, as well as using the maximum fee in each period instead of the mean. The results are unchanged and are available upon request.

[9] For robustness, we also trimmed at the following cutoffs: 0%, 5%. Results are similar and are available upon request

they do not generate any tax benefit, we assume that stock borrowing in this segment is done for speculative reasons.

Because of the tax arbitrage opportunity, the events of IoNE dividend distribution severely disrupted prices and quantities in the lending market.

Figure 1.4 depicts fees on 242 trading days around the IoNE dividend event. Between 21 and 9 days before the event, fees average 2.6%. After that, fees increase, reaching an impressive 10.4% at the record date. Quite interestingly, most the increase in fees occurs during the three-day period that precedes the *record date*.

Figure 1.5 splits the dynamics of loan fees for *tax-arbitrage* and *speculative* loan contracts. The increase in average loan fees in figure 1.4 stems mostly from *tax-arbitrage* contracts.

The pressure in the tax-arbitrage segment spills over to the speculative segment. Albeit less dramatically, fees on *speculative* contracts increase significantly, from 2.66% to 3.68%, a 38% increase.

Figure 1.6 depicts the short-interest around the event. Twenty-one days business days before the *record date* the average short-interest on *tax-arbitrage* contract was, on average, 1.51% of volume. They increase to 3.44% on the *record date*. In contrast, short-interest on *speculative* drops, from 1.47% to 0.96% of volume on average, a 35% drop.

Average changes in short interest and fees suggest that, during IoNE dividend events, the *demand* for borrowing stocks increases in the *tax-arbitrage* segment, and the *supply* of loanable stocks to *speculative* segment contracts.

We go beyond averages. We explore the heterogeneity across IoNE distribution events in the magnitude of the increase in the fees of the *tax-arbitrage* transactions. We show the following chain of causation. During events in which the fees and short-interest in *tax-arbitrage* segment increase more steeply, fees in the *speculative* segment also increase more. Moreover, the more fees increase in the *speculative* segment increase, the larger the increase in stock prices.

In summary, we estimate the causal effect of short selling constraints on returns by using the increases in lending fees in the arbitrage segment as a source of exogenous variation for the fees of the speculative segment.

The identification hypothesis is that, during IoNE distribution events, contracts in arbitrage segment do not reflect investors expectations about future returns. This is a reasonable assumption because in this period lending fees in the tax-arbitrage segment are just too high to rationalize short selling stocks (if borrowers short sell the borrowed stock they are not entitled to the dividend payment). Around IoNE dividend events, one can safely assume that stocks borrowed for tax-arbitrage purposes are not sold in the spot market. Therefore, it reasonable to assume that the changes in fees in the tax-arbitrage segment are exogenous to return expectations in the spot market.

We also deal with the incomplete *ex-dividend* price drop, a threat to identification in studies that use dividend distribution events that pressure lending markets (Frank and Jagannathan (1998),Elton, Gruber, and Blake (2005)). The incomplete *ex-dividend* drop produces abnormal positive returns in spot market in periods around dividend distribution events. Differently from the previous literature, we explore the fact that spot and lending transactions are settled on different dates. Spot market operations take three business days to be settled, which implies that the *ex-dividend* date is two days before the dividend *record date*. In contrast, lending operations are settled on the same business day, and most of the arbitrage driven loans actually take place on the *record date* and not on the *ex-date*. We avoid capturing the incomplete price drop at ex-dividend date by restricting the analysis to cumulative returns accruing from the *ex-dividend* date onward. The assumption is that any information or abnormal return due to the dividend distribution is already factored into prices by the *ex-date*.

1.5.1 Specification

The main object of interest is the impact of speculative (non-arbitrage) fees on returns, which provides a measure of the impact of short-selling restrictions on returns. Our main regression of interest is:

$$Return_{ei} = \alpha + \beta Spec\Delta Fee_{ei} + Controls_{ei} + \varepsilon_{ei} \tag{1.4}$$

The subscript e stands for IoNE event, defined as the 8-days period from the *ex-dividend* date. Subscript i stands for stocks. $Return_{ei}$ is the cumulative return in the event window e and stock i, $Spec\Delta Fee$ is the change in the fee on speculative contracts during the event. Controls include IoNE dividend as a percentage of the ex-price, illiquidity, market cap, turnover, median fee and median short interest. Notice that a lower expected return on the stock should lead to an increase in short-selling activity and an increase in the lending fee in the speculative market. This makes $Spec\Delta Fee$ endogenous, biasing the OLS estimation results.

In order to overcome the endogeneity problem we use two instruments: increases in fees and short-interest on arbitrage transactions. The hypothesis is that changes in the fees and short-interest on arbitrage transactions contain no relevant information for returns in the event window.

The first stage is:

$$\begin{aligned} Spec\Delta Fee_{ei} = \beta_0 + \beta_1 ArbShortInt_{ei} + \beta_2 Arb\Delta Fee_{ei} + \\ Controls_{ei} + \epsilon_{ei} \end{aligned} \tag{1.5}$$

We take a conservative stance and cluster observations at the ticker level when estimating standard deviations.

1.6 Results

1.6.1 An Event Study

The Miller hypothesis implies that events in the lending market spillover into the spot market. Before presenting our main results, we perform an event study for IoNE dividend distribution to calculate the average excess return around the IoNE dividend *record date*. Figure 1.8 depicts the cumulative excess return starting 10 days before the record date. The 10-day cumulative return on the record date is 83bps (significantly positive at 5% confidence level). Inspection the dates around the record date, shows that the positive excess return persists for some days after *record date*.

1.6.2 Reduced Form

Table 1.3 has the results of the reduced-form. In column (1) we include only changes in short-interest among *tax-arbitrage* transactions (besides the controls). The estimated coefficient, 0.062, means that the mean *tax-arbitrage* short-interest - 2.458 percentage point. - is associated with a non-significant abnormal returns of $2.458 \times 0.062 = 0.15$ percentage points cumulated over eight trading sessions. In column (2) we regress the abnormal returns on the size of the increase in fees in *tax-arbitrage* transactions. We find that the mean increase on fee of arbitrage contracts causes a 0.23% abnormal return. When both instruments are included, only Arbitrage Δ fee matters (column (3)). We find no impact on abnormal returns of increases in fees in speculative transactions (column (4)). This is not surprising because most of the variation in fees in speculative transactions is endogenous. Finally, in column 5 we include the instruments and the endogenous variable, and reach the same conclusions as in columns 1 through 4. In summary, changes in loan fee in the *tax-arbitrage* transactions are associated with abnormal returns, which

is compatible with the idea that the tax-arbitrage tightness in lending markets increase share prices.

1.6.3 First Stage

Table 1.4 presents the first stage results. Column (1) presents the results of the regression of variation of *speculative* fees on one instrument, the short-interest of *tax-arbitrage* transactions. Short-interest in *tax-arbitrage* transactions cause increases in fees on *speculative* transactions. The relationship is not strong enough when considering just this instrument (partial F -statistic of 7.68) and weak instruments issues may be a concern. Column (2) is the same regression but the instrument is the change in *tax-arbitrage* transactions. The estimated coefficient means that a typical increase *tax-arbitrage* loan-fee - 5.080 - causes a $(5.080 \times 0.089) * 100 = 45.2$ percentage point increase in fees in *speculative* transactions, a large increase (fees are normally 1 percentage point). In this case, the partial F-statistic is 18.55 and weak instrument is not a concern. In Column (3) we include both instruments.

1.6.4 Second Stage

Table 1.5 presents results for the second stage. The columns are equivalent to the models presented in the last section. We concentrate on the second column, for which the instrument is the Arbitrage Δ Fee. The estimated coefficient on changes in *speculative* fees is 0.524, and quite precisely estimated (p-value less than 1%). The estimated coefficient means that the typical increase in *speculative* fees - 73.5% - causes a 0.385 percentage point abnormal return in share price. The exogenous increase in *speculative* fees is large enough to make short-selling less attractive, producing a large increase in share prices, at least in short-run.

1.6.5 Determinants of Instrument Variation

In order to investigate if the assumptions about our chosen instrument are empirically plausible, we regress the variation in loan fee for the arbitrage contracts on several variables. Our hypothesis is that the tightness in arbitrage market is mainly driven by the amount of income that could be engendered by fiscal arbitrage operations. Since, the amount of taxes on IoNE dividend distribution is the upper bound on the income generated by evaded taxes in those operations (and the tax rate is constant), the IoNE dividend distribution per share, and the market capitalization seem plausible determinants of the tightness of this lending market. Notice, that the IoNE dividend per share distribution is announced at earnings date, which is many days earlier than our event window. Thus, it should not contain any information about stock market return in our event window. Table 1.6 corroborates our hypothesis, since both market capitalization and IoNE dividend per share are positively and significantly related to the loan fee for arbitrage contracts. This result holds, in regressions when each of those variables is the only explanatory variable, when both are combined, or when both are combined with other regressors.

We also test for some characteristics microstructure of the lending market that could affect the impact of the induced arbitrage operations on the lending fee, and for controls considered in the main regressions. We test for illiquidity, turnover, median short-interest over the whole sample, the short interest before the event window, the Herfindal index among lenders' and borrowers' brokers. However, none of those variables seems to be related with our instrument. In summary, one can assume that, conditional on the variables related to the magnitude of the IoNE dividend distribution, the variation left in tightness of the *tax-arbitrage* market is as good as random.

1.7 Falsification Tests

For robustness, we perform two exercises. First, we analyze the returns in different windows around the IoNE dividend record date. Second, we use normal dividends payout as a placebo test for our main exercise.

1.7.1 Persistence Test and Reversal

Figure 1.9 shows the estimated reduced-form parameters over different windows of analysis, always accumulating 8-trading days. We see a clear pattern. The impact of loan fee in the *tax-arbitrage* transactions on abnormal return happens exactly when we should expect, reaching a peak in the window that includes the record date. For other periods, the impact is hardly distinguishable from zero. As shown in Table 1.7, it is very interesting that when considering cumulative returns from 8 to 15 after the recode date, the estimated coefficient of -0.519 (significant at 10% level), indicating a reversal of almost the same magnitude as the positive abnormal returns around the record date.

Figure 1.10 shows the estimated structural parameters over different windows of analysis, always accumulating 8-trading days. Again, the pattern is clear. The structural parameter, the causal effect of *speculative* transactions' changes in fees on abnormal return happens exactly when we should expect, reaching a peak two days after the record date. For other periods, the impact is not distinguishable from zero, but in the end of the period. Precise coefficients are shown in Table 1.8.

1.7.2 Placebo Test

One may suspect that our empirical findings are a feature of the dividend events, being unrelated to the tax arbitrage mechanism we proposed. In order to disentangle this alternative hypothesis from the tax arbitrage effect, we repeat the same tests around ordinary dividend dates. Since ordinary dividends are not taxable in Brazil, there are no tax arbitrage operations involving stock loans around ordinary dividend dates. Therefore, normal dividend events are ideal to construct a placebo test for the effect of tax arbitrage.

We repeat the same methodology described in Section 1.5, now considering just normal dividends events. We do not consider normal dividends events for which the record date is less than 21 days of IoNE dividend date, since those events would be contaminated by the short-restriction caused by the arbitrage opportunity. All others filters are exactly the same as it was done for IoNE events.

In figure 1.11 and 1.12 we can see that around the record date of the normal dividend there is no particular spike either on the loan fee neither on the short interest.

We present the summary statistics for the normal dividends events in Table 1.9. The results of the reduced form regressions are shown in Table 1.10. It is interesting to notice that in this case Arbitrage Short Interest and Arbitrage Δ Fee are endogenous variables, since they should reflect the willingness to short of investors. Contrary to our main results, here we found that they have negative impact on returns, which is the result normally found in the literature. The interpretation is that the trading activity of informed short sellers is more intense when they have negative information about the stock performance. Recall that in our initial experiment, a higher arbitrage fee (or short interest) indicated an exogenous constraint to short selling. The increased demand for short selling due to tax arbitrage left less room for investors willing to short sell due to speculative reasons. The difference in outcomes between the IoNE and common dividend exercises should reflect the importance of tax arbitrage in generating our main results.

1.8 Conclusions

Despite an intense debate about the impact of short-selling restriction price, the evidence regarding the effects of short sales on asset prices is mixed. The lack of consensus is in large part due to two difficulties. First, identifying a clear source of exogenous variation in supply and demand for stock loan. Second, lack of data availability on lending due to the OTC nature of these markets in most countries. Our study estimates the causal impact of short-selling restrictions on returns by taking advantage of a unique dataset and a unique source of exogenous variation in rental fees. One standard-deviation increase in loan fees causes a 1.11% of returns above the market over 8 trading days. This is strong evidence in favor of Miller's Hypothesis.

1.9 Tables

Tabela 1.1: **Summary Statistics: Brazilian Equity Loan Market**
Number of stocks is the number of different shares traded in stock loan market in a certain year. Volume is the financial volume, price times number of shares, of all stocks lent. Average fee is the mean of stocks daily loan fee, while daily loan fee is a value weighted average loan fee each day for each stock that accounts for the lenders fee and commissions fees. Short-interest is number of shares held in loan contracts normalized by the total number of shares outstanding, this measure is calculated daily for each stock and the average presented is the simple mean.

Year	Num. Stocks	Volume (BRL$ bi)	Volume (US$ bi)	Median Fee (%)	Average Short-Interest(%)
2007	323	272.473	142.106	5	0.611
2008	313	303.505	174.568	3.655	0.540
2009	323	258.912	137.483	2.850	0.365
2010	346	465.605	265.892	2.990	0.427
2011	353	732.750	436.302	2.842	0.675
2012	330	785.927	405.854	2.252	0.849
2013 (until June)	294	500.269	246.665	2.148	1.236

Tabela 1.2: **Summary Statistics for Arbitrage Contracts** Return is defined as stock return minus IBRX50 return cumulated for 8 days beging at two days before the record date. Arbitrage loan contracts are the ones that generated tax benefits - i.e, borrowers are mutual funds and lenders are retail investors or foreign investors. Arbitrage Short-Interest is the total shares involved in arbitrage loan contracts divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the IoNE dividend value as a percentage of the ex-date price. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample.

	mean	min	median	max	std.dev	obs
Cumulative Returns (%)	0.481	-9.757	0.387	13.683	4.378	583
Spec Δ Fee	0.735	-0.859	0.080	19.946	2.126	583
Arbitrage Short Interest	2.458	0.0001	1.300	27.489	3.760	583
Arbitrage Δ Fee	5.080	-0.960	1.123	62.826	9.311	583
IoNE Dividend	0.232	0.001	0.128	2.540	0.312	583
Illiquidity	0.022	0.00002	0.001	1.023	0.083	583
log(MC)	23.230	19.605	22.869	26.722	1.633	583
Turnover	0.281	0.0003	0.216	1.698	0.249	583
Median Fee	2.334	0.283	2.120	10.817	1.967	583
Median SI	1.342	0.003	0.856	14.815	1.643	583

Tabela 1.3: **Reduced Form Results** Arbitrage Short-Interest is the total shares involved in arbitrage loan contracts divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. Robust standard errors corrected for clustering at the ticker level.

	Dependent variable:		
	Cumulative Returns from -2 to 5		
	(1)	(2)	(3)
Arbitrage Short Interest	0.062		0.041
	(0.084)		(0.084)
Arbitrage Δ Fee		0.047***	0.043***
		(0.018)	(0.015)
IoNE Dividend	−0.602	−0.920	−0.925
	(0.642)	(0.635)	(0.656)
Illiquidity	1.303	0.911	1.033
	(2.264)	(2.287)	(2.279)
log(MC)	0.037	−0.054	−0.017
	(0.129)	(0.137)	(0.128)
Turnover	−2.049*	−2.130*	−2.158*
	(1.203)	(1.165)	(1.179)
Median Fee	0.192	0.220*	0.241*
	(0.124)	(0.123)	(0.126)
Median SI	0.161	0.195	0.179
	(0.241)	(0.240)	(0.236)
Constant	−0.511	1.525	0.556
	(3.232)	(3.356)	(3.162)
Observations	583	583	583
R^2	0.016	0.021	0.022
Adjusted R^2	0.004	0.010	0.009
Note:		*p<0.1; **p<0.05; ***p<0.01	

Tabela 1.4: **First Stage.** Arbitrage Short-Interest is the total shares involved in arbitrage loan contracts divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. Robust standard errors clustered at the ticker level in parentheses.

	Dependent variable:		
	Spec ΔFee		
	(1)	(2)	(3)
Arbitrage Short Interest	0.111***		0.070**
	(0.040)		(0.031)
Arbitrage Δ Fee		0.089***	0.083***
		(0.021)	(0.020)
IoNE Dividend	0.961***	0.349	0.340
	(0.347)	(0.216)	(0.218)
Illiquidity	0.574	−0.156	0.054
	(0.610)	(0.501)	(0.508)
log(MC)	0.289***	0.121	0.185*
	(0.102)	(0.085)	(0.097)
Turnover	0.033	−0.130	−0.178
	(0.458)	(0.427)	(0.431)
Median Fee	−0.120**	−0.063*	−0.027
	(0.049)	(0.038)	(0.037)
Median SI	−0.065	−0.003	−0.030
	(0.066)	(0.057)	(0.060)
Constant	−6.134**	−2.419	−4.081*
	(2.399)	(1.964)	(2.272)
Partial F test	7.68	18.55	11.41
Observations	583	583	583
R^2	0.142	0.225	0.237
Adjusted R^2	0.131	0.216	0.227
Note:	*p<0.1; **p<0.05; ***p<0.01		

Tabela 1.5: **Unweighted Second Stage Regressions.** This table shows the estimates of equation 1.4. The dependent variable is the cumulative returns.The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample. Robust standard errors are clustered at the ticker level.

	Dependent variable:		
	Cumulative Returns from -2 to 5		
	One Instrument: SI	One Instrument: Fee	Two Instruments
	(1)	(2)	(3)
Spec Δ Fee	0.560	0.524**	0.530*
	(0.810)	(0.241)	(0.292)
IoNE Dividend	−1.140	−1.103*	−1.109
	(1.050)	(0.666)	(0.688)
Illiquidity	0.982	0.993	0.991
	(2.415)	(2.384)	(2.388)
log(MC)	−0.125	−0.118	−0.119
	(0.253)	(0.163)	(0.173)
Turnover	−2.067*	−2.062*	−2.063*
	(1.214)	(1.193)	(1.196)
Median Fee	0.260	0.253*	0.254**
	(0.168)	(0.130)	(0.128)
Median SI	0.197	0.197	0.197
	(0.246)	(0.244)	(0.244)
Constant	2.921	2.793	2.813
	(5.470)	(3.880)	(4.061)
Weak Instruments	7.68	18.55	11.41
Observations	583	583	583
R^2	-0.035	-0.028	-0.030
Adjusted R^2	-0.048	-0.041	-0.042
Note:			*p<0.1; **p<0.05; ***p<0.01

Tabela 1.6: **Instrument** IoNE is the IoNE dividend value as a percentage of the ex-date price. Log MC is the log of Market Cap. Illiiquidity is defined in Equation 1.3. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median SI is the median short-interest for the whole sample. Herfindahl Index (HI) is the sum of squares of each broker market share. Increases in the index indicate a decrease in competition. SI Before is the average short interest from 21 to 17 days before the event. Robust standard errors corrected for clustering at the ticker level.

	Dependent variable:				
	Arbitrage Δ Fee				
	(1)	(2)	(3)	(4)	(5)
IoNE Dividend	8.565***		7.324***	7.416***	7.423***
	(2.956)		(2.440)	(2.427)	(2.423)
log(MC)		1.831***	1.628***	1.719***	1.739***
		(0.560)	(0.531)	(0.543)	(0.557)
Illiquidity				3.982	3.308
				(3.507)	(3.716)
Turnover				4.160	4.440
				(3.470)	(3.551)
Median SI				−0.488	−0.706
				(0.357)	(0.569)
SI Before					0.254
					(0.438)
HI lenders					1.667
					(6.250)
HI borrower					0.057
					(7.157)
Constant	3.093***	−37.458***	−34.430***	−37.177***	−38.087***
	(0.820)	(12.664)	(12.135)	(12.797)	(12.773)
Observations	583	583	583	583	559
R^2	0.082	0.103	0.162	0.169	0.171
Adjusted R^2	0.081	0.102	0.159	0.162	0.159
Note:					*p<0.1; **p<0.05; ***p<0.01

Tabela 1.7: **Reduced Form: Analyses of Persistence.** The dependent variable is the cumulative in different periods.$Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample.. Robust standard errors corrected for clustering at the ticker level.

	Dependent variable:				
			r		
	(-12 to -5)	(-7 to 0)	(-2 to 5)	(3 to 10)	(8 to 15)
	(1)	(2)	(3)	(4)	(5)
Arbitrage Δ Fee	−0.023	−0.029	0.047***	0.005	−0.051**
	(0.018)	(0.020)	(0.018)	(0.016)	(0.022)
IoNE Dividend	1.129**	0.771	−0.920	1.146*	0.932
	(0.516)	(0.956)	(0.635)	(0.585)	(1.095)
Illiquidity	1.669	4.745*	0.911	3.415	−0.653
	(1.761)	(2.576)	(2.287)	(2.164)	(2.146)
log(MC)	0.204	0.148	−0.054	−0.217	−0.321**
	(0.135)	(0.119)	(0.137)	(0.147)	(0.143)
Turnover	2.470	0.738	−2.130*	−0.921	−0.413
	(1.613)	(1.251)	(1.165)	(1.334)	(1.328)
Median Fee	−0.040	0.172	0.220*	−0.072	−0.260**
	(0.150)	(0.166)	(0.123)	(0.117)	(0.103)
Median SI	−0.416**	−0.087	0.195	−0.177	0.016
	(0.184)	(0.251)	(0.240)	(0.200)	(0.144)
Constant	−4.675	−3.311	1.525	5.277	8.590**
	(3.500)	(2.999)	(3.356)	(3.659)	(3.517)
Observations	579	583	583	580	574
R^2	0.021	0.018	0.021	0.026	0.019
Adjusted R^2	0.009	0.006	0.010	0.014	0.007
Note:					*p<0.1; **p<0.05; ***p<0.01

Tabela 1.8: **Second Stage Regressions: Analyses of Persistence** This table shows the estimates of equation 1.4. The dependent variable is the cumulative returns in different periods. Fee on Non-Arbitrage (Spec) Transactions is instrumented by the delta fee in arbitrage transactions. Robust standard errors are clustered at the ticker level.The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the payout value as a percentage of the ex-date price. Liquidity is defined in Equation 1.3. Log MC is the log of Market Cap. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample.

	Dependent variable:				
			r		
	(-12 to -5)	(-7 to 0)	(-2 to 5)	(3 to 10)	(8 to 15)
	(1)	(2)	(3)	(4)	(5)
Spec Δ Fee	−0.267	−0.326	0.524**	0.058	−0.519*
	(0.226)	(0.244)	(0.241)	(0.186)	(0.270)
IoNE Dividend	1.224**	0.886	−1.103*	1.126*	1.095
	(0.537)	(0.995)	(0.666)	(0.602)	(1.153)
Illiquidity	1.646	4.694*	0.993	3.424	−0.791
	(1.794)	(2.658)	(2.384)	(2.172)	(2.057)
log(MC)	0.236*	0.187	−0.118	−0.224	−0.264
	(0.135)	(0.128)	(0.163)	(0.153)	(0.173)
Turnover	2.437	0.695	−2.062*	−0.912	−0.500
	(1.638)	(1.281)	(1.193)	(1.328)	(1.344)
Median Fee	−0.060	0.152	0.253*	−0.068	−0.283***
	(0.160)	(0.175)	(0.130)	(0.121)	(0.107)
Median SI	−0.415**	−0.088	0.197	−0.176	0.014
	(0.187)	(0.251)	(0.244)	(0.200)	(0.147)
Constant	−5.331	−4.100	2.793	5.418	7.425*
	(3.461)	(3.159)	(3.880)	(3.749)	(4.114)
Observations	579	583	583	580	574
R^2	0.001	-0.008	-0.028	0.026	-0.031
Adjusted R^2	-0.012	-0.021	-0.041	0.015	-0.043
Note:					*p<0.1; **p<0.05; ***p<0.01

Tabela 1.9: **Placebo test using normal dividends payout - Summary Statistics for "Arbitrage Contracts"** Return is defined as stock return minus IBRX50 return cumulated for 5 days beging at two days before the record date. Arbitrage loan contracts are the ones that generated tax benefits - i.e, borrowers are mutual funds and lenders are retail investors or foreign investors. Arbitrage Short-Interest is the total shares involved in arbitrage loan contracts divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts. The method for calculating $Spec\Delta Fee$ is analogous. IoNE is the IoNE dividend value as a percentage of the ex-date price. Turnover is the average daily turnover in the 21 business days before the IoNE dividend record date. The daily turnover is the traded volume in a day normalized by the market capitalization. Median loan fee is the median of average daily loan fee in the whole sample, whereas average daily loan fee is the value weighted average loan fee in a day. Median SI is the median short-interest for the whole sample.

	mean	min	median	max	std.dev	obs
"Spec Δ Fee"	-0.001	-0.743	-0.005	1.850	0.291	255
"Arbitrage SI"	1.062	0.0005	0.716	13.955	1.358	255
"Spec Δ Fee"	0.003	-0.837	-0.027	2.705	0.401	255
IoNE Dividend	0.601	0	0.271	5.230	0.822	255
Illiquidity	0.008	0.00001	0.001	1.193	0.076	255
log(MC)	22.883	19.753	22.877	26.108	1.016	254
Turnover	0.398	0.001	0.314	1.509	0.313	255
Median Fee	2.629	0.276	1.942	18.565	2.762	255
Median SI	2.049	0.005	1.193	14.815	2.395	255

Tabela 1.10: **Placebo test using normal dividends payout - Reduced Form Results** "Arbitrage Short-Interest"are loan contracts among arbitrage groups for the IoNE dividends divided by the total shares outstanding in the record day. $Arbitrage\Delta Fee$ is the percentage increase of the loan fee in the event window for arbitrage contracts in the case of IoNE dividends. The method for calculating $Spec\Delta Fee$ is analogous. Normal Dividend is the payout value as a percentage of the ex-date price. Turnover is the average daily turnover in the 21 days before the IoNE dividend event. Average daily loan fee is the value weighted average loan fee in a day. Median loan fee is the median of average daily loan fee in the whole sample.Liquidity is defined in Equation 1.3. Robust standard errors corrected for clustering at the ticker level. Median SI is the median short-interest for the whole sample.

	Dependent variable:				
	Cumulative Returns from -2 to 5				
	Reduced Form			OLS	
	(1)	(2)	(3)	(4)	(5)
"Arbitrage Short Interest"	−0.697**		−0.689**		−0.705**
	(0.287)		(0.291)		(0.300)
"Arbitrage Δ Fee"		−0.858	−0.826		−1.256
		(1.018)	(1.005)		(1.264)
"Spec ΔFee"				−0.075	1.043
				(0.871)	(1.225)
Normal Dividend	−0.652	−0.544	−0.590	−0.603	−0.637
	(0.422)	(0.430)	(0.434)	(0.433)	(0.439)
Illiquidity	−1.160	−1.004	−1.275	−0.881	−1.350
	(1.188)	(1.170)	(1.173)	(1.184)	(1.226)
log(MC)	−0.247	−0.174	−0.217	−0.203	−0.218
	(0.357)	(0.349)	(0.353)	(0.354)	(0.354)
Turnover	−1.080	−1.383	−1.110	−1.357	−1.094
	(1.432)	(1.486)	(1.446)	(1.480)	(1.429)
Median Fee	0.108	0.137	0.108	0.138	0.101
	(0.096)	(0.098)	(0.095)	(0.099)	(0.101)
Median SI	0.255	−0.036	0.239	−0.023	0.248
	(0.299)	(0.246)	(0.295)	(0.253)	(0.303)
Constant	6.932	5.099	6.255	5.762	6.302
	(8.467)	(8.273)	(8.373)	(8.389)	(8.384)
Observations	254	254	254	254	254
R^2	0.047	0.036	0.052	0.031	0.055
Adjusted R^2	0.020	0.009	0.021	0.004	0.020
Note:					*p<0.1; **p<0.05; ***p<0.01

1.10 Figures

Figura 1.2: **Distribution of Loan Fees By Firms by Year**
Figure shows the distribution of the average daily value-weighted loan fees in percentage points per year between January 2007 and June 2013. The vertical axis shows the frequency of firms with loan fees in the interval reported on the horizontal axis.

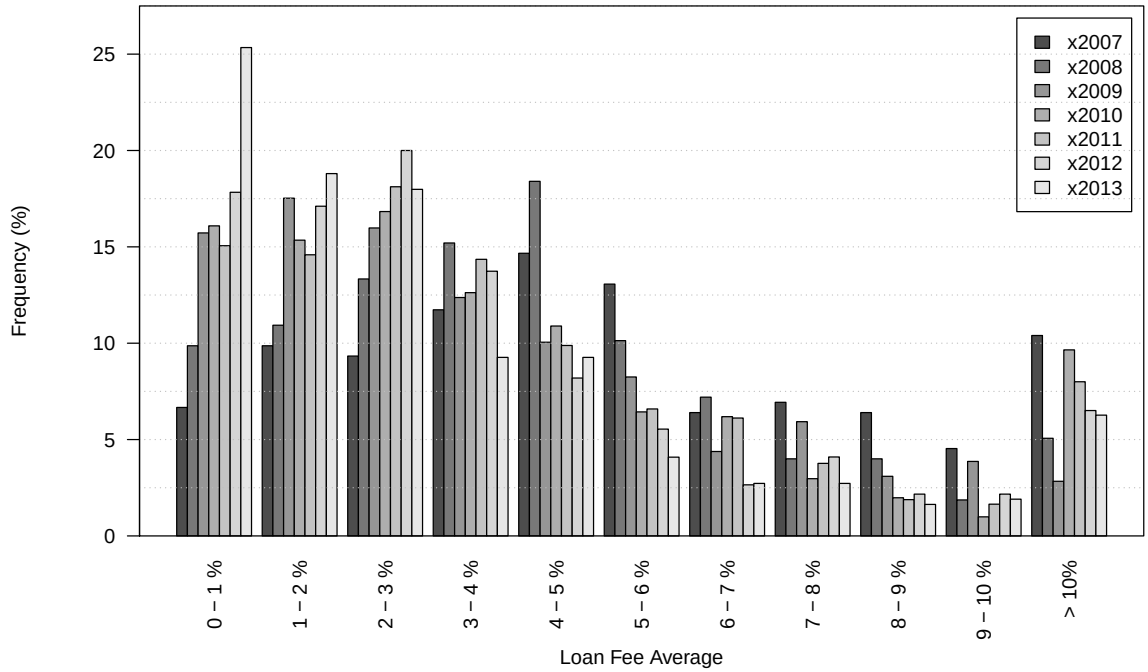


Figura 1.3: Payout Announcement

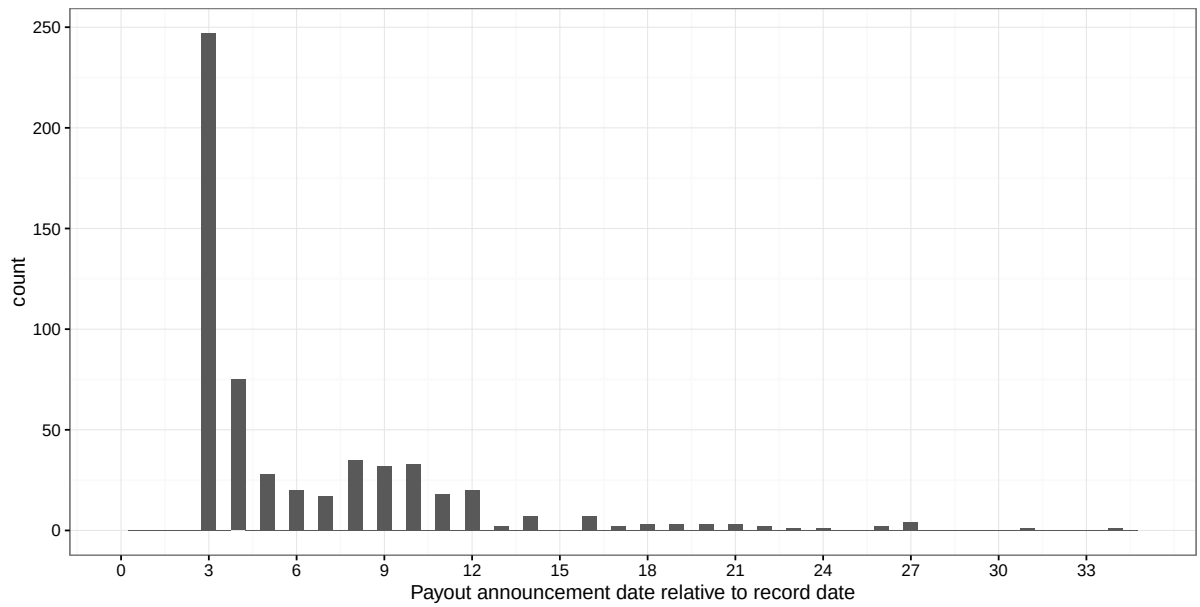


Figura 1.4: **Loan Fee around Record Date of an IoNE dividend event**
Figure shows the average loan fee around the record date of the IoNE dividend event. The daily loan fee for each share is the value weighed fee among all contracts in a certain day. The average loan fee is the common mean among shares. We consider 487 IoNE dividend events from January 2010 until June 2013.

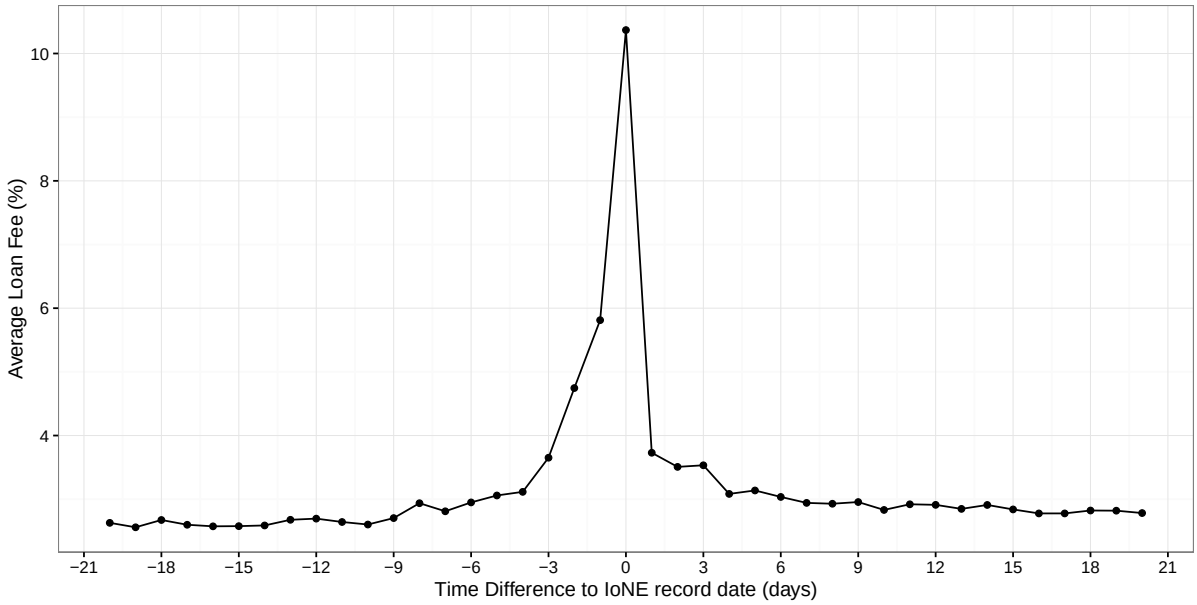
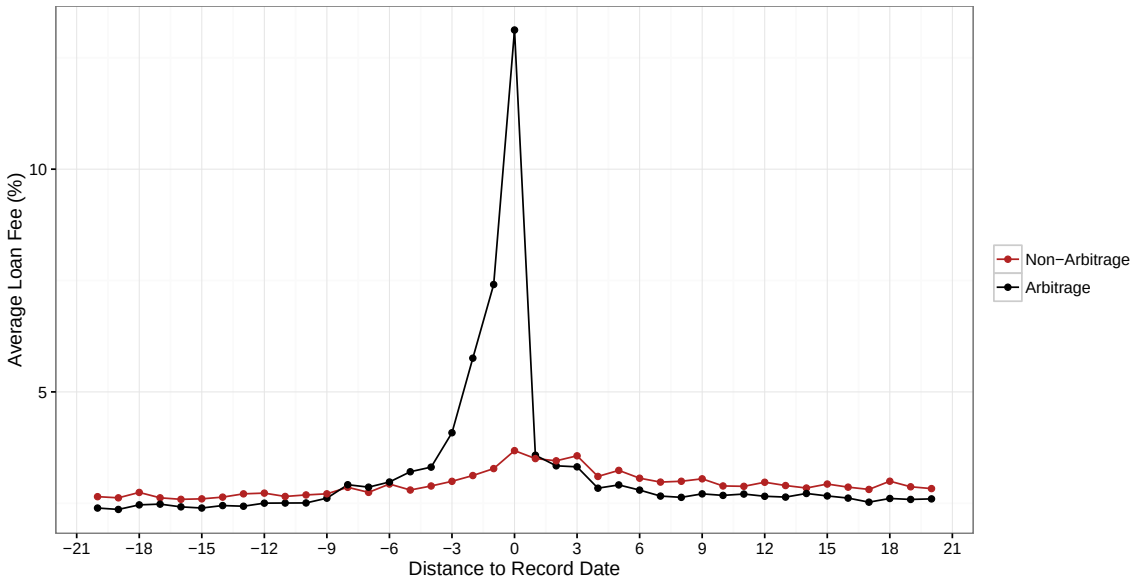


Figura 1.5: **Loan Fee around Record Date of an IoNE dividend event by Groups**
Arbitrage loan contracts are the ones that have tax benefits - i.e, borrowers are mutual funds and lenders are retail investors or foreign investors that took place before the record date and were liquidated after it. Speculative contracts are the other contracts. We define the daily loan fee as the value weighed average loan fee for each stock and each day around the IoNE date. The figure shows the average loan fee among all stocks for each day around the IoNE record date.

(a) Loan Fee for Groups of Contracts



(b) Loan Fee for Speculative contracts

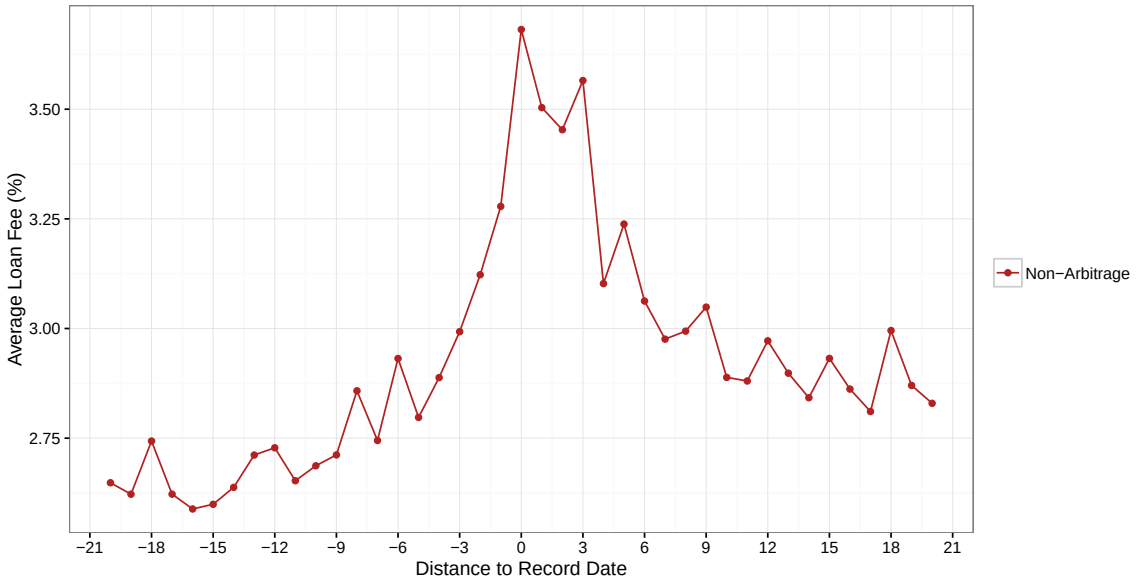


Figura 1.6: **Short-Interest around Record Date of an IoNE Event**
Figure shows the average short-interest around the record date of the IoNE dividend event. The average short-interest is the common mean among shares.

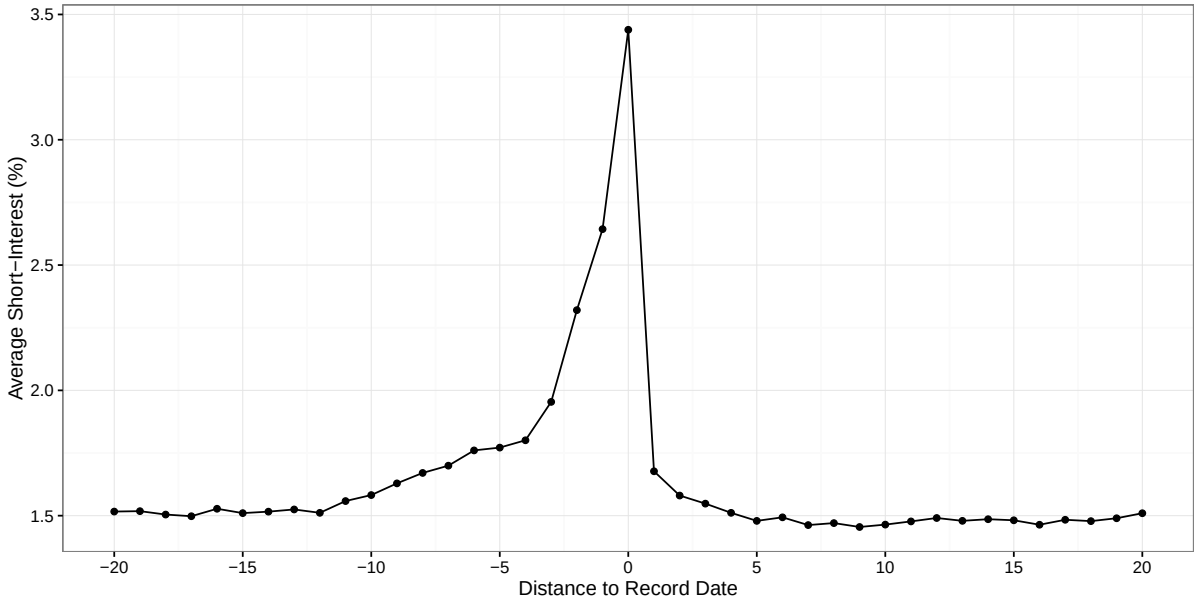
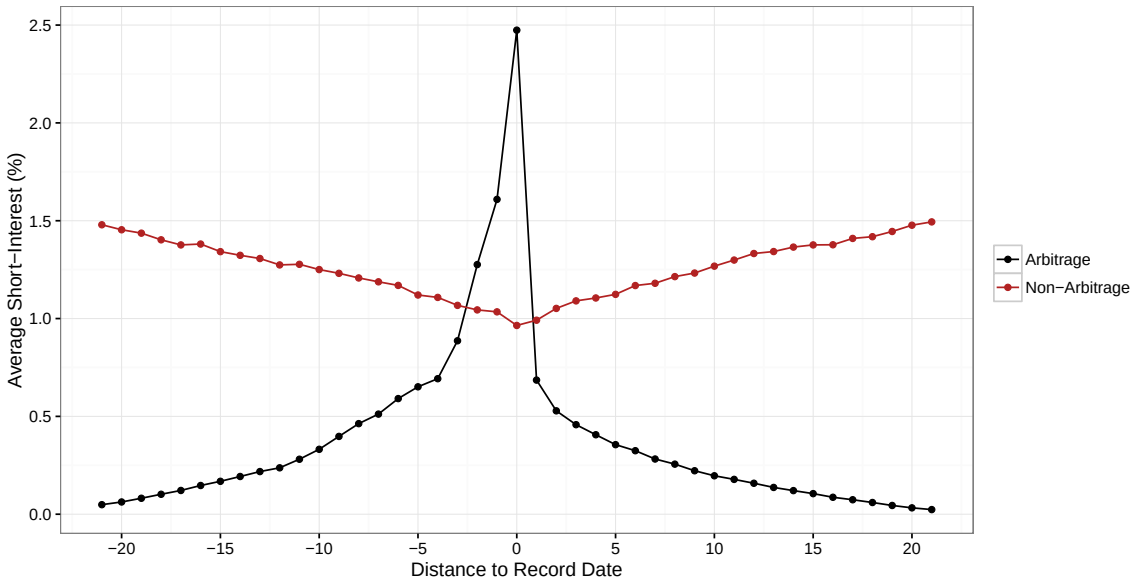


Figura 1.7: **Short Interest around Record Date of an IoNE dividend event by Groups** Arbitrage loan contracts are the ones that have tax benefits - i.e, borrowers are mutual funds and lenders are retail investors or foreign investors that took place before the record date and were liquidated after it. Speculative contracts are the other contracts. We define the daily loan fee as the value weighed average loan fee for each stock and each day around the IoNE date. The figure shows average short interest among stocks for each day around the IoNE dividend record date.

(a) Short Interest for Groups of Contracts



(b) Loan Fee for Speculative contracts

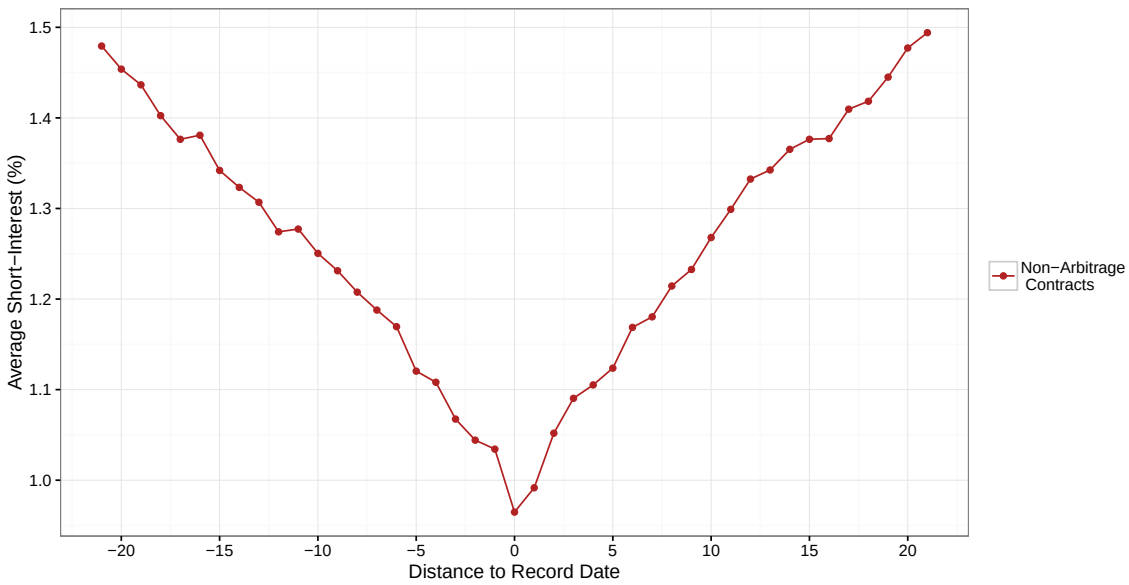


Figura 1.8: **Abnormal Returns around the IoNE Dividend Date**
Figure show the cumulative abnormal returns starting from 10 days previous the record date. Abnormal return is calculated as the stock return minus the stock loading on the market portfolio IBX50, that accounts for the first 50 biggest stocks in market capitalization. The dotted line is the 95% percent confidence interval.

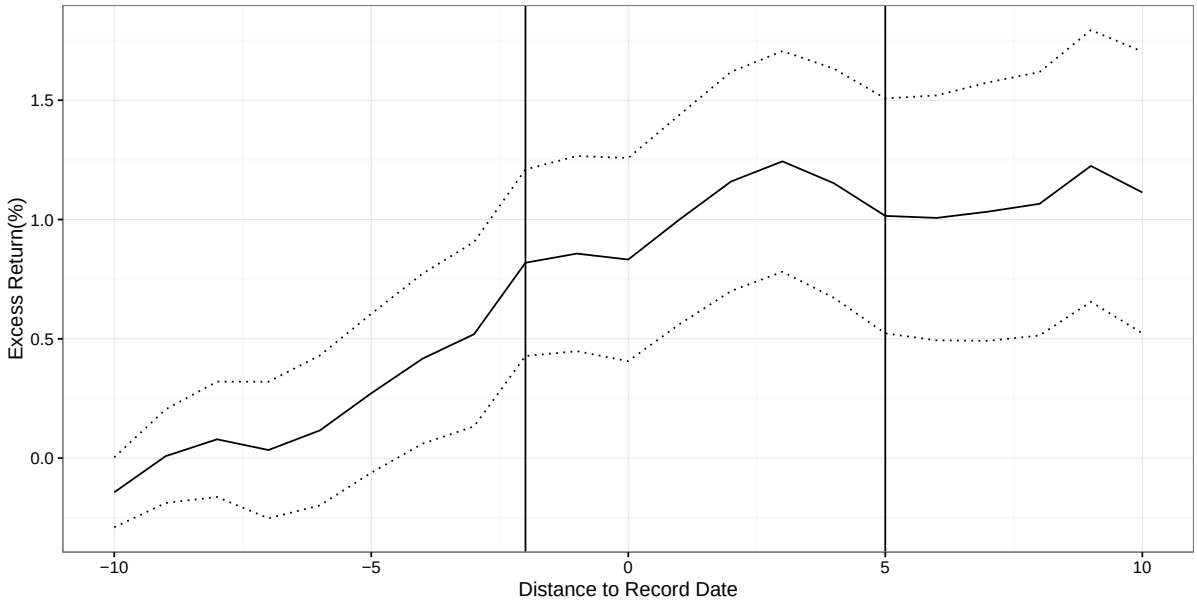


Figura 1.9: **Reduced Form Coefficient for Different Return Windows**
Figure shows the estimated coefficient for different periods of 8-days cumulative returns. X-axis indicates the final date of 8-days cumulative returns, y-axis is the coefficient of $Arb\Delta Fee$ in reduced form regressions when considering as independent variables $Arb\Delta Fee$ and controls. The dotted line is the 95% percent confidence interval.

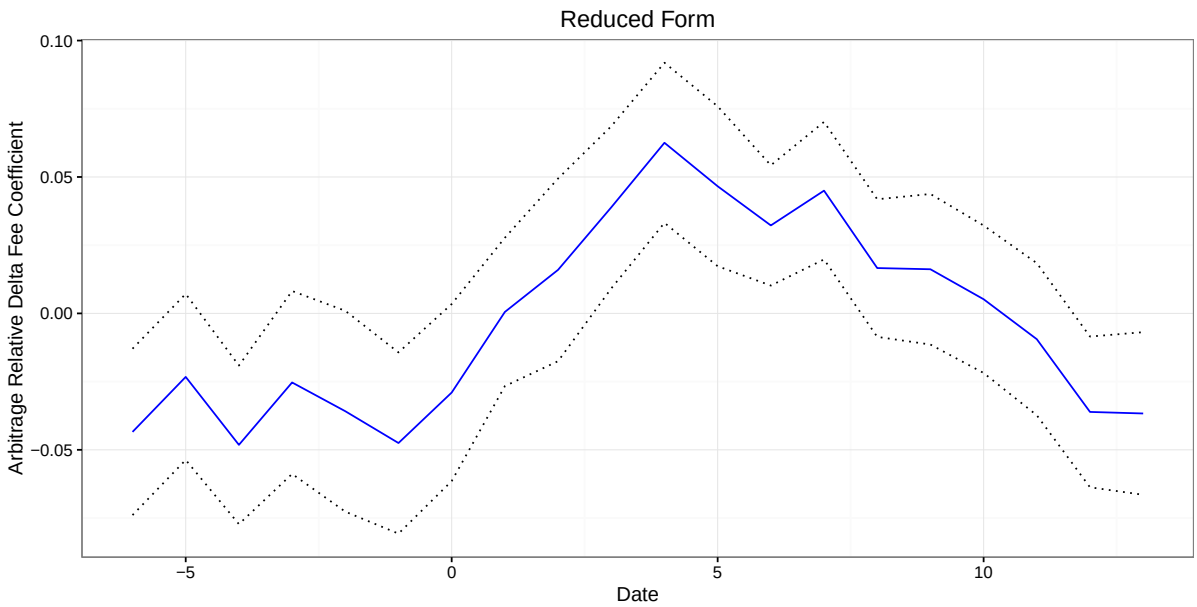


Figura 1.10: **Second Stage Coefficient for Different Return Windows**
Figure shows the estimated coefficient for different periods of 8-days cumulative returns. X-axis indicates the final date of 8-days cumulative returns, y-axis is the coefficient of $Spec\Delta Fee$ in structural form regressions when considering just one instrument: $Arb\Delta Fee$ and controls. The dotted line is the 95% percent confidence interval.

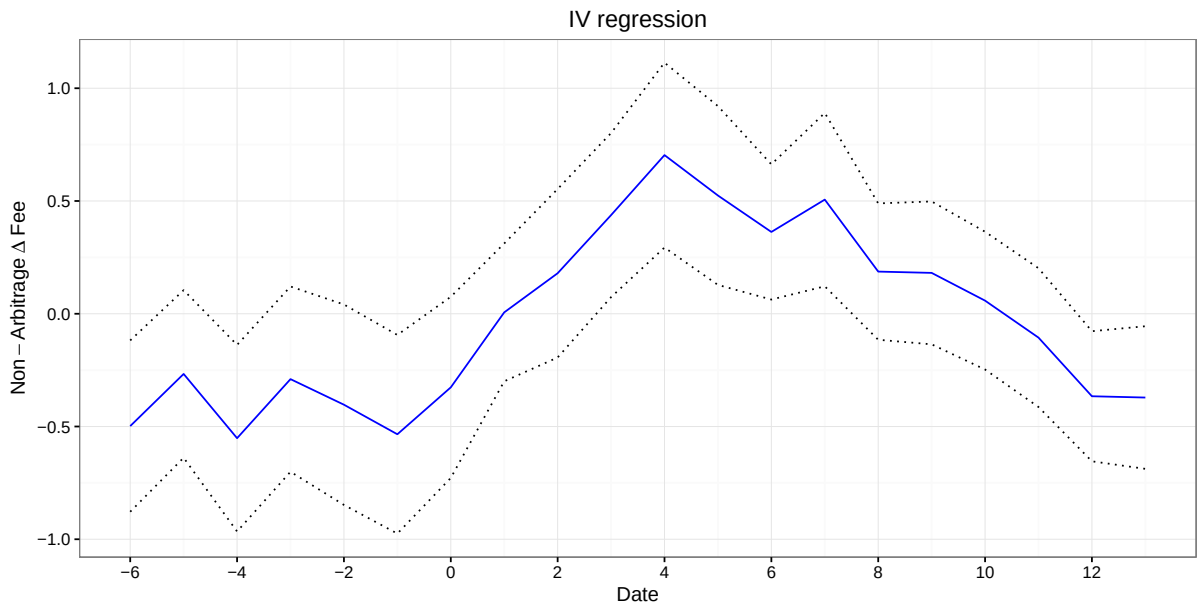


Figura 1.11: **Loan Fee around Record Date of an normal dividend event**
Figure shows the average loan fee around the record date of the normal dividend event. The daily loan fee for each share is the value weighed fee among all contracts in a certain day. The average loan fee is the common mean among shares. We consider 487 IoNE dividend events from January 2010 until June 2013.

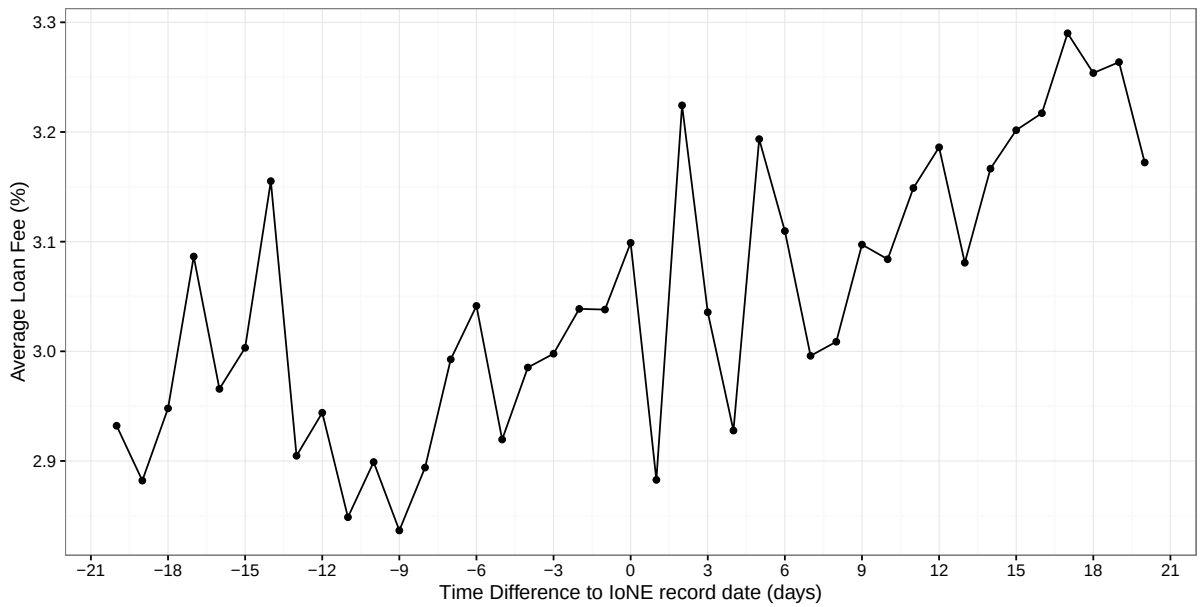


Figura 1.12: **Short-Interest around Record Date of an IoNE Event**
Figure shows the average short-interest around the record date of the IoNE dividend event. The average short-interest is the common mean among shares.

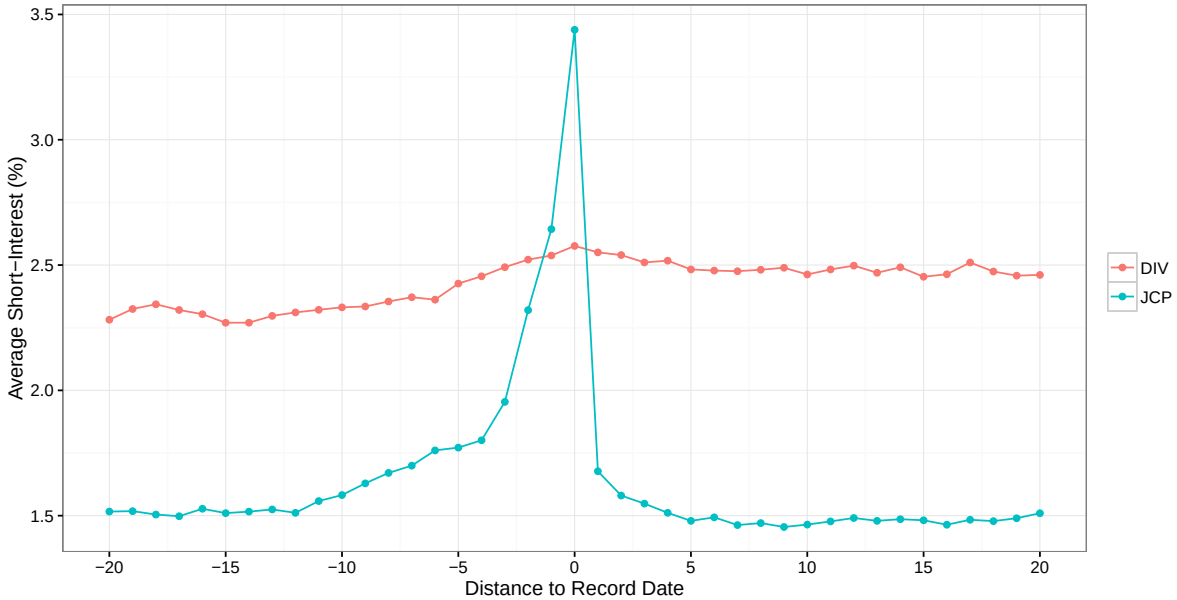
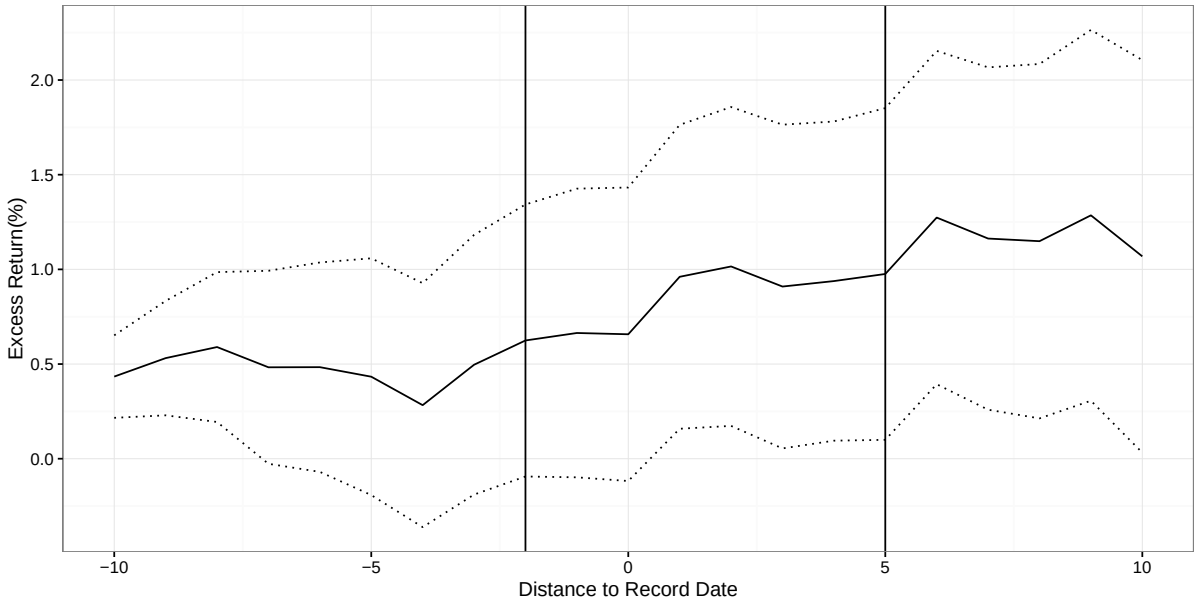


Figura 1.13: **Abnormal Returns around the Normal Dividend Date**
Figure show the cumulative abnormal returns starting from 10 days previous the record date. Abnormal return is calculated as the stock return minus the stock loading on the market portfolio IBX50, that accounts for the first 50 biggest stocks in market capitalization. The dotted line is the 95% percent confidence interval.



Capítulo 2

Loan Fee Dispersion and the Cross-Section of Returns [1]

ABSTRACT

The lack of transparency of stock loan OTC market generates loan fee dispersion across loan contracts for the same stock at the same day. Using an unique database that unifies all stock loan transactions in Brazil from January 2007 to June 2013, we show that loan fee dispersion is the best predictor of the cross-section of stock returns, when compared to traditional short-sale related measures in the literature: loan fee average, short-interest or days to cover. We document a newer version of the "short premium": the small dispersion-minus-large dispersion (sDmlD) portfolio of stocks has a monthly average excess return of 1.03% and a 0.79% four-factor alpha.

[1] This is a joint work with Marco Bonomo and Fernando Barbosa.

Recent literature has extensively addressed the predictive power of short sale related metrics over returns. Measures like short-interest [2], institutional ownership [3], loan-fee [4] and days-to-cover (DTC) [5] were shown to be good forecasts of the cross-section and time-series of returns. Overall, the suggested interpretation of those results is that stocks that are more short-sale constrained have lower future returns than less constrained ones, the so called short premium. Although several authors have argued that short sale constraints reduce the price efficiency [6], there is still no consensus on what would be the appropriate measure of short-sale constraints. In particular, there is little discussion on the interaction between short-sale constraints and the risk involved in short-selling activity.

The stock lending market is most often opaque, the over the counter (OTC) structure gives rise to problems such as search frictions and imperfect competition. In this context the law of one price is not expected to hold. In this paper we document that there is price dispersion in loan fee, i.e., for contracts with the same conditions for the same stock in the same day stock investors pay different loan fees. Furthermore, we document time variation on the loan fee dispersion. Panel regressions show that the loan fee dispersion is correlated with characteristics of the stock and the underlying market structure statistics. More specifically, the loan fee dispersion is higher for stocks that are smaller in market capitalization, have more volatile returns, and higher turnover. When we consider the characteristics from the stock loan market, loan fee dispersion is positively correlated with short-interest and lender brokers concentration and negatively correlated with borrowers brokers concentration. We also calculate the network degree for lenders and borrowers of each stock. We find that loan fee dispersion is higher when market concentration is higher, i.e. when the lenders Herfindahl index is higher. Finally, we show that the loan fee dispersion is the best predictor of the cross-section of asset returns when compared to other short-selling related measures.

We present three possible explanations for why loan fee dispersion should forecast returns. The first one is related to search friction and trading cost. Higher search costs are associated with higher loan fee dispersion. Therefore agents will require a premium for short selling these stocks with higher search costs. The second concerns the additional risk of the short position. If the loan contract ends (or is terminated) but the investor still wants to hold a short position, there is the risk of not finding a new loan at a reasonable price. The third explanation is that loan fee dispersion reflects heterogeneous beliefs about stock returns. If agents with inside information are willing to short sell as fast as possible, at any cost, a higher loan fee dispersion could indicate the presence of well informed arbitrageurs and consequently predict lower future prices. All three mechanisms will be addressed in more detail later. Before that, a literature review must be provided.

Duffie et al. (2002) present a theoretical framework for stock lending market dynamics in the light of search frictions. They develop a dynamic model in which search frictions limit the frequency at which share lenders and borrowers are able to find one another. If lenders have some bargaining power they are able to charge loan fee between zero and the reservation price of the short seller. At the same time, short sellers are willing to pay the fee because, if they refuse, they might not be able to find another lender before the termination time when the fundamental price is uncovered and would lose their expected surplus. In other words, search frictions allow lenders to charge a positive fee; consequently, investors are willing to pay more than their valuations due to the expectation to profit from lending the stock.

Kolasinski, Reed, and Ringgenberg (2013) empirically study some related issues. Using data from 12 lenders, they study the determinants of the supply and dispersion of fees in equity lending market. They find that both loan supply curve and loan fee dispersion are increasing

[2] Figlewski (1981)
[3] Asquith, Pathak, and Ritter (2005b)
[4] Drechsler and Drechsler (2014)
[5] Hong, Li, Ni, Scheinkman, and Yan (2015)
[6] Asquith et al. (2005b), Saffi and Sigurdsson (2011) and Boehmer and Wu (2013)

in the search cost proxies [7]. They also found that substantial variation in the lending fees charged by the same lender, and showed that the in-lender dispersion in lending fees is positively related to the stock average lending fee. However, their data do not allow them to investigate the determinants of different lending fees paid by borrowers. Chague, De-losso, De Genaro, and Giovannetti (2016), using a Brazilian data base similar to ours, proposes a proxy for search costs at the borrower-stock-day level, constructed through the connectedness in the network formed by borrower with brokers and lenders of the stock. They found that for the same stock, on the same day, less connected borrowers - the ones with higher search costs - pay higher loan fees. This relation holds even for the same broker, who discriminates borrowers according to their connectedness, charging a higher lending fee from the less connected borrowers.

The next step to give reasoning to the finding that loan fee dispersion forecast negative returns in the cross section is to understand how short sale constraint can affect prices. In the model of Drechsler and Drechsler (2014) stocks that have high loan fee also have higher prices, because the additional expected rent from lending out should be incorporated into prices. Though, Drechsler and Drechsler (2014) show that there is short premium even net of fees, hence the loan fee different than zero are not sufficient to explain the short premium encountered on the data.

An important dimension of the cost of short-selling is the additional risk engendered by the opacity of the stock loan market. The typical stock loan contract is short termed and re-callable, i.e., the lender of a stock has the right to close the loan contract at any time. If the loan contract is terminated while the short position still open, the short seller must find a new loan that might have different fee. Notice that if the short seller is incapable of finding a new loan or its cost are prohibitive, she will be forced to close the short position and might incur in losses. The risk of not finding a new loan at reasonable price is non-diversifiable because the hedging asset would have to be correlated with the loan fee but also with the availability of stocks to rent. On top of that, the loan fee is the cost payed by the short seller and affects its wealth. Therefore we would have a non-diversifiable risk that is correlated with wealth, hence priced in the cross-section. A formal model that presents this idea is described in the appendix of Drechsler and Drechsler (2014).

Another explanation dates back to Miller (1977), he argues that the interaction between heterogeneous valuations and short sale constraint leads to prices superior to the average valuation. When investor opinions are disperse, to restrict short selling is by definition to exclude pessimist from the market. Consequently, just the beliefs of the more optimistic investors are reflected on prices. In this case, short selling constraints causes the overprice of constrained stocks, therefore we expect future negative returns when price converges to fundamentals. [8]

Empirically, many short related measures were also shown to be good forecasts of returns. Short interest (SI), defined as the ratio of shares shorted to the shares outstanding, is the most popular metric of short selling activity. Papers typically find that stocks with a high SI significantly under-perform stocks with a low one. Hong et al. (2015) recently introduce a new measure: days to cover (DTC), defined as SI divided by the average daily turnover. They show that DTC captures the cost of entering crowded trades. According to this proposed measure two stocks with the same SI may have different associated costs, which would be higher for the stock with lower turnover. Consistent with their theory, they found that a strategy long low DTC decile stocks and short high DTC decile stocks yields 1.19% per month with a t-statistic of 6.67. Finally, Drechsler and Drechsler (2014) used the loan fee level as method to access the risk to short. They find that high fee stocks earn lower returns than low fee stocks, both gross and net of fees. Moreover, even when the cost of these fees is taken into account, short-sales constraints appear unable to explain anomalies because their average returns significantly exceed their fees.

[7] For a nice literature review on search cost and price dispersion please refer to Baye, Morgan, and Scholten (2006).

[8] Chen, Hong, and Stein (2002b) show that even if short is allowed mispricing can prevail if at least one group of investor are not allowed to short.

They also find that the long-short portfolio based on loan fee sort is strongly related to 8 well known anomalies in the cross-section, including value, profitability and momentum.

All the introduced measures are "average"measures and they do not consider an important characteristic of the market to borrow stock: it is opaque. Stock loans transactions are over the counter (OTC) and information is disperse. For instance, usually SI and DTC is a public available metric, whereas loan fee is not. In most financial markets, lending services are provided over the counter by big custodian banks, and there is no institution that centralizes the registration of the loan contracts. In this context, the loan fee average, for example, misses an important dimension of the security loan markets, since a small percentage of trades are done at this rate. The loan faced by a borrower will depend on many other factors that may include her search costs, her access to different brokers, the supply concentration of the specific stock, and brokers network. The loan fee dispersion is a simple measure that captures the uncertainty about short cost.

The OTC market configuration makes data aggregation virtually impossible, and many times jeopardizes the reliability of the results of papers that tried to address this manner. This paper uses a unique database to circumvent the difficulties with data and identification. In Brazil the equity lending market transactions are registered in a centralized platform - the only stock exchange presently operating in Brazil, the BM&FBovespa. Thus, we observe all rental transactions, which mitigates possible biases due to selective data availability. Our data is recoded at transaction level.

Finally, we study the cross-section of stocks by loan fee dispersion using portfolio sorts on short sale characteristics. We sort stocks into three quantiles - small (bottom 30%), middle (40%), large (top 30%) - based on their loan fee dispersion at the end of each month. We examine the average returns on these quantiles portfolios over the following month. The average return on the large quantile is -0.58% per month, while the average return on a portfolio long the stocks small quintile and short the stocks in the large quantile portfolio (the large dispersion-minus-small dispersion (lDMsD)) is a highly significant 1.03% per month, with a four factor (FF4) alpha of 0.79% per month.

For comparison, we repeat the same methodology for each of the short measures described above: average loan fee, short interest and days to cover. We observe that the long-short portfolio based on loan fee dispersion is the one with higher alpha and the only one where the alpha is statistically significant.

We are also interested on the predictive power of loan fee dispersion on returns when controlling for the other short measures. We test our hypothesis using the Fama-Macbeth regression methodology. We find that one loan fee standard deviation forecasts -1.19% returns in the following month. Most import, though, is that loan fee dispersion is the the best predictor among all the short measures discussed.

The remaining of the paper is organized as follows: Section 2.1 describe the data used as well and summary statistics, Section 2.2 describes the loan fee dispersion, Section 2.3 connects the loan fee dispersion with the cross-sectional returns of stocks, Section 2.4 concludes.

2.1 Data, Variables and Summary Statistics

Unlike most countries, where the market for renting stocks is completely decentralized, in Brazil, every transaction must be reported to BM&FBovespa. That means, although the market is still opaque from an investor perspective, the stock exchange acts as clearing and keep the margins. Consequently, the Brazilian framework configures a unique opportunity to academic research on securities lending and short selling. We observe the whole process: offers, contracts, and liquidation. We also observe information about investors, brokers, and maturity. Our data run from January 2007 to June 2013. A typical observation contains:

1. Date when the loan contract was open/closed;

2. Stock ID;
3. Type of investor lending the stock, e.g. Mutual Fund, Retail Investor, etc.;
4. Unique anonymous code identifying the brokers intermediating the transaction;
5. Type of investor borrowing the stock;
6. Number of shares;
7. Lending fee;
8. Broker's commission rates.

Information on stocks returns and volume are from Economática. In conformance with CRSP returns, we consider returns are holding period returns from month-end to month-end, with ordinary dividends reinvested at month-end.

To be included in the final data set, we require that the stock have a valid price all days in the month, have valid turnover measure and short-interest measure. In order to consider a daily short-interest measure valid, we require that at least 5 contracts were closed that day. The monthly measure is considered valid only if there were at least 10 trading days with a valid daily measure. [9] For each contract we consider the loan fee as the fee required by the owner of the stock plus the commissions fees charged by the intermediaries brokers[10]. All loan fee variables are annual rates.

Our main variable of interest is monthly *SD Fee*. It is calculated as the average of daily standard deviation. To be specific, the dataset contains the loan fee charged for every contract initiated in a day. With that information, we calculate the daily standard deviation. *SD Fee* is simply the monthly average of daily standard deviation. We opt for averaging daily standard deviations rather than calculating the monthly standard deviation to reduce the impact of fees time-series volatility and to focus on cross-sectional variation.

Similarly, the daily average loan fee is the volume weighted daily average loan fee in each day. The monthly average loan fee(*Avg Fee*) is the average daily loan fee in the month. Daily short interest (SI) in the quantity of shares in loan contracts divided by the number of shares outstanding. Monthly SI is the average SI in the month. In the same fashion, daily days to cover (DTC) is the daily SI divided by the daily turnover, whereas monthly DTC is the average DTC in the month.

Table 2.1 reports summary statistics for the main variables in our study. Average Short Interest is 2.23% with a 2.58 standard deviation. Our main variable, the standard deviation of shorting fees(*Std Dev Fee*), presents mean equal to 0.927 ranging between 0.045 and 12.192 . Summary statistics for other common measures of short selling restriction, such as *Days to Cover*, are also presented. Table 2.3 reports yearly summary statistics. Note how the yearly average of *Std Dev Fee* appears to be stable. On the other hand, the average *Short Interest Ratio* changes considerably between years. Finally, Table 2.4 reports the correlation matrix between different measures of short sale constraints.

2.2 Loan Fee Dispersion

It should be no surprise that we find evidence of loan fee dispersion, after all we are looking at an opaque market. However, due to the difficulty in finding data, the distribution of fees have not

[9] We also exclude dates that are at least 5 days close to a "interest on net equity"(IoNE) window due to a tax arbitrage that could distort loan fees and SI interest in this period. IoNE is a type of dividend which receives tax treatment similar to interest payments, both form firms and investors. Investors were subject to 15% tax, but domestic funds were tax exempt. Thus, there was a tax arbitrage operation, which consisted of funds borrowing from other type of investors before the dividend record date, reverting the loan just after it.

[10] The ability to distinguish both is one other advantage of our data set.

received much attention in the short selling literature. In this section we describe fee dispersion and the underlying mechanisms driving it.

Observing only the daily average fee means ignoring important aspects of market structure. In figure 2.1 we see the histogram of fees for Petrobras on April 30, 2008 and April 1, 2009. In both days, the average loan fee was very similar 0.71% and 0.75% respectively. However, the distribution was vastly different, with interquartile range being respectively 0.09 p.p. and 0.66 pp. As figure 2.1 shows, the selected date in 2009 depicts much more dispersed fees than in April 2008. Keep in mind that here we are looking at days with almost the same average loan fee and at loan contracts for the exact same stock.

Part of the dispersion is explained by price discrimination. In order to have a broad picture of loan fees received and charged by different types of investor, we calculate the loan fee spread between loan contracts between different types of investors. We take the mutual fund as our benchmark investor, and calculate the loan fee spread as:

$$\text{Relative Spread}_{i,t,p} = \frac{\text{Loan Fee}_{i,t,p} - \text{Loan Fee}_{i,t,\text{fund}}}{\text{Loan Fee}_{i,t,\text{fund}}} \tag{2.1}$$

Figures 2.3 and 2.2 plot the time series of the spread between fees paid (received) by mutual funds, retail investors and foreign investors. It is clear that mutual funds pay lower loan fees than retail investors when they are borrowing, and receive higher loan fees when they are lending. The spread between mutual funds and foreign investors is smaller. However, it is interesting to notice the widening of the gap between fees paid by retail investor and foreign investors, specially after 2008. According to Chague et al. (2016), mutual funds have access to greater number of brokers and potentially more connected ones than retail investors, which is consistent with them paying lower fees.

Yet, it is interesting to notice that inter-group price dispersion is also relevant. In section 2.3 we show that loan fee dispersion is the best predictor of cross sectional returns among stock loan market statistics. Notice that the intra-group price dispersion rather than inter-group could be the main driver of the results we obtain. If we restrict the sample to only contracts where the borrower is a mutual fund the new *SD Fee* will have 0.9 correlation with the old measure.

Table 2.2 uses a panel regression to illustrate empirical properties of the fee dispersion statistics. All three columns represent regressions with time and individual fixed effects. The covariates in the first column are spot market statistics, e.g. price volatility and market capitalization. The regression in the second column, on the other hand, aims to get a better picture on how the OTC market structure relates to the fees dispersion in that same market. It is intuitive that a higher volatility should be associated with higher dispersion in charged fees. We find no significant relation whatsoever between book to market and loan fee dispersion. Interestingly , turnover correlates positively with our dependent variable. This may reflect that both turnover and the loan fees dispersion could move with dispersion of beliefs. To measure market concentration we use the Herfindahl Index (HI), which is the sum of squares of the market shares of the brokers. Increases in the index indicate a decrease in competition and an increase of market power. Our last covariate, *Degree*, relates to network theory. It is simply the number of different brokers lending (borrowing) that stock in the last month.

Chague et al. (2016) using a different extraction from the same database we use, find compelling evidence that higher search costs are associated with higher fee dispersion. They create a measure of borrower-stock connectedness(BC), as a proxy for the search cost an investor faces when borrowing a specific stock. They also show that BC is positively related with loan fee dispersion: "loan fee standard deviation and loan fee range among borrowers in the low BC-group are respectively 46% and 135% higher than those among borrowers in the high-BC group.

Figure 2.4 depicts the time series of aggregate measures. The green and more stable line represents the aggregate *SD fee*. Table 2.3 reports yearly summary statistics. Note how the yearly average of *SD Fee* appears to be more stable than the average *short interest* which changes

considerably between years.

2.3 Loan Fee Dispersion and Returns

In this section we analyze how loan fee dispersion affect returns. Our hypothesis is that, everything else equal, the higher the loan fee dispersion is, the lower are future returns. We have three underlying reasons for this hypothesis : (i) the higher the loan fee dispersion the larger the search costs; (ii) loan fee dispersion captures short-selling risk, (iii) loan fee dispersion reflects heterogeneous beliefs about the stock returns. We are unable to differentiate between those possible mechanisms, since all of them are consistent with the predictive power of loan fee dispersion on the cross-section of returns.

In this section we first present the time series results of portfolio formed on loan fee dispersion sorts, than we present the Fama and MacBeth regressions.

2.3.1 Sorts on Loan Fee Dispersion

In this section, we show that stocks sorted on loan fee dispersion generate larger and more significant return spreads than stocks sorted other short related measures. We conduct the test as follows. At the end of each month from January 2007 to June 2013, we independently sort all valid stocks in three quantiles based on their loan fee dispersion, average loan fee, short interest and days to cover. For each short characteristic we form value weighted portfolios based on the market capitalization of the end of that month. Holding period is one month. We also form a zero-investment long-short portfolio that buys the low characteristics portfolio and sells the high characteristics. Since all characteristics considered are related to short sale constraints, our hypothesis is that the long-short portfolios have positive average returns.

Table 2.5 presents value-weighted average returns and characteristics for each portfolio over the sample. In Table 2.6 we can see that the annualized Sharpe Ratio of the long short portfolio constructed based on fee dispersion is 0.742 , with average annual returns equals to 12.4%, performing better than the remaining long short portfolios based on usual measures of short selling restriction. The cumulative returns of all strategies is plotted in Figure 2.6.

Now we consider the returns of the long-short portfolio when hedging out the exposure of the market, Fama-French portfolios, momentum portfolios and liquidity portfolio. Those portfolios have been called risk factors in the literature of asset pricing. We use them as benchmarks and as usual, we call them factors. The factors returns are extracted from NEFIN. Tables 2.8, 2.9, 2.10 and 2.11 show the time series regressions of each portfolio on the five factors. We observe that in the case of the loan fee dispersion long-short portfolio the four-factor alpha is 0.79% with t-stat of 2.019. We repeat the exercise for all short measures. Loan fee dispersion is the single measure that has a significant alpha.

For comparison, we also report the same results for equal weighted portfolios.

A concern in the analysis of long-short portfolios is the turnover of each portfolio. Our portfolios are rebalanced every month. To be more specific, each month stocks are sorted accordingly to their loan fee dispersion. If the stocks relative position did not vary through time, turnover would be very low. The opposite would be true if the loan fee dispersion was very instable. For this end we calculate the transition matrix of each portfolio. Figure 2.7 shows that if the stock is at portfolio with the lowest loan fee dispersion, sd_fee1, the probability to remain at same portfolio for the next month is 69%, to be placed on portfolio sd_fee2 is 26% and to jump to sd_fee3 is 5%. In other words, 31% of the stocks belonging to portfolio sd_fee1 on a given month are expected to change portfolio in the next month.

Furthermore, as observed in the summary statistics, each year we have around 100 stocks on average with valid data. Figure 2.5 shows the number of stocks of each portfolio for each ranking month. On average we have 30 stocks in each portfolio, hence reasonably diversified. Notice that

double-sorts would lead to very non-diversified portfolios, because of the low number of stocks that attend the criteria of having enough transactions both on the spot market and on stock lending market. In this sense Fama and MacBeth regressions are more elucidative to analyze the *additional* predictive power of the loan fee dispersion. The Fama and MacBeth results are presented in the next section.

2.3.2 Fama and MacBeth Regressions

We use Fama and MacBeth (1973) regression methodology as an alternative attempt to access the predictive power of loan fee dispersion on returns. The advantage of this methodology is that it allows us to examine the predictive power of loan fee dispersion while controlling for other short measures as well as for known predictors of cross-sectional stock returns. We conduct the Fama and Macbeth regressions in the usual way. Each month we run a cross-sectional regression of stock returns on last month: loan fee dispersion, average loan fee, short interest and days to cover and a set of control variables known to predict returns that includes log of market capitalization, log of book to market, reversal (defined as past month return) and momentum (defined as the cumulative holding-period return from month t-12 and t-2).

Table 2.12 shows the Fama-Macbeth results. Loan fee dispersion is the stronger predictor of negative future returns when compared to the other measures. When it is the only short related measure included in the regression, its coefficient is -0.606. Controlling for all other short related measures, its impact on returns increases to -1.187, highly significant. Notice that we include both the loan dispersion and loan fee average, the loan fee dispersion wipes out all the negative impact of average loan fee on returns. This corroborates to our hypothesis that the average loan fee neglects some important aspects of the market which are captured by the dispersion of fees.

Notice that both loan fee and days to cover seems to have a negative impact on returns, although the coefficient on the regressions are not significant, those results are different than Drechsler and Drechsler (2014) and Hong et al. (2015).

Table 2.4 shows that there is a high correlation between *Average Fee* and *Loan Fee Dispersion*, however, our results suggest *Loan Fee Dispersion* is a better predictor of future low stock returns.

2.4 Conclusion

We show that loan fee dispersion, which is the average standard deviation of the loan fee, is a well-motivated measure of short risk across stocks. Consistent with our hypothesis loan fee dispersion is the best predictor of future return in the cross-section when compared to traditional short related measures.

2.5 Tables

Tabela 2.1: **Summary Statistics** This table reports summary statistics for the main variables in our study. Short interest is the total borrowed shares over total shares outstanding. Each month, SD Fee is calculated as the monthly average of daily standard deviation of fees. Average Fee is the average value weighted loan fee. Short interest is the total of borrowed shares over total shares outstanding. Market Cap is represented in Millions. Book-to-market ratio is calculated in the end of December of each year. The cases with negative book value are deleted. Herfindahl Index (HI) is the sum of squares of the market shares of the brokers. Increases in the index indicate a decrease in competition and an increase of market power

	mean	min	median	max	std.dev	var
Return	0.007	-0.620	0.007	1.041	0.110	0.012
Volatility	0.024	0.001	0.021	0.123	0.012	0.0001
Price	20.437	0.432	16.580	340.200	17.057	290.940
Short Interest	2.232	0.001	1.390	39.294	2.586	6.685
Turnover	0.005	0.00000	0.003	0.118	0.005	0.00002
Avg Fee	3.863	0.261	2.598	47.356	4.014	16.112
SD Fee	0.927	0.045	0.662	12.192	0.882	0.777
SD Fee (MF)	0.648	0	0.364	11.065	0.871	0.758
HI Lenders	0.128	0.043	0.103	0.904	0.084	0.007
HI Borrowers	0.125	0.049	0.104	0.934	0.075	0.006

Tabela 2.2: **Panel Regression** Panel Regression with individual and time fixed effects. Log MC is the log of market cap. BM stands for Book to Market. Turnover is calculated by dividing the total number of shares traded in a year by the number of shares outstanding. Herfindahl Index (HI) is the sum of squares of each broker market share. Increases in the index indicate a decrease in competition. Degree is simply the number of different brokers lending (borrowing) that stock in the last month.

	Dependent variable:		
	sd_fee		
	(1)	(2)	(3)
Log MC	−0.079**		−0.042
	(0.034)		(0.034)
BM	25.255		23.081
	(41.885)		(39.714)
Volatility	13.497***		12.441***
	(1.513)		(1.445)
Turnover	31.422***		19.300***
	(3.236)		(3.159)
Short Interest		0.032***	0.030***
		(0.006)	(0.006)
HI Lenders		0.297	0.444**
		(0.211)	(0.222)
HI Borrowers		−0.229	−0.031
		(0.219)	(0.232)
Degree Lenders		0.052***	0.046***
		(0.002)	(0.002)
Degree Borrowers		−0.012***	−0.013***
		(0.003)	(0.003)
Observations	5,245	6,391	5,245
R^2	0.059	0.125	0.156
Adjusted R^2	0.057	0.120	0.149
F Statistic	79.089*** (df = 4; 5030)	175.510*** (df = 5; 6158)	103.208*** (df = 9; 5025)
Note:			*p<0.1; **p<0.05; ***p<0.01

Tabela 2.3: **Average Statistics per Year** This table reports Number of Observations and the average values of our main variables for each year. N Obs is the number of observations,i.e. datapoints. N Stocks is the number of stocks composing the dataset that year. BM stands for Book to Market. Short interest is the total borrowed shares over total shares outstanding. Each month, SD Fee is calculated as the monthly average of daily standard deviation of fees. All values presented as year average

Year	N Obs	N Stocks	Market Cap	BM	Short Interest	Avg fee	SD fee
2007	745	78	30405		2.62	5.49	1.18
2008	870	93	29570	0.47	2.24	4.11	0.92
2009	885	99	27907	0.63	1.51	3.59	0.91
2010	990	105	30543	0.66	1.57	3.43	0.87
2011	1235	129	26728	0.53	2.13	3.66	0.93
2012	1328	131	26479	0.64	2.40	3.50	0.81
2013	564	120	26301	0.75	3.84	3.82	0.98

Tabela 2.4: **Correlation Matrix** This table shows the correlation between the variables used to sort the long short portfolios. SD Fee is the monthly average of daily standard deviation of fees; Avg Fee is the monthly average of daily average fees. Short interest is the total borrowed shares over total shares outstanding. DTC represents Days to Cover, i.e. the ratio between short interest and turnover. Each data-point here represent a monthly measure for a specific stock.

	SD Fee	Avg Fee	Short Interest	DTC
SD Fee	1			
Avg Fee	0.757	1		
Short Interest	0.218	0.176	1	
DTC	0.015	0.065	0.477	1

Tabela 2.5: **Summary of Portfolio characteristics** For each month we calculate the monthly (market cap weighted) average of each variable. This table displays the time series average for each variable.

Portfolio	Return	Volatility	Short Interest	Turnover	W Avg Fee	MC	BM
sd_fee1	0.009	0.022	1.186	0.002	0.805	176.417	0.539
sd_fee2	0.016	0.023	2.020	0.004	1.703	131.043	0.598
sd_fee3	0.015	0.026	2.183	0.004	5.651	41.272	0.618

Tabela 2.6: **Returns to Long Short Strategies** Our main portfolio is based on loan fee dispersion, SdMLd, Small dispersion Minus Large dispersion. CMEw stands for Cheap Minus Expensive and relates to average loan fee. The remaining small minus large portfolios, SsiMLsi and SdtcMLdtc, are based on Short Interest and Days to Cover. Sharpe index based on annualized returns.

	PANEL A: Value Weighted Portfolios			
portfolio	Avg	Sd	Sharpe	T-stat
SdMLd	0.124	0.167	0.742	6.422
CMEw	0.071	0.191	0.371	3.211
SsiMLsi	0.073	0.149	0.493	4.267
SdtcMLdtc	0.042	0.174	0.242	2.099
	PANEL B: Equal Weighted Portfolios			
portfolio	Avg	Sd	Sharpe	T-stat
SdMLd	0.094	0.135	0.693	6.002
CMEw	0.054	0.142	0.378	3.276
SsiMLsi	0.090	0.147	0.611	5.295
SdtcMLdtc	0.007	0.124	0.057	0.497

Tabela 2.7: **Market Factor Returns** This table presents average returns, standard deviation of returns and sharpe ratio for the main market factors in Brazil. HML stands for High Minus Low ; SMB represents Small minus Big; Winners minus Losers is abbreviated as WML and finally IML relates to liquidity, Illiquid Minus Liquid. Those factors relate to Book to Market, Market Cap, Momentum and Liquidity, respectively

Factor	Avg	Sd	Sharpe	T-stat
Market	0.002	0.202	0.010	0.087
HML	0.031	0.123	0.253	2.216
SMB	0.003	0.173	0.017	0.153
WML	0.134	0.177	0.754	6.619
IML	0.025	0.122	0.202	1.772

Tabela 2.8: **Returns to portfolio strategies based on Loan Fee Dispersion** This table provides portfolio alphas and loading, sorted on Loan Fee Dispersion. At the end of each month, all the stocks are sorted into three quantile: small (bottom 30%), Middle (40%), large (top 30%) deciles based on their loan fee dispersion at the end of each month. Portfolio returns are computed over the next month minus the monthly risk free rate - 30 day DI Swap. For the analysis we consider 5 factors: Market, Small Minus Big Factor (SMB), High Minus Low (HML), Winners Minus Losers (WML), Illiquid Minus Liquid (IML). The factors considered were extracted from NEFIN. Data runs from January 2007 to June 2013. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	PANEL A: Value Weighted Portfolios					
Portfolios	alpha	Market	HML	SMB	WML	IML
sd fee1	0.004	1.112***				
	t = 1.510	t = 32.301				
sd fee2	−0.004	0.991***				
	t = −1.626	t = 30.908				
sd fee3	−0.006**	0.853***				
	t = −2.258	t = 10.612				
SdMLd	0.010**	0.259***				
	t = 2.498	t = 2.911				
sd fee1	0.004	1.133***	−0.053	−0.077		
	t = 1.608	t = 26.586	t = −0.808	t = −1.571		
sd fee2	−0.004*	1.012***	0.095*	−0.020		
	t = −1.807	t = 25.884	t = 1.730	t = −0.323		
sd fee3	−0.007**	0.865***	0.207	0.042		
	t = −2.525	t = 9.397	t = 1.482	t = 0.407		
SdMLd	0.011***	0.267***	−0.260	−0.119		
	t = 2.838	t = 2.706	t = −1.596	t = −0.933		
sd fee1	0.003	1.145***	−0.043	−0.049	0.051	
	t = 1.487	t = 26.441	t = −0.578	t = −0.893	t = 1.017	
sd fee2	−0.003	0.974***	0.065	−0.102	−0.152**	
	t = −1.257	t = 22.513	t = 1.095	t = −1.353	t = −2.501	
sd fee3	−0.004	0.812***	0.165	−0.072	−0.215**	
	t = −1.563	t = 10.332	t = 1.404	t = −0.611	t = −2.458	
SdMLd	0.008**	0.333***	−0.208	0.023	0.267**	
	t = 2.019	t = 4.215	t = −1.352	t = 0.153	t = 2.296	
sd fee1	0.003	1.165***	−0.051	−0.122	0.037	0.107
	t = 1.456	t = 24.533	t = −0.682	t = −1.278	t = 0.704	t = 1.117
sd fee2	−0.003	1.034***	0.043	−0.318***	−0.195***	0.321***
	t = −1.230	t = 22.862	t = 0.857	t = −3.164	t = −3.190	t = 3.144
sd fee3	−0.004*	0.691***	0.212***	0.367*	−0.127	−0.651***
	t = −1.719	t = 7.379	t = 2.933	t = 1.756	t = −1.472	t = −3.469
SdMLd	0.007**	0.475***	−0.262**	−0.489*	0.164	0.758***
	t = 2.056	t = 4.257	t = −2.307	t = −1.663	t = 1.364	t = 2.706

	PANEL B: Equal Weighted Portfolios					
Portfolios	alpha	Market	HML	SMB	WML	IML
sd fee1	0.004	0.905***				
	t = 1.626	t = 27.443				

sd fee2	0.002	0.895***				
	t = 0.687	t = 15.401				
sd fee3	−0.004	1.052***				
	t = −0.899	t = 8.551				
SdMLd	0.008*	−0.147				
	t = 1.911	t = −1.179				
sd fee1	0.004*	0.839***	−0.015	0.175***		
	t = 1.942	t = 28.492	t = −0.368	t = 3.549		
sd fee2	0.001	0.855***	0.162***	0.170*		
	t = 0.580	t = 19.192	t = 3.056	t = 1.921		
sd fee3	−0.005	0.913***	0.242	0.470***		
	t = −1.295	t = 8.490	t = 1.504	t = 3.557		
SdMLd	0.008**	−0.074	−0.258	−0.295***		
	t = 2.322	t = −0.642	t = −1.521	t = −2.787		
sd fee1	0.004*	0.845***	−0.011	0.189***	0.026	
	t = 1.872	t = 30.268	t = −0.255	t = 2.975	t = 0.528	
sd fee2	0.004	0.805***	0.124*	0.062	−0.202**	
	t = 1.560	t = 18.537	t = 1.917	t = 0.801	t = −2.401	
sd fee3	−0.0002	0.816***	0.167	0.260**	−0.394***	
	t = −0.070	t = 8.855	t = 1.376	t = 2.245	t = −2.891	
SdMLd	0.004	0.030	−0.177	−0.071	0.420***	
	t = 1.144	t = 0.300	t = −1.434	t = −0.845	t = 3.813	
sd fee1	0.004*	0.835***	−0.006	0.228***	0.033	−0.058
	t = 1.942	t = 26.043	t = −0.146	t = 2.862	t = 0.723	t = −0.642
sd fee2	0.004	0.798***	0.126**	0.087	−0.197**	−0.037
	t = 1.551	t = 15.922	t = 1.978	t = 0.812	t = −2.325	t = −0.292
sd fee3	0.0001	0.722***	0.202**	0.600***	−0.326**	−0.504**
	t = 0.017	t = 7.676	t = 2.176	t = 3.196	t = −2.355	t = −2.491
SdMLd	0.004	0.113	−0.209**	−0.373**	0.360***	0.446***
	t = 1.152	t = 1.235	t = −2.142	t = −2.309	t = 3.146	t = 2.629

Tabela 2.9: **Returns to portfolio strategies based on Average Loan Fee** This table provides portfolio alphas and loading, sorted on Average Loan Fee. At the end of each month, all the stocks are sorted into three quantile: small (bottom 30%), Middle (40%), large (top 30%) deciles based on their average loan fee at the end of each month. Portfolio returns are computed over the next month minus the monthly risk free rate - 30 day DI Swap. For the analysis we consider 5 factors: Market, Small Minus Big Factor (SMB), High Minus Low (HML), Winners Minus Losers (WML), Illiquid Minus Liquid (IML). The factors considered were extracted from NEFIN. Data runs from January 2007 to June 2013. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	PANEL A: Value Weight Portfolios					
Portfolios	alpha	Market	HML	SMB	WML	IML
wavg fee1	0.001	1.123***				
	t = 0.207	t = 30.392				
wavg fee2	−0.001	0.756***				
	t = −0.283	t = 17.539				
wavg fee3	−0.005	0.827***				
	t = −1.322	t = 8.358				
CMEw	0.006	0.297**				
	t = 0.913	t = 2.373				
wavg fee1	0.001	1.146***	−0.021	−0.071		
	t = 0.225	t = 30.632	t = −0.279	t = −1.288		
wavg fee2	−0.001	0.771***	0.156*	0.016		
	t = −0.476	t = 20.652	t = 1.868	t = 0.184		
wavg fee3	−0.006	0.822***	0.208**	0.091		
	t = −1.436	t = 8.319	t = 2.077	t = 0.988		
CMEw	0.006	0.325***	−0.229	−0.162		
	t = 1.011	t = 2.676	t = −1.620	t = −1.195		
wavg fee1	0.001	1.144***	−0.023	−0.076	−0.011	
	t = 0.265	t = 27.448	t = −0.279	t = −1.561	t = −0.137	
wavg fee2	−0.0003	0.754***	0.142	−0.022	−0.072	
	t = −0.118	t = 18.018	t = 1.557	t = −0.224	t = −0.714	
wavg fee3	−0.003	0.762***	0.162*	−0.037	−0.241**	
	t = −0.851	t = 8.638	t = 1.719	t = −0.424	t = −2.220	
CMEw	0.004	0.382***	−0.185	−0.039	0.230	
	t = 0.648	t = 3.323	t = −1.162	t = −0.301	t = 1.334	
wavg fee1	0.001	1.164***	−0.031	−0.151*	−0.026	0.110
	t = 0.231	t = 26.614	t = −0.367	t = −1.698	t = −0.311	t = 0.935
wavg fee2	−0.0002	0.739***	0.148*	0.031	−0.061	−0.078
	t = −0.094	t = 13.909	t = 1.732	t = 0.276	t = −0.640	t = −0.510
wavg fee3	−0.003	0.714***	0.180**	0.136	−0.206*	−0.257
	t = −0.794	t = 6.552	t = 2.059	t = 0.684	t = −1.803	t = −1.100
CMEw	0.004	0.450***	−0.211	−0.287	0.180	0.367
	t = 0.595	t = 3.162	t = −1.352	t = −1.024	t = 0.994	t = 1.064

	PANEL B: Equal Weight Portfolios					
Portfolios	alpha	Market	HML	SMB	WML	IML
wavg fee1	0.003**	1.016***				
	t = 2.419	t = 19.541				

wavg fee2	−0.0002	0.814***				
	t = −0.064	t = 17.958				
wavg fee3	−0.001	1.050***				
	t = −0.192	t = 7.626				
CMEw	0.005	−0.033				
	t = 0.883	t = −0.239				
wavg fee1	0.003**	0.973***	0.122***	0.163***		
	t = 2.499	t = 28.163	t = 3.427	t = 3.739		
wavg fee2	−0.0004	0.769***	0.141**	0.175**		
	t = −0.186	t = 17.511	t = 2.133	t = 2.043		
wavg fee3	−0.001	0.892***	0.128	0.479***		
	t = −0.252	t = 8.502	t = 1.010	t = 3.073		
CMEw	0.004	0.082	−0.006	−0.316**		
	t = 1.030	t = 0.722	t = −0.047	t = −2.006		
wavg fee1	0.004***	0.949***	0.103***	0.109*	−0.100**	
	t = 3.158	t = 26.548	t = 3.073	t = 1.828	t = −2.192	
wavg fee2	0.001	0.736***	0.116*	0.105	−0.131*	
	t = 0.404	t = 17.969	t = 1.794	t = 1.145	t = −1.650	
wavg fee3	0.003	0.801***	0.058	0.285**	−0.364**	
	t = 0.709	t = 9.437	t = 0.527	t = 2.479	t = −2.322	
CMEw	0.001	0.147	0.045	−0.175	0.264*	
	t = 0.312	t = 1.393	t = 0.394	t = −1.376	t = 1.707	
wavg fee1	0.004***	0.942***	0.105***	0.133*	−0.095**	−0.035
	t = 3.193	t = 22.985	t = 3.040	t = 1.712	t = −2.081	t = −0.519
wavg fee2	0.001	0.706***	0.127**	0.214*	−0.110	−0.162
	t = 0.421	t = 13.814	t = 2.024	t = 1.787	t = −1.495	t = −1.012
wavg fee3	0.003	0.732***	0.084	0.536***	−0.314*	−0.373*
	t = 0.735	t = 7.703	t = 0.874	t = 3.349	t = −1.935	t = −1.741
CMEw	0.001	0.210*	0.021	−0.404***	0.219	0.338
	t = 0.259	t = 1.843	t = 0.207	t = −2.674	t = 1.361	t = 1.622

Tabela 2.10: **Returns to portfolio strategies based on Short Interest** This table provides portfolio alphas and loading, sorted on Short Interest. At the end of each month, all the stocks are sorted into three quantile: small (bottom 30%), Middle (40%), large (top 30%) deciles based on their short interest at the end of each month. Portfolio returns are computed over the next month minus the monthly risk free rate - 30 day DI Swap. For the analysis we consider 5 factors: Market, Small Minus Big Factor (SMB), High Minus Low (HML), Winners Minus Losers (WML), Illiquid Minus Liquid (IML). The factors considered were extracted from NEFIN. Data runs from January 2007 to June 2013. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	PANEL A: Value Weighted Portfolios					
Portfolios	alpha	Market	HML	SMB	WML	IML
si1	0.001	1.097***				
	t = 0.492	t = 30.687				
si2	−0.001	0.956***				
	t = −0.230	t = 29.394				
si3	−0.005*	1.050***				
	t = −1.731	t = 13.579				
SsiMLsi	0.006	0.047				
	t = 1.508	t = 0.450				
si1	0.001	1.105***	0.055	−0.002		
	t = 0.450	t = 28.679	t = 0.674	t = −0.040		
si2	−0.001	0.997***	−0.012	−0.119**		
	t = −0.284	t = 26.182	t = −0.195	t = −2.236		
si3	−0.005*	1.059***	0.048	−0.007		
	t = −1.699	t = 12.777	t = 0.437	t = −0.081		
SsiMLsi	0.006	0.046	0.007	0.006		
	t = 1.477	t = 0.435	t = 0.044	t = 0.056		
si1	0.001	1.096***	0.048	−0.021	−0.036	
	t = 0.559	t = 30.149	t = 0.518	t = −0.334	t = −0.454	
si2	−0.0004	0.992***	−0.017	−0.131**	−0.023	
	t = −0.169	t = 21.794	t = −0.255	t = −2.373	t = −0.390	
si3	−0.003	0.999***	0.002	−0.137*	−0.242***	
	t = −0.702	t = 13.425	t = 0.015	t = −1.655	t = −3.402	
SsiMLsi	0.004	0.097	0.046	0.116	0.207	
	t = 0.743	t = 1.074	t = 0.281	t = 1.001	t = 1.575	
si1	0.001	1.134***	0.033	−0.160*	−0.064	0.206*
	t = 0.463	t = 29.465	t = 0.364	t = −1.736	t = −0.734	t = 1.953
si2	−0.001	1.012***	−0.024	−0.205**	−0.038	0.110
	t = −0.191	t = 19.394	t = −0.371	t = −2.573	t = −0.647	t = 1.470
si3	−0.002	0.925***	0.030	0.129	−0.189**	−0.394**
	t = −0.656	t = 10.678	t = 0.299	t = 0.836	t = −2.552	t = −2.254
SsiMLsi	0.003	0.209*	0.004	−0.289	0.126	0.600**
	t = 0.658	t = 1.816	t = 0.024	t = −1.289	t = 0.867	t = 2.264

	PANEL B: Equal Weighted Portfolios					
Portfolios	alpha	Market	HML	SMB	WML	IML
si1	0.004	0.856***				
	t = 1.467	t = 14.983				

si2	0.0001	0.892***				
	t = 0.033	t = 17.024				
si3	−0.003	1.105***				
	t = −0.764	t = 10.635				
SsiMLsi	0.008**	−0.249***				
	t = 2.024	t = −2.649				
si1	0.004	0.783***	0.095	0.237**		
	t = 1.637	t = 20.615	t = 1.061	t = 2.503		
si2	−0.00004	0.835***	0.091	0.189***		
	t = −0.021	t = 19.890	t = 1.132	t = 2.594		
si3	−0.004	0.995***	0.191	0.372***		
	t = −1.083	t = 10.697	t = 1.484	t = 2.805		
SsiMLsi	0.008**	−0.213**	−0.096	−0.135		
	t = 2.174	t = −2.447	t = −0.546	t = −1.060		
si1	0.005**	0.762***	0.079	0.192*	−0.084	
	t = 1.977	t = 17.993	t = 0.834	t = 1.917	t = −0.947	
si2	0.001	0.800***	0.064	0.114	−0.140*	
	t = 0.790	t = 21.891	t = 0.845	t = 1.532	t = −1.646	
si3	0.0004	0.905***	0.121	0.179	−0.363***	
	t = 0.107	t = 11.779	t = 1.216	t = 1.324	t = −3.400	
SsiMLsi	0.005	−0.143*	−0.042	0.013	0.279**	
	t = 1.488	t = −1.688	t = −0.267	t = 0.086	t = 2.339	
si1	0.005**	0.779***	0.072	0.129	−0.097	0.093
	t = 2.017	t = 17.680	t = 0.756	t = 0.861	t = −1.105	t = 0.578
si2	0.002	0.767***	0.077	0.236**	−0.116	−0.180
	t = 0.812	t = 15.531	t = 1.119	t = 2.308	t = −1.396	t = −1.358
si3	0.001	0.818***	0.155*	0.496***	−0.299***	−0.470***
	t = 0.182	t = 10.937	t = 1.830	t = 2.644	t = −2.808	t = −3.016
SsiMLsi	0.004	−0.038	−0.082	−0.367	0.203*	0.563***
	t = 1.321	t = −0.628	t = −0.594	t = −1.637	t = 1.800	t = 3.196

Tabela 2.11: **Returns to portfolio strategies based on Days to Cover** This table provides portfolio alphas and loading, sorted on Days to Cover. At the end of each month, all the stocks are sorted into three quantile: small (bottom 30%), Middle (40%), large (top 30%) deciles based on their days to cover at the end of each month. Portfolio returns are computed over the next month minus the monthly risk free rate - 30 day DI Swap. For the analysis we consider 5 factors: Market, Small Minus Big Factor (SMB), High Minus Low (HML), Winners Minus Losers (WML), Illiquid Minus Liquid (IML). The factors considered were extracted from NEFIN. Data runs from January 2007 to June 2013. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	PANEL A: Value Weighted Portfolios					
Portfolios	alpha	Market	HML	SMB	WML	IML
dtcover1	0.002	1.139***				
	t = 0.550	t = 15.171				
dtcover2	−0.002	1.036***				
	t = −0.708	t = 21.427				
dtcover3	−0.001	0.899***				
	t = −0.428	t = 15.518				
SdtcMLdtc	0.003	0.240**				
	t = 0.570	t = 2.150				
dtcover1	0.002	1.143***	−0.084	−0.043		
	t = 0.590	t = 15.331	t = −0.777	t = −0.538		
dtcover2	−0.002	1.061***	0.102	−0.031		
	t = −0.883	t = 25.975	t = 1.285	t = −0.493		
dtcover3	−0.002	0.951***	0.095	−0.108*		
	t = −0.520	t = 13.627	t = 1.381	t = −1.804		
SdtcMLdtc	0.004	0.192	−0.179	0.065		
	t = 0.630	t = 1.599	t = −1.168	t = 0.568		
dtcover1	0.003	1.124***	−0.099	−0.084	−0.077	
	t = 0.788	t = 13.522	t = −0.882	t = −0.952	t = −0.745	
dtcover2	−0.0002	1.011***	0.063	−0.138***	−0.201***	
	t = −0.057	t = 21.272	t = 0.956	t = −2.850	t = −3.077	
dtcover3	−0.003	0.982***	0.119	−0.041	0.126	
	t = −0.853	t = 11.384	t = 1.639	t = −0.653	t = 1.336	
SdtcMLdtc	0.006	0.142	−0.218	−0.043	−0.203	
	t = 0.903	t = 0.990	t = −1.349	t = −0.347	t = −1.122	
dtcover1	0.003	1.161***	−0.113	−0.220	−0.104	0.201
	t = 0.703	t = 14.074	t = −0.968	t = −1.304	t = −0.949	t = 1.198
dtcover2	−0.0001	1.000***	0.068	−0.098	−0.193***	−0.060
	t = −0.045	t = 18.666	t = 0.967	t = −0.784	t = −2.938	t = −0.412
dtcover3	−0.003	0.985***	0.118	−0.050	0.124	0.013
	t = −0.845	t = 10.107	t = 1.626	t = −0.464	t = 1.302	t = 0.109
SdtcMLdtc	0.006	0.176	−0.231	−0.170	−0.228	0.187
	t = 0.857	t = 1.165	t = −1.392	t = −0.729	t = −1.215	t = 0.741

	PANEL B: Equal Weighted Portfolios					
Portfolios	alpha	Market	HML	SMB	WML	IML
dtcover1	0.003	1.087***				
	t = 0.664	t = 12.458				

dtcover2	−0.002	0.961***				
	t = −0.882	t = 16.906				
dtcover3	0.002	0.788***				
	t = 0.733	t = 13.427				
SdtcMLdtc	0.0004	0.298***				
	t = 0.093	t = 4.707				
dtcover1	0.003	0.933***	0.113	0.463***		
	t = 0.955	t = 17.695	t = 1.245	t = 4.080		
dtcover2	−0.003	0.895***	0.103	0.219***		
	t = −1.214	t = 19.494	t = 1.398	t = 2.990		
dtcover3	0.002	0.774***	0.171***	0.103		
	t = 0.572	t = 13.948	t = 2.680	t = 1.131		
SdtcMLdtc	0.001	0.159***	−0.058	0.360***		
	t = 0.221	t = 3.814	t = −0.668	t = 5.668		
dtcover1	0.005*	0.875***	0.069	0.340***	−0.232**	
	t = 1.784	t = 18.281	t = 0.724	t = 4.592	t = −2.050	
dtcover2	−0.0001	0.841***	0.062	0.103	−0.218***	
	t = −0.074	t = 21.023	t = 1.106	t = 1.260	t = −2.952	
dtcover3	0.003	0.745***	0.149**	0.040	−0.118	
	t = 0.975	t = 14.096	t = 2.401	t = 0.436	t = −1.241	
SdtcMLdtc	0.002	0.131***	−0.080	0.299***	−0.114	
	t = 0.581	t = 2.581	t = −0.831	t = 4.631	t = −1.320	
dtcover1	0.005*	0.841***	0.082	0.465***	−0.207*	−0.185
	t = 1.831	t = 16.412	t = 0.814	t = 3.638	t = −1.786	t = −1.092
dtcover2	0.00002	0.791***	0.081**	0.283***	−0.182***	−0.267**
	t = 0.010	t = 17.771	t = 2.089	t = 2.637	t = −2.603	t = −2.537
dtcover3	0.003	0.732***	0.153***	0.086	−0.109	−0.068
	t = 0.958	t = 10.707	t = 2.625	t = 0.691	t = −1.158	t = −0.456
SdtcMLdtc	0.002	0.109**	−0.071	0.378***	−0.098	−0.117
	t = 0.615	t = 2.187	t = −0.685	t = 3.189	t = −1.048	t = −0.845

Tabela 2.12: **: Fama-MacBeth Regressions** This table reports results from Fama-Macbeth (1973) regression of monthly stock returns on Fee Dispersion, Average Fee, Short-Interest and Days to Cover. Fee Dispersion is the average daily standard deviation. Average Fee is the average value weighted loan fee. Short interest is the total of borrowed shares over total shares outstanding. Days-to-cover is short interest ratio over daily turnover. Size (log(MC)) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (log(BM)) is the natural log of book-to-market ratio calculated as end of December of each year. The cases with negative book value are deleted. The short term reversal measure (Reversal) is the lagged monthly return. All the t-statistics are Newey and West (1987) adjusted to control for heteroskedasticity and autocorrelation. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	Monthly Return (%)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fee Dispersion	−0.606**				−0.995**	−0.518**	−0.577**	−0.820*	−1.187**
	(0.238)				(0.419)	(0.218)	(0.242)	(0.422)	(0.581)
Average Fee		−0.045			0.123			0.105	0.190
		(0.056)			(0.109)			(0.113)	(0.159)
Short Interest			−0.142			−0.080		−0.086	0.092
			(0.110)			(0.099)		(0.096)	(0.248)
Days to cover				−0.001			−0.001		−0.001
				(0.001)			(0.001)		(0.001)
Log(MC)	−0.250	−0.239	−0.248	−0.190	−0.161	−0.267	−0.244	−0.192	−0.142
	(0.185)	(0.188)	(0.189)	(0.194)	(0.201)	(0.182)	(0.180)	(0.203)	(0.211)
Log(BM)	−0.148	−0.130	−0.101	−0.116	−0.151	−0.136	−0.117	−0.130	−0.096
	(0.233)	(0.233)	(0.250)	(0.242)	(0.248)	(0.238)	(0.229)	(0.250)	(0.238)
Reversal	−0.021	−0.023	−0.024	−0.026	−0.020	−0.019	−0.024	−0.018	−0.027
	(0.030)	(0.030)	(0.028)	(0.030)	(0.031)	(0.029)	(0.030)	(0.030)	(0.029)
Constant	5.965	5.498	6.148	4.584	3.837	6.586	6.150	4.843	4.062
	(4.389)	(4.588)	(4.489)	(4.909)	(4.763)	(4.237)	(4.430)	(4.757)	(4.776)
Observations	5,193	5,193	5,193	5,193	5,193	5,193	5,193	5,193	5,193
R^2	0.349	0.342	0.350	0.341	0.359	0.363	0.359	0.373	0.393

2.6 Figures

Figura 2.1: **Petrobras Loan Fee** Histogram of fees paid by investors renting PETR4 (Petrobras) on 30th April 2008 and April first 2009. The y-axis represents the number of contracts, x-axis shows the fee.

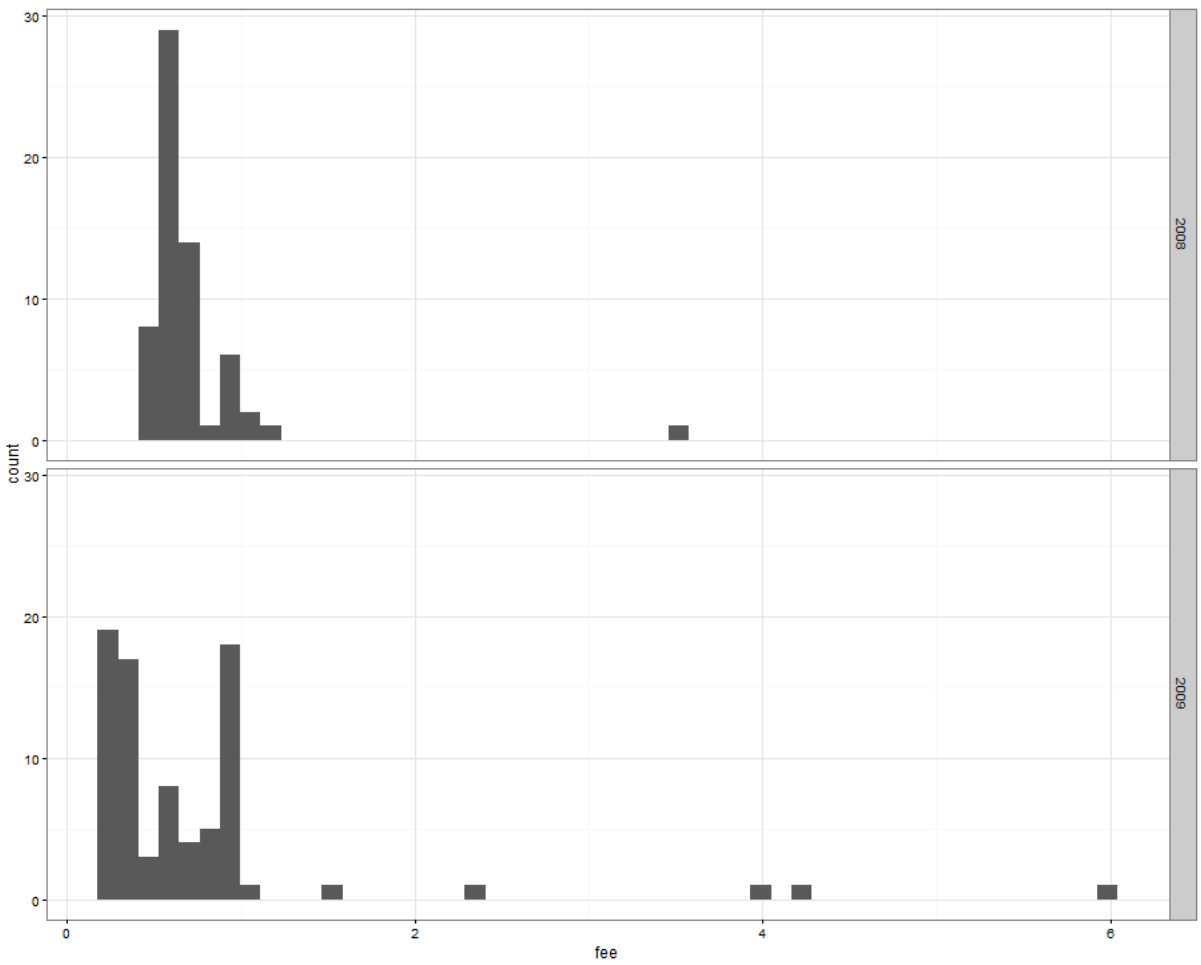


Figura 2.2: **Mutual Funds Fee Relative Spread: Lender.** Evolution of relative spread of loan fees received by individual person and foreign investors when compared to mutual funds

Figura 2.3: **Mutual Funds Fee Relative Spread: Borrower.** Evolution of relative spread of loan fees paid by individual person and foreign investors when compared to mutual funds

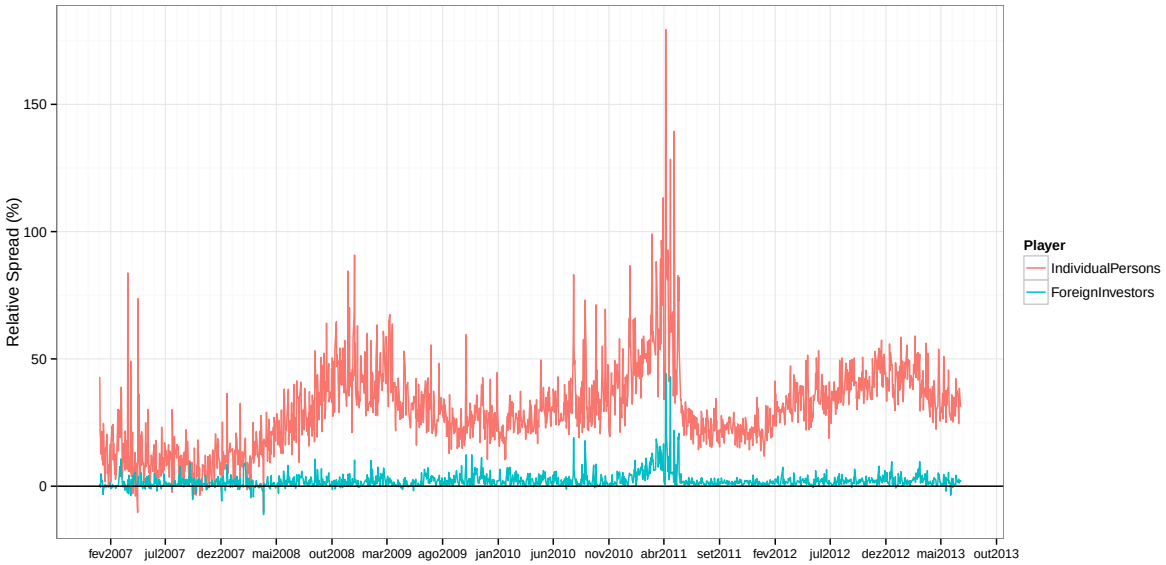


Figura 2.4: **Loan fee Measures Over Time.** Time series of aggregate measures. Each company daily average loan fee is calculated as the mean fee for contracts opened that day. The equal weighted cross sectional average composes our aggregated measure. The other aggregated measures are analogous.

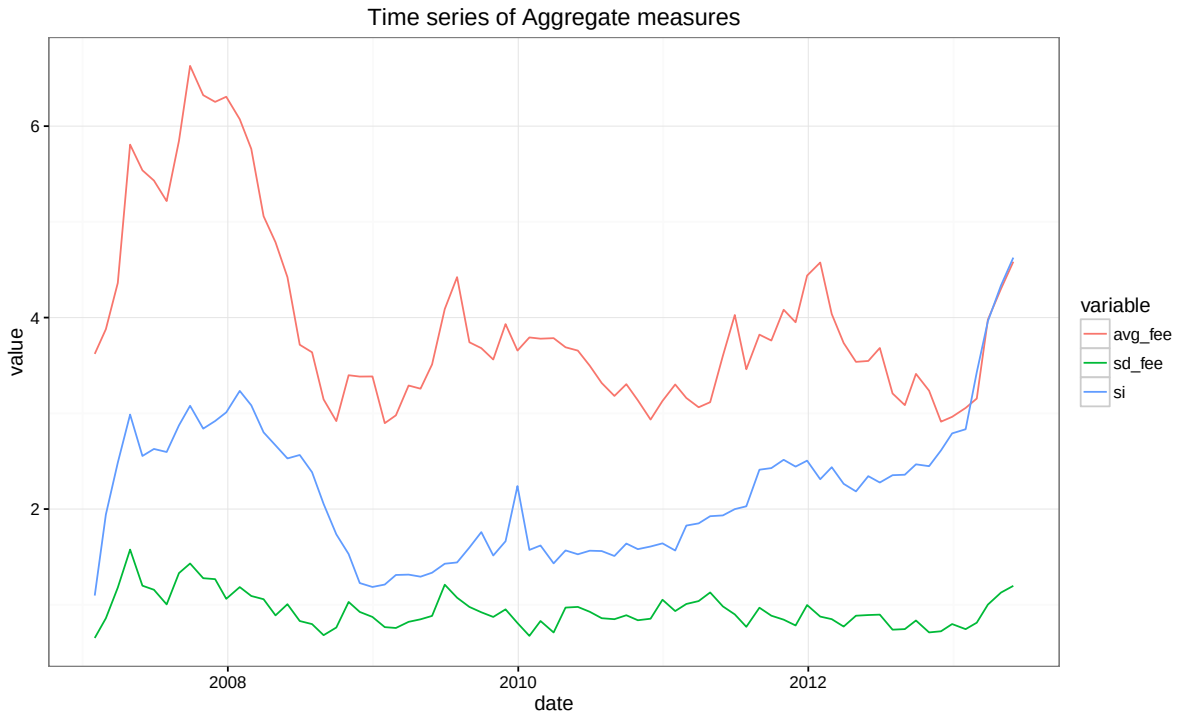


Figura 2.5: **Number of stocks for each portfolio.**

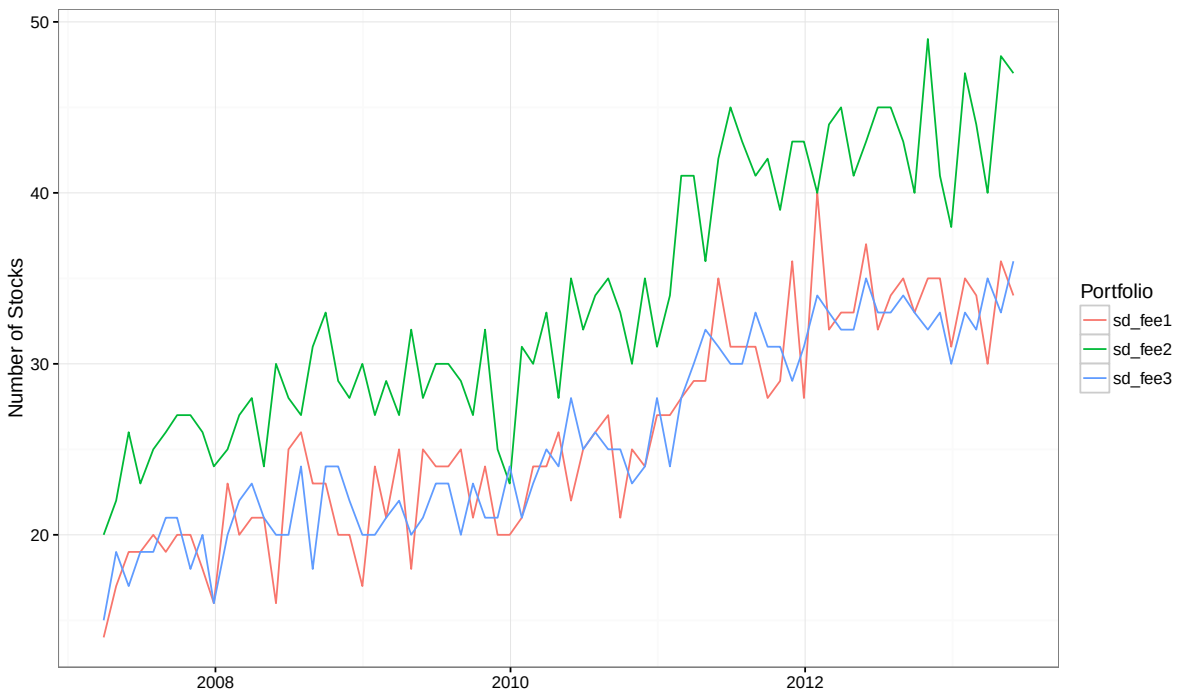


Figura 2.6: **Cumulative Returns for each long-short portfolio.**

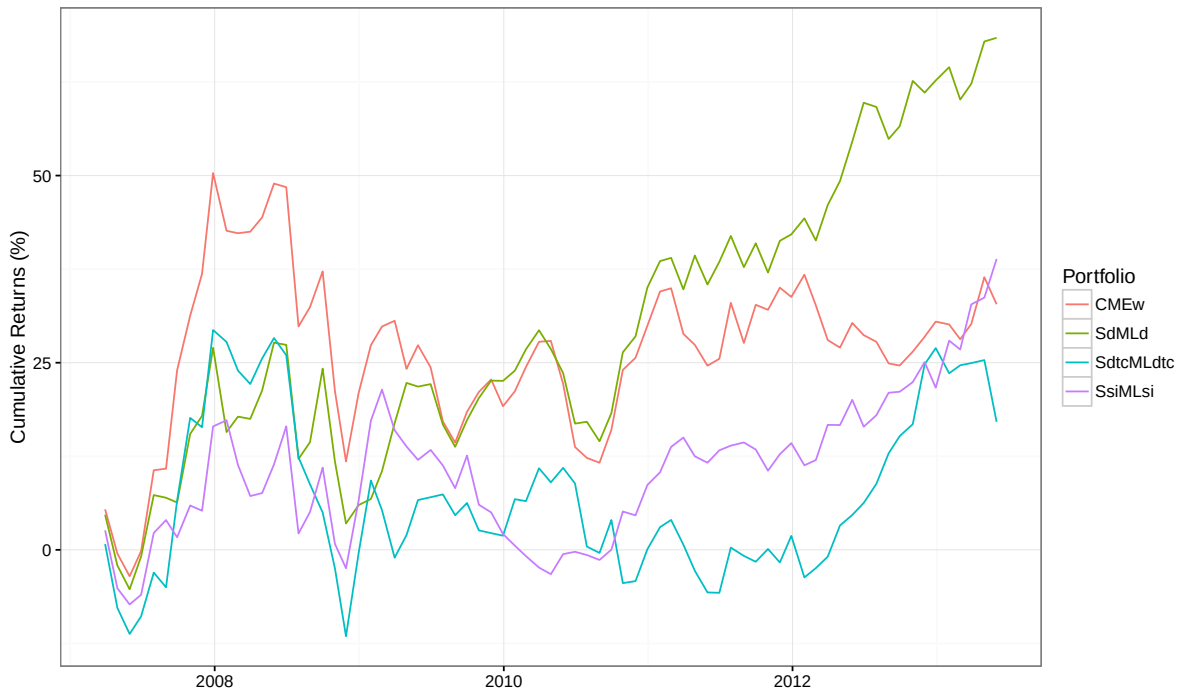
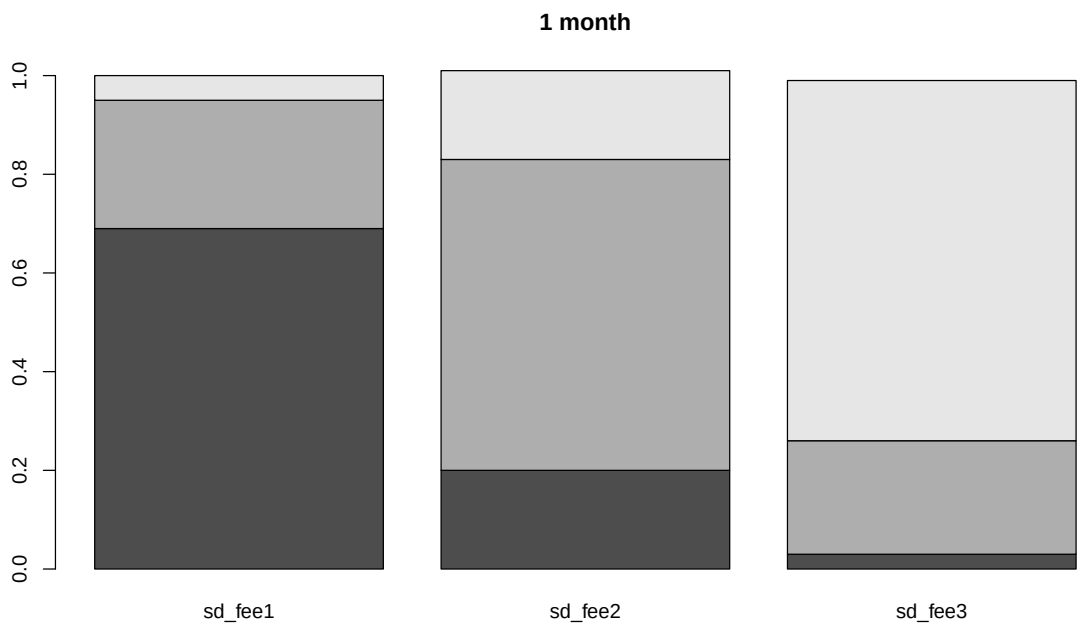


Figura 2.7: **Transition Matrix** The figure plots the transition probabilities of stocks across the three sd-fee sorted buckets over the period of one month. For example, the first column shows the percentage of stocks in the first bucket that end up in each of the three buckets after one month.



Capítulo 3

The Cross-Section of Risk and Return [1]

ABSTRACT

In the finance literature, a common practice is to create *factor-portfolios* by sorting on characteristics (such as book-to-market, past return, or profitability) associated with average returns. The goal of this exercise is to create a parsimonious set of factor-portfolios that explain the cross-section of average returns, in the sense that the returns of these factor-portfolios span the mean-variance efficient portfolio. We argue that this is unlikely to be the case, as factor portfolios constructed in this way fail to incorporate information about the covariance structure of returns. By using a high statistical power methodology to forecast future covariances, we are able to construct a set of portfolios which captures the characteristic premia, but hedges out much of the factor risk. We apply our methodology to hedge out unpriced risk in the Fama and French (2015) five-factors. We find that the squared Sharpe ratio of the optimal combination of the resulting hedged-factor portfolios is 2.15, compared with 1.37 for the unhedged portfolios, and is highly statistically significant.

3.1 Introduction

A common practice in the academic finance literature has been to create *factor-portfolios* by sorting on characteristics positively associated with expected returns. The resulting set of zero-investment factor-portfolios, which go long a portfolio of high-characteristic firms and short a portfolio of low-characteristic firms, then serve as a model for returns in that asset space. Prominent examples of this are the three- and five-factor models of Fama and French (1993, 2015), but there are numerous others, developed both to explain the equity market anomalies, and also the cross-section of returns in other asset classes.[2]

Consistent with this, Fama and French (2015, FF) argue that a standard dividend-discount

[1] Joint work with Kent Daniel, Simon Rottke, Tano Santos.

[2] Examples are: UMD (Carhart, 1997); LIQ (Pastor and Stambaugh, 2003); BAB (Frazzini and Pedersen, 2014); QMJ (Asness, Frazzini, and Pedersen, 2013); and RX and HML-FX (Lustig, Roussanov, and Verdelhan, 2011). We concentrate on the factors of Fama and French (2015). However, the critique we develop in Section 3.2 applies to any factors constructed using this method.

model implies that a combination of individual-firm metrics based on valuation, profitability and investment should forecast these firms' average returns. Based on this they develop a five factor model—consisting of the Mkt-Rf, SMB, HML, RMW, and CMA factor-portfolios—and argue that this model does a fairly good job explaining the cross-section of average returns for a variety of test portfolios, based on a set of time-series regressions like:

$$\begin{aligned} R_{p,t}-R_{f,t} &= \alpha_p + \beta_m \cdot (R_{m,t}-R_{f,t}) + \beta_{HML} \cdot HML_t + \beta_{SMB} \cdot SMB_t \\ &\quad +\beta_{CMA} \cdot CMA_t + \beta_{RMW} \cdot RMW_t + \epsilon_{p,t} \end{aligned}$$

where a set of portfolios is chosen for which the excess returns, $R_{p,t}-R_{f,t}$, exhibit a considerable average spread.[3]

Standard projection theory shows that the αs from such regressions will all be zero for all assets if and only if the mean-variance efficient (MVE) portfolio is in the span of the factor portfolios, or equivalently if the maximum Sharpe ratio in the economy is the maximum Sharpe-ratio achievable with the factor portfolios alone. For the case of the five factor-portfolios examined by Fama and French (2015), the *ex-post* optimal combination of these five-factors has an annualized Sharpe ratio of 1.17 over 1963:07-2014:12 time period. Despite several critiques of this methodology it remains popular in the finance literature.[4]

The objective of this paper is to refine our understanding of the relationship between firm characteristics and the risk and average returns of individual firms. Our argument is that, if characteristics are a good proxy for expected returns, then forming factor portfolios by sorting on characteristics will generally *not* explain the cross-section of returns in the way proposed in the papers in this literature.

The argument is straightforward, and based on the early insights of Markowitz (1952) and Roll (1977): suppose a set of characteristics are positively associated with average returns, and a corresponding set of long-short factor-portfolios are constructed by buying high-characteristic stocks and shorting low-characteristic stocks. This set of portfolios will explain the returns of portfolios sorted on the same characteristics, but are unlikely to span the MVE portfolio of all assets, because they do not take into account the asset covariance structure. The intuition underlying this comes from a stylized example: assume there is a single characteristic which is a perfect proxy for expected returns, i.e., $\mathbf{c} = \kappa\boldsymbol{\mu}$, where $\mathbf{c}$ is the characteristic vector, $\boldsymbol{\mu}$ is a vector of expected returns and κ is a constant of proportionality. A portfolio formed with weights proportional to firm-characteristics, i.e., with $\mathbf{w}^c \propto \mathbf{c} = \kappa\boldsymbol{\mu}$, will be MVE only if $\mathbf{w}^c \propto \mathbf{w}^* = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. In Section 3.2, we develop stylized model where we develop this argument formally.

When will $\mathbf{w}^c$ be proportional to $\mathbf{w}^*$? That is, when will the characteristic sorted portfolio be MVE? As we show in Section 3.2, this will be the case only in a few selected settings. For example, it will always be true in a single factor world framework in which the law of one price holds. However, it will not generally hold in settings where the number of factors exceeds the number of characteristics. Specifically, we show that any cross-sectional correlation between firm-characteristics and firm exposures to unpriced factors will result in the factor-portfolio being inefficient.

Of course our theoretical argument does not address the *magnitude* of the inefficiency of the characteristic-based factor portfolios. Intuitively, our theoretical argument is that forming factor-portfolios on the basis of characteristics alone leads to these portfolios being exposed to unpriced factor risk, risk which is hedged-out in the MVE portfolio. In Sections 3.3 and 3.4 we

[3] The Fama and French (2015) test portfolios SMB, HML, RMW, and CMA are formed by sorting on various combinations of firm size, valuation ratios, profitability and investment respectively.

[4] Daniel and Titman (1997) critique the original Fama and French (1993) technique. Our critique here is closely related to that paper. Also related to our discussion here are Lewellen, Nagel, and Shanken (2010) and Daniel and Titman (2012) who argue that the space of test assets used in numerous recent asset pricing tests is too low-dimensional to provide adequate statistical-power against reasonable alternative hypotheses. Our focus in this paper is also expanding the dimensionality of the asset return space, but we do so with a different set of techniques.

address respectively, the questions of how large the loadings on unpriced factors are likely to be, and how much improvement in the efficiency of the factor-portfolios can be obtained by hedging out the unpriced factor risk.

As we discuss in Section 3.3, extant evidence on the value effect suggests that the industry component of many characteristic measures, such as book-to-price, are not helpful in forecasting average returns. This suggests that any exposure of HML to industry factors is unpriced. Therefore, if this exposure were hedged out, it would result in a factor-portfolio with lower risk, but the same expected return, i.e., with a higher Sharpe ratio. Our analysis in Section 3.3 shows that the HML exposure on industry factors varies dramatically over time, but that at selected times the exposure can be very high. We highlight two episodes in particular in which the correlation between HML and industry factors exceeds 95%: in late-2000/early-2001 as the prices of high-technology firms earned large negative returns and became highly volatile, and 2008-2009 during the financial-crisis, a parallel episode for financial firms. In both of these episodes the past return performance of the industry led to the vast majority of the firms in the industry becoming either growth or value firms—that is, there was a high cross-sectional correlation between valuation ratios and industry membership—leading to HML becoming highly correlated with that industry factor.

However the evidence that the FF factor-portfolios sometimes load heavily on presumably unpriced industry factors, while suggestive, does not establish that these portfolios are inefficient. Therefore in Section 3.4, we address the question of what fraction of the risk of the FF factor-portfolios is unpriced and can therefore be hedged out, and how much improvement in Sharpe-ratio results from doing so. The method that we use for constructing our hedge portfolio builds on that developed in Daniel and Titman (1997). However, through the use of higher frequency data, industry adjustment, differential windows for calculating volatilities and correlations, and other improvements we are able to construct hedge portfolios that have both a higher spread in factor loadings and lower idiosyncratic risk. That is, they are more efficient hedge portfolios. Using this technique, we construct hedge portfolios for the five factor portfolios of Fama and French (2015). We are conservative in the way that we construct these portfolios; consistent with the methodology employed by Fama and French, we form these portfolios once per-year, in July, and hold the composition of the portfolios fixed for 12 months. The portfolios are value-weighted buy-and-hold portfolios. Except for the size (SMB) hedge portfolio, these all earn economically and statistically significant five-factor alphas. Using the combined Market-, HML-, RMW- and CMA-hedge portfolios, we construct a combination portfolio that has zero exposure to any of the five FF factors, and yet earns an annualized Sharpe-ratio of 0.883, close to that of the 1.17 Sharpe-ratio of the *ex-post* optimal combination of the five FF factor-portfolios. Thus, by hedging the unpriced factor risk in the FF portfolios, we increase the squared-Sharpe ratio of this optimal combination from from 1.37 to 2.15.

This result is important for several reasons. First it increases the hurdle for standard asset pricing models, in that pricing kernel variance that is required to explain the returns of our hedge factor portfolios is about double what is required to explain the returns of the Fama and French (2015) five factor-portfolios.

Second, while the characteristics approach to measure managed portfolio performance (see, e.g., Daniel, Grinblatt, Titman, and Wermers, 1997) has gained some popularity, the regression based approach initially employed by Jensen (1968) (and later by Fama and French (2010) and numerous others) remains the more popular. A good reason for this is that the characteristics approach can only be used to estimate the alpha of a portfolio when the holdings of the managed portfolio are known, and frequently sampled. In contrast, the Jensen-style regression approach can be used in the absence of holdings data, as long as a time series of portfolio returns are available.

However, as pointed out originally by Roll (1977), to use the regression approach, the multi-factor benchmark used in the regression test must be efficient, or the conclusions of the regression

test will be invalid. What we show in this paper is that, with the historical return data, efficiency of the proposed factor-portfolios can be rejected. However, the hedged versions of the factor-portfolio, that we construct here and which incorporate the information both from the characteristics and from the historical covariance structure, are efficient with respect to both of these information sources. Thus, alphas equivalent to what would be obtained with the DGTW characteristics-approach can be generated with the regression approach, if the hedged factor portfolios are used, without the need for portfolio holdings data.

The layout of the remainder of the paper is as follows: In Section 3.2 we lay out the underlying econometric theory that motivates our analysis. Section 3.3 provides a descriptive analysis of the industry loadings of the Fama and French factors. In Section 3.4 we perform the construction of the hedge portfolios, and empirically test the effectiveness of this hedging. Section 3.5 concludes.

3.2 Theory

Consider a single-period setting, with N risky assets and risk-free asset whose returns are generated according to a K factor structure:

$$\tilde{\mathbf{r}}_t = \mathbf{b}_{t-1}(\tilde{\mathbf{f}}_t + \boldsymbol{\lambda}_t) + \tilde{\boldsymbol{\epsilon}}_t \tag{3.1}$$

where $\mathbf{r}_t$ is N×1 vector of the period t realized excess returns of the N assets; $\tilde{\mathbf{f}}_t$ is a K×1 vector of the period t unanticipated factor returns, with $\mathbb{E}_{t-1}[\mathbf{f}_t] = \mathbf{0}$, and $\boldsymbol{\lambda}_t$ is the K×1 vector of premia associated with these factors. $\mathbf{b}_{t-1}$ is the N×K matrix of factor loadings, and $\tilde{\boldsymbol{\epsilon}}_t$ is the N×1 vector of (uncorrelated) residuals. We assume that $N \gg K$, and that N is sufficiently large so that well diversified portfolios can be constructed with any factor loadings.[5]

Factor Representation: As is well known, there is a degree of ambiguity in the choice of the factors. Specifically, any set of the factors that span the K-dimensional space of non-diversifiable risk can be chosen, and the factors can be arbitrarily scaled. Therefore, without loss of generality, we rotate and scale the factors so that:[6]

$$\boldsymbol{\lambda}_{t-1} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Omega}_{t-1} = \mathbb{E}_{t-1}[\tilde{\mathbf{f}}\tilde{\mathbf{f}}'] = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_K^2 \end{bmatrix} \tag{3.2}$$

We further define:

$$\begin{aligned} \boldsymbol{\mu}_{t-1} &= \mathbb{E}_{t-1}[\tilde{\mathbf{r}}] \\ \boldsymbol{\Sigma}^{\epsilon}_{t-1} &= \mathbb{E}_{t-1}[\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}'] \\ \boldsymbol{\Sigma}_{t-1} &= \mathbb{E}_{t-1}[\tilde{\mathbf{r}}\tilde{\mathbf{r}}'] = \mathbf{b}\boldsymbol{\Omega}_{t-1}\mathbf{b}' + \boldsymbol{\Sigma}^{\epsilon} \end{aligned}$$

where $\boldsymbol{\mu}_{t-1}$ and $\boldsymbol{\sigma}^2_{\epsilon}$ are N×1 vectors. Given we have chosen the K factors to summarize the asset covariance structure, $\boldsymbol{\Sigma}^{\epsilon}_{t-1} = \mathbb{E}_{t-1}[\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}']$ is a diagonal matrix, (i.e., with the residual variances on the diagonal, and zeros elsewhere).

[5] We note that, in a finite economy, the breakdown of risk into systematic and idiosyncratic is problematic. See Grinblatt and Titman (1983), Bray (1994) and others.

[6] The rotation is such that the first factor captures all of the premium. The scaling of the first factor is such that its expected return is 1. The other factors form an orthogonal basis for the space of non-diversifiable risk, but the scaling for all but the first factor is arbitrary

3.2.1 Characteristic-Based Return Factors

Over that last several decades, academic studies have documented that certain characteristics (market capitalization, price-to-book values ratios, past returns, etc.) are related to expected returns. In response to this evidence, Fama and French (1993; 2015), Carhart (1997), Pastor and Stambaugh (2003), Frazzini and Pedersen (2014) and numerous other researchers have introduced "return factors" based on characteristics. The literature has then tested whether these characteristic-weighted factors can explain the cross-section of returns, in the sense that some linear combination of the factor portfolios is mean-variance-efficient.

What we'll assume going forward is that we can identify a vector of characteristics that perfectly captures expected returns, that is such that: $\mathbf{c}_{t-1} = \kappa \boldsymbol{\mu}_{t-1}$. We'll further assume that $\mathbf{c}_{t-1}$ is an N×1 vector, that is that a single characteristic summarizes expected returns. However, this is to simplify things; it should be a straightforward extension when there are multiple characteristics that summarize returns.[7]

What we'll assume here is that a factor-portfolio is formed based on our single vector of characteristics $\mathbf{c}$. That is, the weights of the portfolio are assumed to be proportional to the characteristic. We normalize this portfolio so as to guarantee that it has a unit expected return:[8]

$$\mathbf{w}_c = \kappa \left(\frac{\mathbf{c}}{\mathbf{c}'\mathbf{c}} \right) = \frac{\boldsymbol{\mu}}{\boldsymbol{\mu}'\boldsymbol{\mu}} \tag{3.3}$$

Note that, given this normalization, $\mathbf{w}_c'\boldsymbol{\mu} = 1$, as desired.

3.2.2 Relation between the characteristic-weighted and MVE portfolio

Assuming no arbitrage in the economy, there exists a stochastic discount factor that prices all assets, and a corresponding mean-variance-efficient portfolio. In our setting the weights of the MVE portfolio, scaled so as to give the portfolio unit expected return, has weights:

$$\mathbf{w}_{\text{MVE}} = \left(\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)^{-1} \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \tag{3.4}$$

The variance of the portfolio is $\sigma^2_{\text{MVE}} = \left(\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)^{-1}$, so the Sharpe-ratio of the portfolio is $SR_{\text{MVE}} = \sqrt{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}$.

Aside: The MVE portfolio and the SDF

The pattern of returns of the MVE portfolio across states is important to understanding preferences of the marginal investor in the economy. From Hansen and Jagannathan (1991), in a frictionless economy for an optimizing investor with stochastic discount factor m, and for any zero investment portfolio with return r_p:

$$\begin{aligned}
\mathbb{E}[m r_p] &= 0 \\
cov(m, r_p) + \mathbb{E}[m]\mathbb{E}[r_p] &= 0 \\
\sigma_m \sigma_p \rho_{m,p} &= -\mathbb{E}[m]\mathbb{E}[r_p] \\
\Rightarrow \frac{\mathbb{E}[r_p]}{\sigma_p} &= \rho_{m,p} R_f \sigma_m
\end{aligned}$$

The LHS of the final equation is the Sharpe-ratio of the portfolio. As noted by Hansen and Jagannathan (1991), this relation implies that the maximum Sharpe-ratio in the economy is

[7] For example, from multiple characteristics, one could form a single variable which is a linear combination of the various characteristics.

[8] The typical normalization in building factor portfolios is that they are "$1-long, $1 short" zero investment portfolios. However since we are dealing with excess returns, this normalization is arbitrary and has no effect on the ability of the factor-portfolios to explain the cross-section of average returns.

bounded by $\sigma_m R_f$, where R_f—the gross risk free rate—is close to 1. Also, the portfolio that has the highest Sharpe-ratio in the economy must necessarily be the portfolio with the most negative correlation with the stochastic discount factor. Also, the magnitude of the Sharpe-ratio places a bound on the volatility of the stochastic discount factor in the economy: σ_m must be greater than the maximum Sharpe-ratio in the economy (scaled by the gross risk-free rate).

The characteristic-scaled and MVE portfolios

Given our scaling of returns, the βs of the risky asset w.r.t the MVE portfolio are equal to the assets' expected excess returns:[9]

$$\begin{aligned}\boldsymbol{\beta}_{\text{MVE}} &= \frac{cov(\mathbf{r}, r_{\text{MVE}})}{var(r_{\text{MVE}})} \\ &= \frac{\boldsymbol{\Sigma}\mathbf{w}_{\text{MVE}}}{\mathbf{w}'_{\text{MVE}}\boldsymbol{\Sigma}\mathbf{w}_{\text{MVE}}} \\ &= \boldsymbol{\mu}\end{aligned}$$

We can then project each asset's return onto the MVE portfolio:

$$\begin{aligned}\mathbf{r} &= \boldsymbol{\beta}_{\text{MVE}} r_{\text{MVE}} + \mathbf{u} \\ &= \boldsymbol{\mu} r_{\text{MVE}} + \mathbf{u}\end{aligned} \tag{3.5}$$

$\mathbf{u}$ is the component of each asset's return that is uncorrelated with the return on the MVE portfolio, which is therefore unpriced risk.

Given the structure of the economy laid out in equations (3.1) and (3.2),

$$r_{\text{mve}} = (1 + f_1)$$

where f_1 denotes the first element of $\mathbf{f}$ (and the only priced factor). This means that, referencing equation (3.5),

$$\boldsymbol{\beta}_{\text{MVE}} = \boldsymbol{\mu} = \mathbf{b}_1 = \frac{1}{\kappa}\mathbf{c} \tag{3.6}$$

Finally, this means that we can write the residual from the regression in equation (3.5) as:

$$\mathbf{u} = \mathbf{b}_U \mathbf{f}_U + \boldsymbol{\epsilon}$$

where $\mathbf{b}_U$ is the N×(K-1) matrix which is $\mathbf{b}$ with the first row deleted (i.e., the loadings of the N assets on the (K-1) unpriced factors), and $\mathbf{f}_U$ is the (K-1)×1 vector consisting of the 2nd through Kth elements of $\mathbf{f}$ (i.e., the *U*npriced factors).

We will use this projection to study the efficiency of the characteristic-weighted portfolio. Since both the characteristic-weighted and MVE portfolio have unit expected returns, the increase in variance in moving from the MVE portfolio to the characteristic portfolio can tell us how inefficient the characteristic-weighted portfolio is. From equations (3.3) and (3.5), we have:

$$\begin{aligned}r_c &= \mathbf{w}_c \mathbf{r} \\ r_c &= r_{MVE} + (\boldsymbol{\mu}'\boldsymbol{\mu})^{-1}\boldsymbol{\mu}'\mathbf{u} \\ r_c &= r_{MVE} + (\mathbf{b}_1'\mathbf{b}_1)^{-1}\mathbf{b}_1'\mathbf{u} \\ r_c - r_{MVE} &= (\boldsymbol{\mu}'\boldsymbol{\mu})^{-1}\boldsymbol{\mu}'\mathbf{u} \\ &= (\boldsymbol{\mu}'\boldsymbol{\mu})^{-1}\boldsymbol{\mu}'\mathbf{b}_U\mathbf{f}_U + (\boldsymbol{\mu}'\boldsymbol{\mu})^{-1}\boldsymbol{\mu}'\boldsymbol{\epsilon}\end{aligned}$$

Thus

$$var(r_c - r_{\text{MVE}}) = \sum_{k=2}^{K} [\underbrace{(\mathbf{c}'\mathbf{c})^{-1}(\mathbf{c}'\mathbf{b}_{U,k})}_{\equiv \beta_{k,c}}]^2 \sigma_k^2 \tag{3.7}$$

[9] For the third equality, just substitute $\mathbf{w}_{\text{MVE}}$ from equation 3.4 into the second.

What is the interpretation of (3.7)? $\beta_{k,c}$ is the coefficient from a cross-sectional regression of the kth (unpriced) factor loading on the characteristic.[10] Even though the K factors are uncorrelated, the *loadings on the factors in the cross-section* are potentially correlated with each other, and this regression coefficient could potentially be large for some factors. Indeed, the necessary and sufficient conditions for the characteristic-sorted portfolio to price all assets are that

$$\beta_{k,c} = 0 \quad \forall \quad k \in \{2, \ldots, K\}.$$

This condition is unlikely hold even approximately. For example, as we show later, in the middle of the financial crisis, many firms in the financial sector were high expected return (high μ). However, these firms also had a high loading on the finance industry factor (σ_k^2 was high). Because $\boldsymbol{\mu}$ (the expected return based on the characteristics) and $\mathbf{b}_{U,k}$ (the loading on the unpriced finance industry factor) were highly correlated, the characteristics-sorted portfolio has high industry factor risk, meaning that it has a lower Sharpe-ratio than the MVE portfolio. Because σ_k^2 was quite high in this period, the extra variance of the characteristic-sorted portfolio was arguably also large. In Section 3.4, we show how this extra variance can be diagnosed and taken into account.

3.2.3 An optimized characteristic-based portfolio

It follows from the previous discussion that the optimized characteristic-based portfolio is

$$\mathbf{w}_c^* = \kappa \left(\frac{\boldsymbol{\Sigma}^{-1}\mathbf{c}}{\mathbf{c}'\boldsymbol{\Sigma}^{-1}\mathbf{c}} \right) = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \tag{3.8}$$

Clearly the challenge is the actual construction of such a portfolio. For instance, there are well known issues associated with estimating $\boldsymbol{\Sigma}$ and using it to do portfolio formation. In the next subsection, we develop an alternative approach for testing portfolio optimality.

Assuming the characteristics model is correct, and one observes the characteristics, it is straightforward to test the optimality of the characteristics-sorted portfolio. All that is needed is some (ex-ante) instrument to forecast the component of the covariances which is orthogonal to the characteristics. If the characteristic sorted portfolio is optimal (i.e., MVE) then characteristics must line up with betas with the characteristics sorted-portfolio *perfectly*. If they don't (and the characteristics model holds) then the portfolio can't be optimal.

Moreover, one can improve on the optimality of the portfolio by following the procedure advocated in this paper, by, first, identifying assets with high (low) alphas relative to the characteristic-sorted portfolio (again based on the characteristic model) and, second, building a portfolio with the highest possible expected alpha relative to the characteristic sorted portfolio, under the characteristic hypothesis. If this portfolio has a positive alpha then the optimality of the characteristics-sorted portfolio is established. This is the empirical approach we take in this paper.

3.3 Industry Factor Loadings

Asness, Porter, and Stevens (2000) and Cohen and Polk (1995) and others show that if book-to-price ratios are decomposed into an industry-component and a within-industry component, then only the within-industry component— that is, the difference between a firm's book-to-price ratio and the book-to-price ratio of the industry portfolio—forecasts future returns.[11] This suggests that any exposure of HML to industry factors is likely unpriced, and therefore that if

[10] Note that we get the same expression, up to a multiplicative constant, if we instead regress the unpriced factor loadings on the the priced factor loadings, or on the expected returns, given the equivalence in equation (3.6).

[11] See also Lewellen (1999) and Cohen, Polk, and Vuolteenaho (2003).

this exposure were hedged out, it would result in a factor-portfolio with lower risk, but the same expected return, i.e., with a higher Sharpe ratio.

Here we show that the HML exposure on industry factors varies dramatically over time, and that at selected times the exposure can be very high. We highlight two episodes in particular in which the correlation between HML and industry factors exceeds 95%: in late-2000/early-2001 as high-technology firms earned large negative returns and became highly volatile, and 2008-2009 during the financial-crisis and a parallel episode for financial firms. In both of these episodes the past return performance of the industry led to the vast majority of the firms in the industry becoming either growth or value firms—that is, there was a high cross-sectional correlation between valuation ratios and industry membership— leading to a high β of HML on that industry. This combination of a high β and high factor volatility leads to the high correlations we observe in the data.

Figure 3.1 plots the adjusted R^2 from 126-day rolling regressions of daily HML returns on the twelve daily value-weighted industry excess returns. The time period is January 1981-December 2015.[12] The plot shows that, while there are short periods where the realized R^2 dips below 0.5, there are also several periods where it exceeds 0.9.

Figure 3.2 plots, for the same set of daily, 126-day rolling regressions, the betas for each of the 12 industries. The upper panel makes it clear that there is considerable time-variation in the industry betas. To provide a little clarity, the lower panel of Figure 3.2 breaks out just two of these industries, "Money" and "Business Equipment." The two industries are selected because, in the post-1995 period, they are generally the industries to which HML has the highest- and lowest- exposures, respectively. Particularly, for the Money/finance industry, the lower panel of Figure 3.2 shows that the HML exposure to the finance industry falls below 0 for a short period shortly after 2000, but then rises dramatically, particularly during the 2008-9 financial crisis, a high of about 0.5.

In contrast, HML has a negative exposure to the BusEq industry, which contains many of the "tech" firms that earned very high returns and reached high valuation ratios leading up to March 2000. Figure 3.3 plots the rolling-126 day volatility of these two industries.[13] Here, two periods in particular stand out for these two industries: the 2000-2002 period for tech, and the 2008-2009 period for the finance industry. For the finance industry, the realized 126-day volatility peaks at approximately 90% (annualized). The high finance industry beta in Figure 3.2 and the high industry volatility at the same time, as seen in Figure 3.3 suggest that in these two periods, HML returns are likely to be highly correlated with the tech industry and the finance industry, respectively.

Figure 3.4 plots the adjusted R^2, and confirms that this is the case for the finance industry in the financial crisis. In fact, over the full 2008:07-2009:06 period, the correlation between the market-adjusted return $R_{HML} - r_{Mkt}$ and market-adjusted financial sector return $r_{Fin} - r_{Mkt}$ is 89.5%. Why is it so high? As of December 2007, the top 4 firms by market cap in the "Money" industry (based on the FF 12-industries) were Bank of America, AIG, Citigroup and J.P. Morgan. Three of these four were in the Big/High-BM (i.e., large value) portfolio. Interestingly, the one that wasn't was AIG – it was in the middle portfolio. While the market capitalization of these firms falls dramatically through 2008, they remain large and, particularly as the volatility of the Money factor increases, these firms and others like them drive the returns both of the HML portfolio and the Money industry portfolio.

However, it is not surprising that there are firms in the Money industry that do not have high valuation ratios, even in the depths of the financial crisis. For example, in 2008 UnitedHealth Group (UNH) and American Express (AXP) (7th and 8th highest by market cap in the Money

[12] The daily HML returns, the daily industry returns and the risk-free data are taken from Ken French's data library at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library. This web site also provides a breakdown of the standard industrial classification (SIC) codes that are included in each

[13] Note that for this plot, like the other "rolling" plots in this section, the x-axis label indicates the date on which the 126-day interval ends.

portfolio) were both "L" (low book-to-market) firms. Yet both UNH and AXP have large positive loadings on HML at this point in time. The reason is that both UNH and AXP covary strongly with money industry returns, as does HML at this point in time.

In Section 3.4 we construct what we call *hedge-portfolios* for each of the Fama and French (2015) factors. For example, what the HML hedge-portfolio does is to short firms like AXP and UNH, for which the characteristic (here B/M) is low and the HML factor loading is high. Similarly, this portfolio would go long value firms in the tech sector (which have a high characteristic and a low HML factor loading). The goal is to build a portfolio with a strongly negative loading on the HML factor, but with an expected return—assuming the characteristics model is correct—of approximately zero. This portfolio can then be combined with the HML portfolio to create a more efficient "hedged" HML with the same expected return, but with lower return variance and therefore a higher Sharpe-ratio.

In the next section, we discuss the construction of this hedge-portfolio, and examine tests of the efficiency of the hedge portfolios.

3.4 Low- and High-Power Test Based Hedge Portfolios

3.4.1 Construction of the factors and the set of test portfolios

Our focus is on the five factor Fama and French (2015) model and our factors are exactly those described in Section 4 of their paper but our test portfolios are rather different. The test portfolios are triple sorted portfolios on size, a characteristic (book-to-market, profitability or investment) and an expected loading on the factor associated with the characteristic (HML, RMW, CMA). We generate two sets of test portfolios, each associated with a specific methodology to estimate the expected loadings, and show that inferences regarding asset pricing models are sensitive to the specific methodology used to instrument for the expected loadings.

The procedure is standard. We first rank NYSE firms by their, for example, book-to-market (BE/ME) ratios at the end of December of a given year and their market capitalization (ME) at the end of June of the following year. Break points are selected at the 33.3% and 66.7% marks for both the book-to-market and market capitalization sorts. Then in July of a given year all NYSE/Amex and Nasdaq stocks are placed into one of the nine resulting bins. There is an important difference though in the way the sorting procedure is implemented relative to Fama and French (1992, 1993 and 2015) or Daniel and Titman (1997) and it is that our characteristics sorted portfolios are industry adjusted. That is, whether a stock has, for example, a high or low book-to-market depends on whether it is above or below the corresponding industry average (see Cohen et al. (2003)). Our industries are the 49 industries Fama and French (1997).

Finally, each of the stocks in one of these nine bins is sorted into one of three additional bins formed based on the stocks' expected future loading on the HML factor. The firms remain in those portfolios between July and June of next year. Sorting on the characteristic and the expected loading itself identifies to what extent the variation in returns is driven by the characteristic or the loading. This last sort results in portfolios of stocks with similar characteristics (BE/ME and ME) but different loadings on HML. The sorts on operating profitability, size and loading on the profitability factor (robust minus weak or RMW) and investment, size and loading on the investment factor (conservative minus aggressive or CMA) are constructed in exactly the same way.

Clearly a key ingredient of the last step of the sorting procedure is the estimation of the expected loading on the corresponding factor. Our purpose is to obtain estimates of the future loadings in the five factor model of Fama and French (2015):

$$\begin{aligned} R_{i,t} - R_{F,t} = a_i + \beta_{Mkt-RF,i}(R_{Mkt,t} - R_{F,t}) + \beta_{SMB,i}R_{SMB_t} \\ +\beta_{HML,i}HML_t + \beta_{RMW,i}RMW_t + \beta_{CM,i}CMA_t + e_{i,t} \end{aligned} \tag{3.9}$$

We instrument future expected loadings with preformation factor loadings. To avoid the danger of "data mining" we search over different methodologies only in what concerns the estimation of the preformation loadings on HML in the context of the Fama and French (1993) three factor model in the 1933-2014 sample. Once the estimation procedure is selected, in a manner to be described shortly, we use it to estimate preformation loadings in the Fama and French (2015) five factor model.

How is then the specific estimation methodology selected? For each estimation procedure we form two portfolios, a portfolio of stocks that have a low loading on HML and another with those that have a high loading. We do this by averaging the returns of the nine book-to-market and size sorted portfolios that have low and high loadings on HML respectively. Then we construct a portfolio that goes long the low loading portfolio and short the high loading portfolio and regress the returns of this zero investment portfolio on the market, SMB and HML. We pick the method that yields the highest t-statistic for the estimate of the loading on HML. Effectively we want to maximize the spread in our estimates of the loadings as this will maximize the power of the asset pricing test.

The resulting estimation method is intuitive and is close to the method proposed by Frazzini and Pedersen (2014). These authors build on the observation that correlations are more persistent than variances (see, among others, de Santis and Gerard (1997)) and propose estimating covariances and variances separately and then combine these estimates to produce the preformation loadings. Specifically, covariances are estimated using a five-year window with overlapping log-return observations aggregated over three trading days, to account for non-synchronicity of trading. Variances of factors and stocks are estimated on daily log-returns over a one-year horizon. In addition, we introduce an additional intercept in the pre-formation regressions for returns in the six months preceding portfolio formation, i.e., from January to June of the rank-year (see Figure 1 in Daniel and Titman (1997) for an illustration). We refer to this estimation methodology as the 'high power' methodology.

This estimation method contrasts with the traditional approach of simply using as instruments for future factor loadings the result of regressing test portfolio excess returns on factors over a moving fixed-sized window based on, e.g., 36 or 60 monthly observations, skipping the most recent 6 months, i.e., those that already fall in the rank-year (see, e.g., Daniel and Titman (1997) or Davis, Fama, and French (2000)).[14] We refer to this method, which is effectively the one used by Daniel and Titman (1997), as the 'low power' method and use it to construct an alternative set of test portfolios. In addition these set of test portfolios are not industry adjusted.

In sum, the high and low power sets of test portfolios differ in two dimensions, the estimation method for the expected loading and whether the characteristics are industry adjusted or not. In what follows we present results for each set and show that standard asset pricing tests yield rather different inferences depending on which set one uses.

Table 3.1 shows the average monthly excess returns for our two sets of test portfolios. Each of the panels corresponds to a sort with respect to a specific loading. For instance Panel A shows the set of test portfolios sorted on size, book-to-market and the expected loading on HML. The top subpanel refers to the low power set of test portfolios, that is the set of test portfolios where the expected loadings are instrumented with the low power methodology, and the bottom corresponds to the test portfolios where the expected loading is estimated with the high power methodology. As we move from the left to the right column in each of the panels we are moving from low loading to high loading portfolios. In general there is a positive relation between the loading and the average return, as shown in the averages computed at the bottom of each of the panels. Still for some of the portfolios the relation is not discernible and there is no clear pattern between high and low power methodologies. For instance, consider the portfolio of small companies of medium operating profitability (portfolio (2,1) in Panel B). In the case of the low power methodology there is a negative relation between average returns and the loading

[14] Notice that in contrast, the high power method avoids discarding the most recent data.

on RMW, whereas there is a weak positive relation in the case of the high power methodology.

Even when there is a positive relation between average returns and factor loadings, a natural concern is whether the sort on loadings is simply a refined sort on the characteristic. We explore this possibility in Table 3.2 where we show the average of the relevant characteristic for each of the test portfolios. For the case of the HML factor the pattern is consistent: As we move from the left to the right column the average book to market of each of the 27 test portfolios increases as well. The pattern is also uniformly consistent for CMA and the low and medium operating profitability test portfolios, but not for the high operating profitability test portfolios. The strong correlation between the factor loading and the characteristic diminishes the power of the test to reject factor models in favor of characteristics based models, as it spuriously assigns variation in average returns to variation in the loading.

3.4.2 Postformation loadings

We estimate the postformation loadings by running a time series regression of the excess returns for each of the test portfolios on the Fama and French (2015) five factors (see equation (3.9)). To compare how our high power methodology results in larger dispersion of the postformation loadings when compared to the low power methodology, Figure 3.5 shows the postformation loadings for each of the 27 test portfolios in the different subpanels of Table 3.1. Panel A and B correspond to the low and high power methodology respectively.

Consider for example the top left panel in Figure 3.5. There are three groups of estimates, each corresponding to a particular book-to-market bin. Each of the groups has in turn three lines, corresponding each to a particular size grouping. Finally each of those lines have three points corresponding to a particular estimate of the postformation loading. The plot thus reports book-to-market on the y-axis for each of the 27 portfolios and the postformation loading on the x-axis.

First, reassuringly, both methodologies generate a positive correlation between pre and postformation loadings for HML and CMA within a characteristic (book-to-market or investment) and size bin. But in the case of the loadings on RMW, the low power methodology does not produce a consistent positive association between pre and postformation loadings. Table 3.2, which again is divided in three panels, reports the point estimates and t-statistics for the postformation loadings on the different factors. Consider Panel B, which reports the postformation loadings on the profitability factor, RMW, and focus on the test portfolios with medium operating profitability. For portfolio (2,2) there is a non-monotone relation between pre and postformation loading when the low power methodology is used to estimate the preformation loadings whereas there is a monotone relation when the high power methodology is used instead.

Second, and most importantly, as it is readily apparent from Figure 3.5 the high power methodology generates substantially more cross sectional dispersion in postformation loadings than the low power methodology, which is key to deliver a higher power test. Each of the panels of Table 3.3 reports the difference in the postformation loadings between the low and and high preformation loading sorted portfolio for each of the characteristic-size bin. Consistently this difference is much larger with the high power methodology than the low. Our high power methodology forecasts future loadings much better than the one used by Daniel and Titman (1997) or Davis et al. (2000).

3.4.3 The pricing of the test portfolios

Of particular interest are the Fama and French (2015) five factor alphas for each of the test portfolios, which are reported in Table 3.3 as well. Consider first Panel A, which is concerned with HML. As we saw, the low power methodology generates little cross sectional dispersion in postformation loadings and leaves five portfolios with statistically significant alphas. Instead, the high power panel takes that number to nine. Moreover, as predicted by our initial hypothesis,

the alphas of the high loading portfolios are all negative whereas they are mostly positive for the low loading portfolios: The five factor model is assigning too high a premium to the high loading portfolios and too low a premium to the low loading portfolios.

This pattern occurs again in the case of CMA (Panel C), when out of the 27 test portfolio the low power methodology leaves three portfolios mispriced whereas the high power one takes that number to twelve. As before the sign of the different alphas is as hypothesized initially, whereas the number of mispriced portfolios for RMW is four in both cases.

The last column, marked 1-3, reports the results associated with what Daniel and Titman (1997) refer to as a characteristic balanced portfolio, a long position on the low loading portfolio and a short position in the high loading portfolio within each of the nine characteristic-size bins. Consider the results on HML. Whereas the low power methodolology results in alphas statistically undistinguishable from zero, the high power methodology features two misspriced characteristics balanced portfolios. For the case of CMA as many as five of these characteristics balanced portfolios are misspriced. The magnitude of the misspricing is relevant: For instance, in the case of the HML factor, the alpha for the characteristic balanced portfolio (3,1) (the small value portfolio) is 0.3% per month or about 3.7% per year.

3.4.4 The main result

Table 3.0 is the main result of the paper. We form long-short portfolios as follows. For each of the five factors in the Fama and French (2015) model we form a portfolio that goes long the low loading on the corresponding factor, averaging across the corresponding characteristic and size, and short the high loading portfolio. For instance consider the line labeled HML. There we take a long position in the low loading portfolios, weighting the corresponding nine book-to-market size sorted portfolios equally, and a short position in the high loading portfolios in the same manner.

We then run a single time series regression of the returns of these portfolios on the five factors. Table 3.0 reports the alphas and loadings as well as the corresponding t-statistics. Panel A focuses on the set of test portfolios where preformation loadings are estimated with the low power methodology and Panel B focuses on the high power one. As before, we focus on the alphas. Whereas, when using the low-power methodology, the Fama and French factors price all these portfolios correctly, the five factor model fails to price four out of five of the high-power long-short test portfolios (the only one for which the Fama and French model cannot be rejected is the "SMB" portfolio.) The last line of each of the panels constructs equal weighted combinations of these portfolios. The alphas for all of them are strongly statistically significant in the high power test whereas this is not the case for the low power methodology.

3.4.5 Industry adjustments

There are two differences in the construction of our set of test portfolios when compared to Daniel and Titman (1997) or Davis et al. (2000): The estimation procedure for the preformation loadings and the fact that the sort on the characteristics is industry adjusted. We revisit next the issue of the industry adjustment. Table 3.1 conducts the same tests as in Table 3.0 with the only difference that in Panel A, where we report the low power procedure results, the test portfolios are now constructed with the industry adjustment and in Panel B, the test portfolios are constructed without the industry adjustment (but still with the high-power loading forecasts).

Consider first panel A. To the low power procedure, whether the test portfolios are constructed with or without the industry adjustment is irrelevant: The alphas are all statistically no different than zero and in these set of test portfolios the Fama and French (2015) five factor model fails to be rejected. When we apply the methodology where the test portfolios are not industry adjusted there is only one significant change: The alpha of the portfolio that goes long the portfolios of stocks that have low loadings on HML and short the portfolios with high loadings on HML is

now statistically no different than zero. Still the five factor model is rejected, as it was the case when the industry adjustment was made, for the combination portfolios, which are reported at the bottom of Panel B. Thus, though industry adjustment ex-ante seems a sensible correction when constructing test portfolios sorted on characteristics, it does not seem to alter the inferences drawn from the asset pricing tests developed in this paper dramatically.

3.5 Conclusions

A set of factor portfolios can only explain the cross-section of average returns if the mean-variance efficient portfolio is in the span of these factor-portfolios. There are numerous sources of information from which to construct such a set of factors. In the cross-sectional asset pricing literature, the most widely utilized source of information used to form factor-portfolios have been observable firm characteristics such as the ones we examine here: firm size, book-to-market ratio, and accounting-based measures of profitability and investment. Portfolios formed going long high-characteristic firms and short low-characteristic firms ignore the forecastable part of the covariance structure, and thus cannot explain the returns of portfolios formed using the characteristics and past-returns. Factor-portfolios formed in this way are therefore inefficient with respect to this information set.

In the empirical part of this paper, we have examined one particular model in this literature: the five-factor model of Fama and French (2015). Our empirical findings show that the factor-portfolios that underlie this model contain large unpriced components, which we show are at least correlated with unpriced factors such as industry risk. When we add information from the historical covariance structure of returns we can vastly improve the efficiency of these factor portfolios, generating a portfolio that is orthogonal to the original five factors and has a Sharpe-ratio of 0.883. It is important to note that we are extremely conservative in the way in which we construct these hedged-portfolios: following Fama and French (1993), we form portfolios annually, and value-weight these portfolios. By hedging out the ex-ante identifiable, unpriced risk in the five-factors, we increase the annualized squared-Sharpe ratio achievable with these factors from 1.37 to 2.15.

Hedged factors like those we construct here raise the bar for standard asset pricing tests. By the logic of Hansen and Jagannathan (1991), a pricing kernel variance of at least 2.15 (annualized) is required to explain the returns of the hedged-factor-portfolios. Also, because the hedged factor portfolios are far less correlated with industry factors, etc., they are also far less likely to be correlated with variables that might serve as plausible proxies for marginal utility.

In addition, the hedged factor portfolios we generate can serve as an efficient set of benchmark portfolios for doing performance measurement using Jensen (1968) style time-series regressions. Such an approach will deliver the same conclusions as the characteristics approach (Daniel, Grinblatt, Titman, and Wermers, 1997), while maintaining the convenience of the factor regression approach.

3.6 Figures

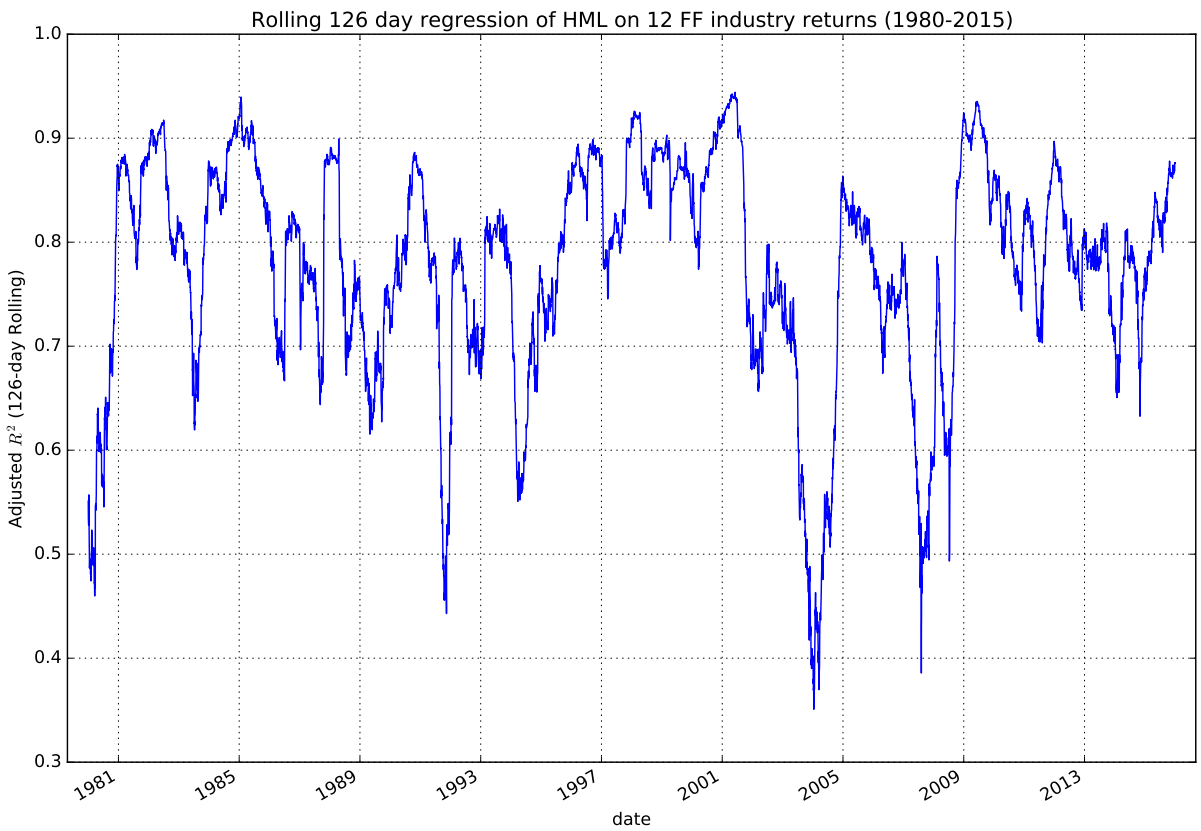


Figura 3.1: **Rolling regression R^2s – HML returns on industry returns** This figure plots the adjusted R^2 from 126-day rolling regressions of daily HML returns on the twelve daily industry excess returns. The time period is January 1981-December 2015.

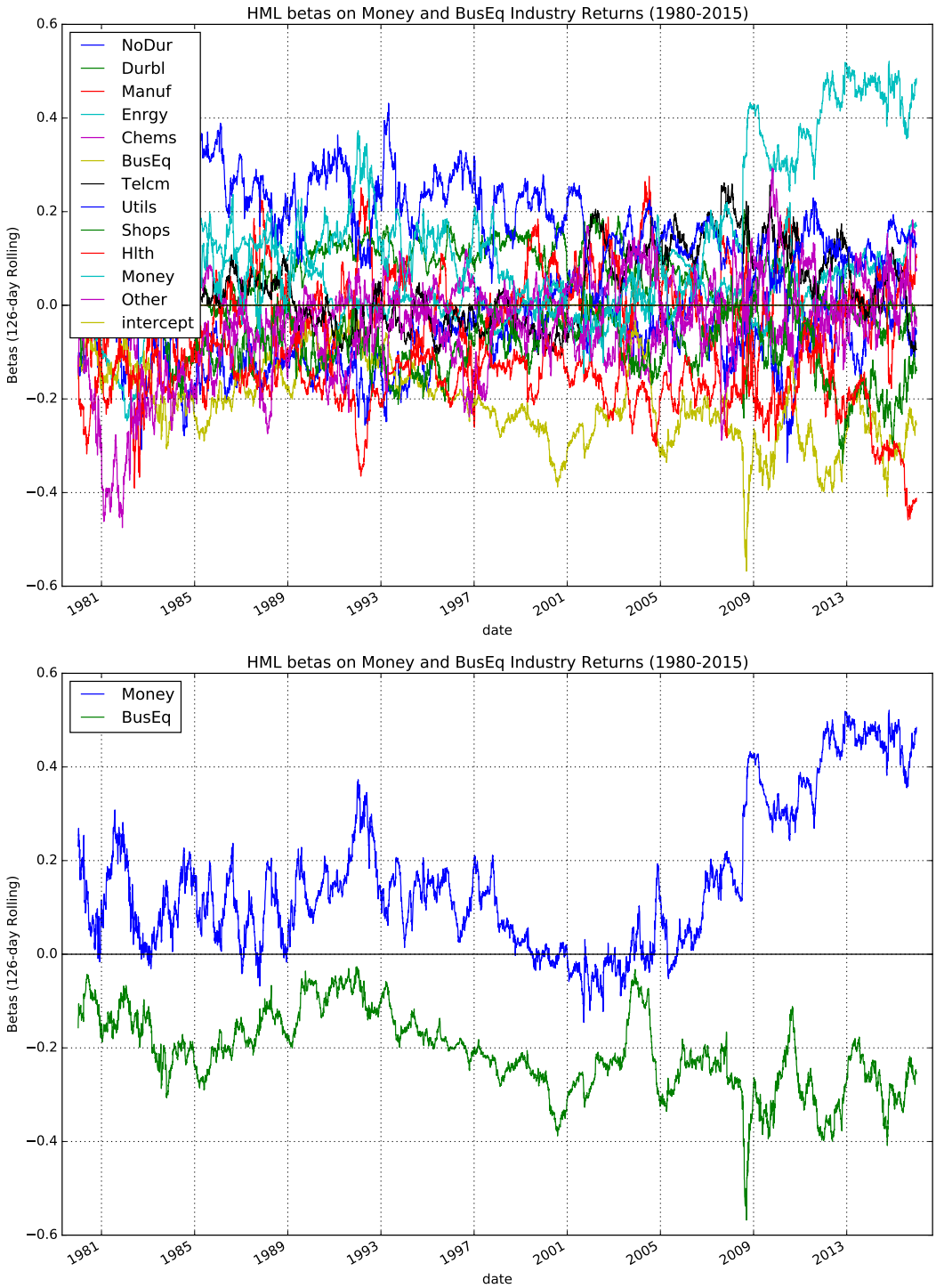


Figura 3.2: **HML loadings on industry factors.** The upper panel of this figure plots the betas from rolling 126-day regressions of the daily returns to the HML-factor portfolio on the twelve daily industry excess returns over the January 1981-December 2015 time period. The lower panel plots only the betas for the Money and Business Equipment industry portfolios, and excludes the other 10 industry factors.

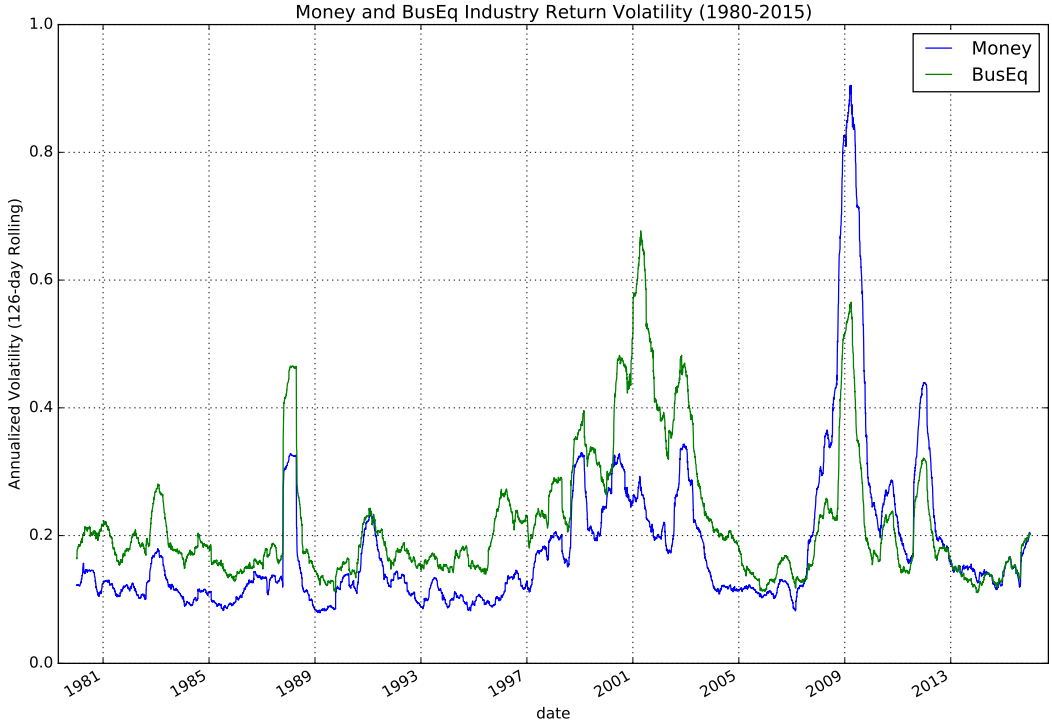


Figura 3.3: **Volatility of the money and business equipment factors.** This figure plots 126-day volatility of the daily returns to the Money and the Business Equipment factors over the January 1981-December 2015 time period.

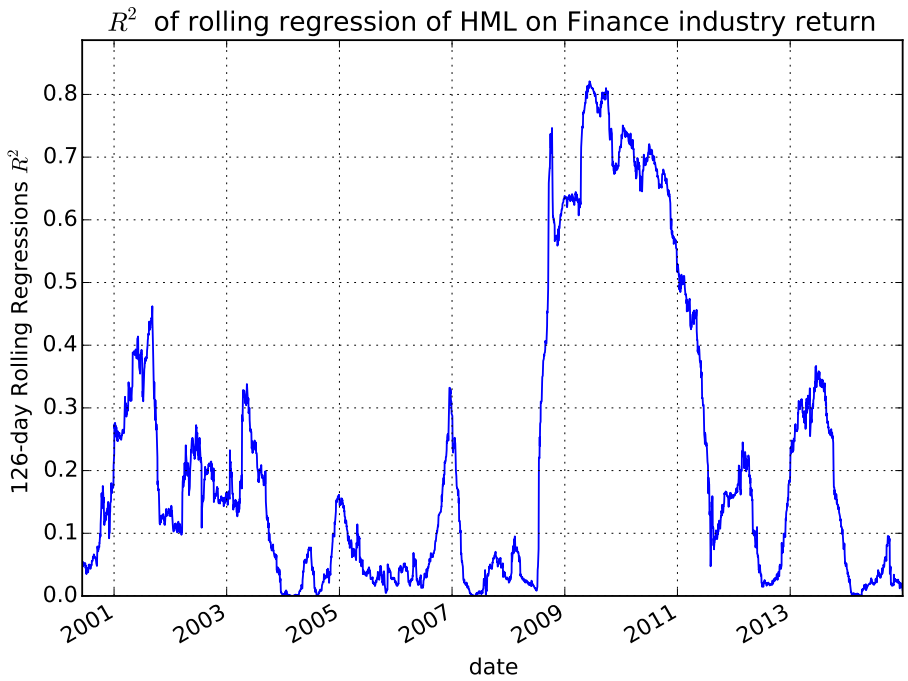


Figura 3.4: **Rolling regression R^2s – HML returns on *Money* industry returns.** This figure plots the adjusted R^2 from 126-day rolling regressions of daily HML returns on the daily *Money* industry returns from the 12 industry returns. The time period is January 1981-December 2015.

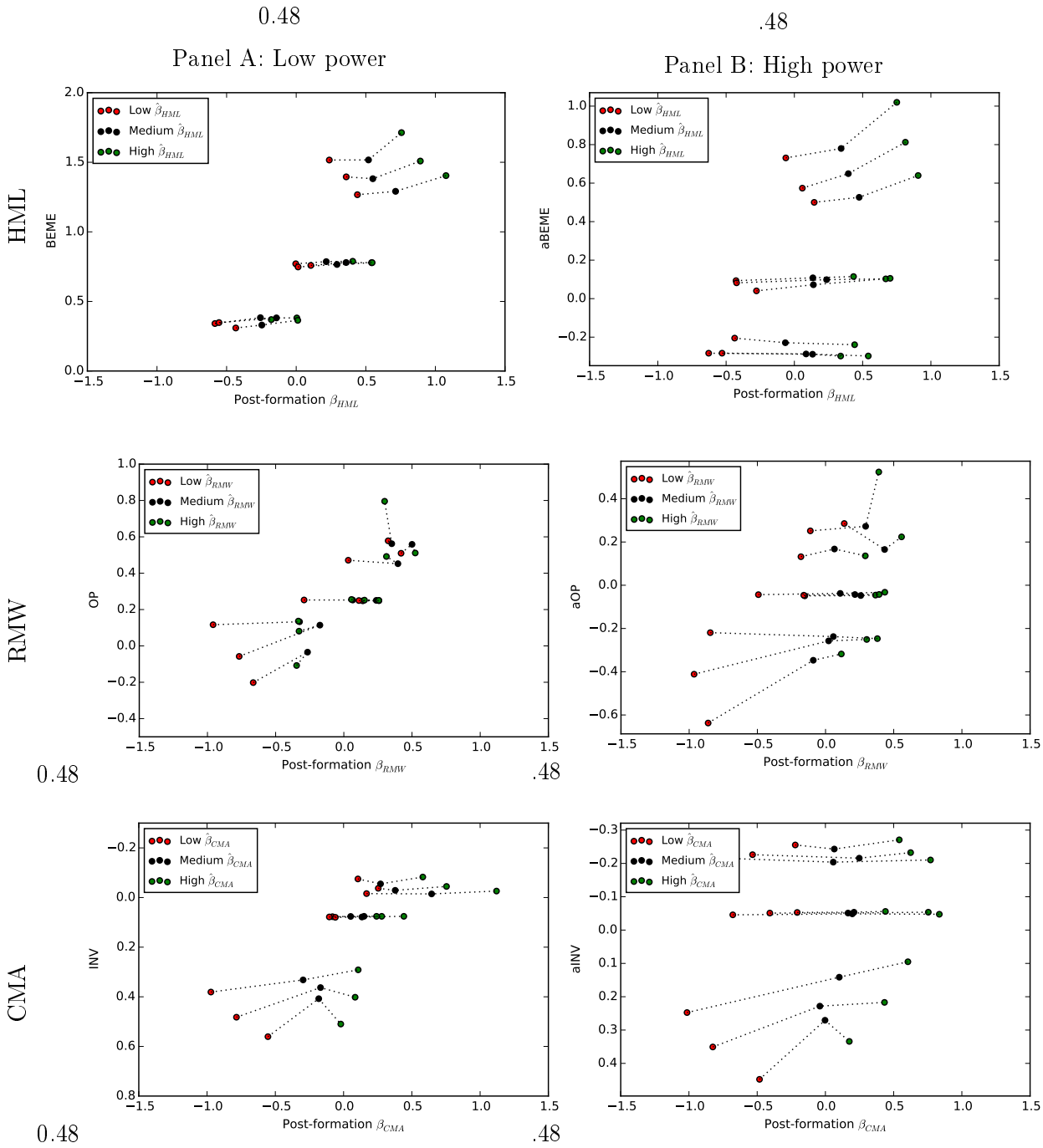


Figura 3.5: **Ex-post loading vs. characteristic.** This figure shows the time series average of post-formation factor-loading on the x-axis and the time series average of the respective characteristic on the y-axis of each of the 27 portfolios formed on size, characteristic and factor-loading. Panels A uses the low power methodology and B uses the high power methodology. The first row uses sorts on book-to-market and HML-loading, the second one operating profitability and RMW-loading and the last one investment and CMA-loading.

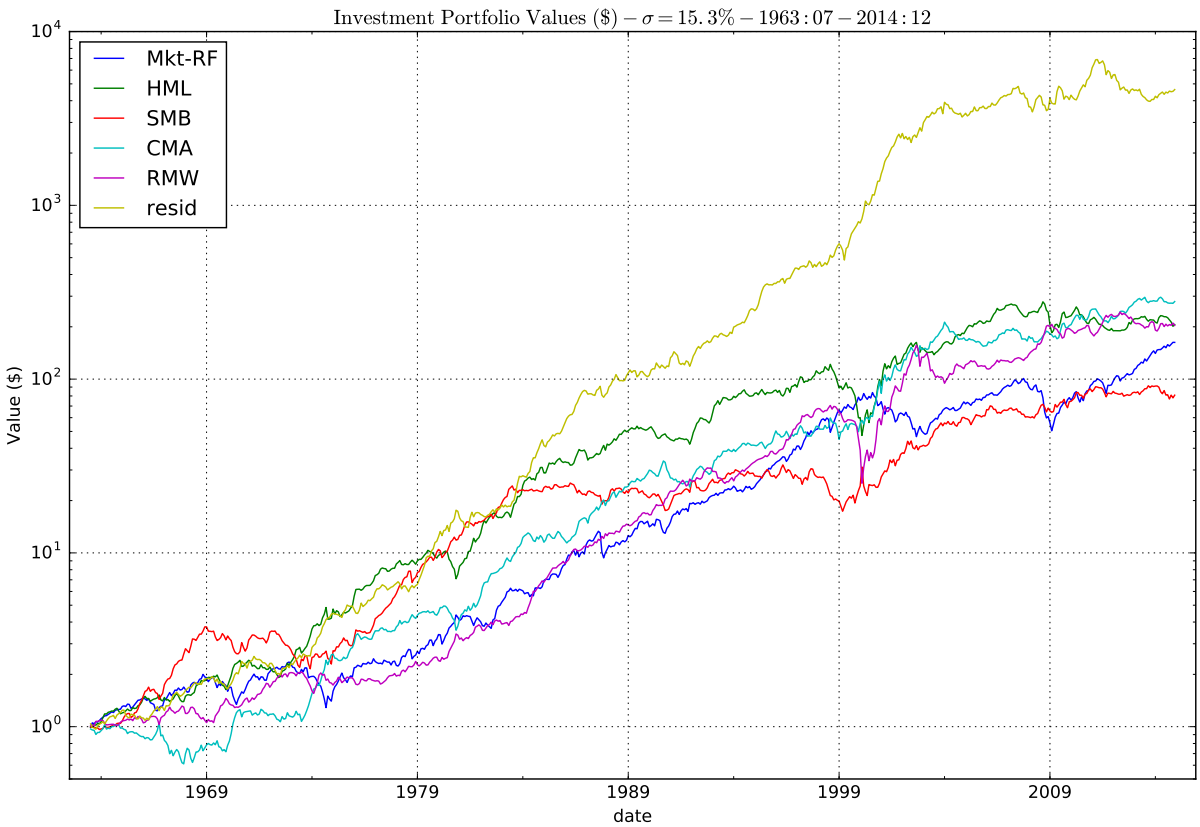


Figura 3.6: **Portfolio Cumulative Returns.** This figure plots the cumulative returns of of the five FF(2015) portfolios, and the residual portfolio. The residual portfolio is the equal-weighted combination of the HML, RMW, and CMA hedge portfolios, orthogonalized to the five-factors. Each portfolio assumes an investment of \$1 at close on the last trading day of June 1963, and earns a return of $(1 + r_{LS,t} + r_{f,t})$ in each month t, where $r_{LS,t}$ is the long-short portfolio return, and $r_{f,t}$ is the one month risk free rate.

3.7 Tables

Tabela 3.1: **Average monthly excess returns for the test portfolios.**

The sample period is July 1963 to December 2014. Stocks are sorted into 3 portfolios based on the respective characteristic - book-to-market (BEME), operating profitability (OP) or investment (INV) and independently into 3 size (ME) groups. These are depicted row-wise and indicated in the first two columns. Last, portfolios are sorted into 3 further porfolios based on the loadings forecast, conditional on the first two sorts. These portfolios are displayed column-wise. The last column shows average returns of all 9 respective characteristic portfolios. The last row shows averages of all 9 respective loadings portfolios. In the top panels we use the low power and in the bottom panels we use the high power methodology.

Panel A: HML

Low power

Char-Portfolio		$\hat{\beta}_{HML}$-Portfolio			
BEME	ME	1	2	3	Avg.
1	1	0.50	0.57	0.58	0.55
	2	0.46	0.65	0.67	
	3	0.53	0.41	0.54	
2	1	0.87	0.89	0.97	0.74
	2	0.81	0.70	0.82	
	3	0.55	0.42	0.66	
3	1	1.07	1.03	1.06	0.90
	2	0.94	0.92	1.05	
	3	0.74	0.58	0.70	
Avg.		0.72	0.68	0.78	

High power

Char-Portfolio		$\hat{\beta}_{HML}$-Portfolio			
BEME	ME	1	2	3	Avg.
1	1	0.34	0.64	0.70	0.55
	2	0.48	0.63	0.76	
	3	0.43	0.51	0.42	
2	1	0.76	0.83	0.99	0.73
	2	0.71	0.75	0.85	
	3	0.56	0.57	0.53	
3	1	1.01	1.08	1.07	0.91
	2	1.00	0.87	1.11	
	3	0.77	0.69	0.59	
Avg.		0.67	0.73	0.78	

Panel B: RMW

Low power

Char-Portfolio		$\hat{\beta}_{RMW}$-Portfolio			
OP	ME	1	2	3	Avg.
1	1	0.65	0.80	0.75	0.59
	2	0.70	0.61	0.66	
	3	0.27	0.40	0.48	
2	1	0.91	0.85	0.84	0.71
	2	0.76	0.75	0.78	
	3	0.48	0.41	0.59	
3	1	0.95	1.07	0.97	0.80
	2	0.76	0.82	0.94	
	3	0.58	0.50	0.60	
Avg.		0.67	0.69	0.73	

High power

Char-Portfolio		$\hat{\beta}_{RMW}$-Portfolio			
OP	ME	1	2	3	Avg.
1	1	0.62	0.87	0.84	0.67
	2	0.65	0.65	0.88	
	3	0.33	0.53	0.66	
2	1	0.86	0.92	0.90	0.72
	2	0.73	0.69	0.85	
	3	0.36	0.59	0.53	
3	1	0.88	0.98	1.07	0.76
	2	0.80	0.76	0.87	
	3	0.52	0.45	0.50	
Avg.		0.64	0.72	0.79	

Panel C: CMA

Low power

Char-Portfolio		$\hat{\beta}_{CMA}$-Portfolio			
INV	ME	1	2	3	Avg.
1	1	0.97	0.98	1.01	0.82
	2	0.83	0.86	0.83	
	3	0.57	0.66	0.66	
2	1	0.99	0.92	0.99	0.78
	2	0.85	0.85	0.83	
	3	0.52	0.43	0.61	
3	1	0.55	0.70	0.63	0.60
	2	0.52	0.73	0.78	
	3	0.43	0.43	0.63	
Avg.		0.69	0.73	0.77	

High power

Char-Portfolio		$\hat{\beta}_{CMA}$-Portfolio			
INV	ME	1	2	3	Avg.
1	1	0.92	0.90	1.02	0.78
	2	0.82	0.93	0.76	
	3	0.51	0.51	0.62	
2	1	0.94	1.01	1.01	0.78
	2	0.91	0.86	0.77	
	3	0.52	0.43	0.56	
3	1	0.61	0.76	0.65	0.64
	2	0.63	0.77	0.66	
	3	0.48	0.57	0.61	
Avg.		0.70	0.75	0.74	

Tabela 3.2: **Average monthly characteristics for the test portfolios.**
Stocks are sorted into 3 portfolios based on the respective characteristic - book-to-market (BEME), operating profitability (OP) or investment (INV) and independently into 3 size (ME) groups. These are depicted row-wise and indicated in the first two columns. Last, portfolios are sorted into 3 further porfolios based on the loadings forecast, conditional on the first two sorts. These portfolios are displayed column-wise. At each yearly formation date, the average respective characteristic (BEME, OP, or INV) for each portfolio is calculated, using value weighting. At each point, the characteristic is divided by the NYSE median at that point in time. The time series from 1963 to 2014 is then averaged to get the numbers that are presented in the table below. Note that, while sorts in the high power panels are based on industry-adjusted characteristics, the averages and medians are calculated based on the unadjusted characteristics. The last column shows average characteristics of all 9 respective characteristic portfolios. The last row shows averages of all 9 respective loadings portfolios. In the top panels we use the low power and in the bottom panels we use the high power methodology.

Panel A: HML

Low power

Char-Portfolio		$\hat{\beta}_{HML}$-Portfolio			
BEME	ME	1	2	3	Avg.
1	1	0.44	0.5	0.48	0.47
	2	0.45	0.5	0.5	
	3	0.41	0.43	0.48	
2	1	1.01	1.03	1.04	1.01
	2	0.99	1.02	1.02	
	3	0.98	1.01	1.02	
3	1	2	2	2.28	1.92
	2	1.85	1.84	2	
	3	1.69	1.71	1.87	
Avg.		1.09	1.12	1.19	

High power

Char-Portfolio		$\hat{\beta}_{HML}$-Portfolio			
BEME	ME	1	2	3	Avg.
1	1	0.44	0.62	0.72	0.58
	2	0.43	0.6	0.76	
	3	0.4	0.51	0.73	
2	1	0.78	0.94	1.05	0.91
	2	0.75	0.93	1.11	
	3	0.64	0.86	1.15	
3	1	1.64	1.79	2.22	1.76
	2	1.46	1.69	2.06	
	3	1.41	1.62	1.97	
Avg.		0.88	1.06	1.31	

Panel B: RMW

Low power

Char-Portfolio		$\hat{\beta}_{RMW}$-Portfolio			
OP	ME	1	2	3	Avg.
1	1	-0.85	-0.19	-0.46	0.06
	2	-0.27	0.45	0.31	
	3	0.46	0.53	0.54	
2	1	0.99	1	1.01	1.01
	2	1	1	1.01	
	3	1.02	1.01	1.02	
3	1	2.36	2.29	3.25	2.22
	2	2.08	2.23	2.08	
	3	1.91	1.83	1.99	
Avg.		0.97	1.13	1.19	

High power

Char-Portfolio		$\hat{\beta}_{RMW}$-Portfolio			
OP	ME	1	2	3	Avg.
1	1	-1.31	0.06	0.31	0.25
	2	-0.36	0.59	0.73	
	3	0.45	0.79	1.03	
2	1	0.89	0.98	1.06	1.02
	2	0.88	1	1.12	
	3	0.86	1.08	1.32	
3	1	2.08	2.21	3.34	2.14
	2	2.16	1.8	2.11	
	3	1.58	1.93	2.02	
Avg.		0.8	1.16	1.45	

Panel C: CMA

Low power

Char-Portfolio		$\hat{\beta}_{CMA}$-Portfolio			
INV	ME	1	2	3	Avg.
1	1	-2.02	-1.6	-2.16	-1.4
	2	-1.31	-1.21	-1.5	
	3	-0.91	-0.78	-1.13	
2	1	1.05	1.02	1.02	1.04
	2	1.05	1.02	0.99	
	3	1.07	1.09	1.03	
3	1	9.4	6.8	8.84	7.08
	2	7.98	6.14	6.73	
	3	7	5.81	5.01	
Avg.		2.59	2.03	2.09	

High power

Char-Portfolio		$\hat{\beta}_{CMA}$-Portfolio			
INV	ME	1	2	3	Avg.
1	1	-0.57	-0.99	-1.96	-0.55
	2	0.34	-0.46	-1.37	
	3	1.11	-0.01	-0.98	
2	1	1.46	1.07	0.67	1.11
	2	1.55	1.08	0.68	
	3	1.85	1.1	0.54	
3	1	9.74	6.75	7.74	6.82
	2	8.5	5.99	5.77	
	3	8.23	4.98	3.68	
Avg.		3.58	2.17	1.64	

Tabela 3.3: **Sorting-factor exposures and five-factor alphas.**
The last column shows the return of long low-loading short high-loading hedge-portfolios. The last row shows averages of all 9 loadings portfolios. In the top panels we use the low power and in the bottom panels we use the high power methodology. Alphas and ex-post loadings on the relevant factor are obtained from a regression of monthly excess returns of the test-portfolios on the 5 Fama and French factors from July 1963 to December 2014.

Panel A: HML

Low power

Char-Portfolio		pre-formation $\hat{\beta}_{\text{HML}}$-sorted portfolios							
BEME	ME	1	2	3	1-3	1	2	3	1-3
		α				$t(\alpha)$			
1	1	0.03	-0.12	-0.19	0.21	0.30	-1.76	-2.41	1.84
	2	-0.02	-0.09	-0.07	0.05	-0.21	-1.52	-0.93	0.49
	3	0.07	-0.08	0.12	-0.05	1.07	-1.42	1.70	-0.43
2	1	0.03	-0.02	0.08	-0.05	0.44	-0.45	1.34	-0.59
	2	-0.02	-0.16	-0.08	0.06	-0.31	-2.46	-1.07	0.62
	3	-0.09	-0.25	-0.01	-0.09	-1.08	-3.13	-0.09	-0.65
3	1	0.15	0.07	-0.02	0.17	2.44	1.34	-0.41	1.97
	2	0.02	0.00	0.05	-0.04	0.23	0.03	0.62	-0.31
	3	-0.06	-0.22	-0.03	-0.03	-0.65	-2.78	-0.30	-0.20
Avg. Portfolio		0.01	-0.10	-0.02	0.03	0.29	-3.62	-0.42	0.42
		post-formation β_{HML}				$t(\beta_{\text{HML}})$			
1	1	-0.58	-0.26	-0.18	-0.40	-13.71	-8.35	-4.96	-7.46
	2	-0.55	-0.14	0.00	-0.56	-14.67	-5.16	0.14	-11.84
	3	-0.43	-0.25	0.01	-0.45	-14.42	-9.60	0.35	-8.51
2	1	0.00	0.22	0.41	-0.41	-0.11	9.18	14.96	-10.26
	2	0.11	0.36	0.54	-0.44	3.25	11.74	15.49	-9.92
	3	0.01	0.29	0.55	-0.53	0.34	7.93	12.96	-8.67
3	1	0.24	0.52	0.76	-0.52	8.49	22.56	26.89	-12.76
	2	0.36	0.55	0.89	-0.53	10.50	17.39	21.79	-9.43
	3	0.44	0.71	1.08	-0.64	10.03	19.66	23.24	-8.77
Avg. Portfolio		-0.05	0.22	0.45	-0.50	0.29	18.03	24.98	-16.80

High power

Char-Portfolio		pre-formation $\hat{\beta}_{\text{HML}}$-sorted portfolios							
aBEME	ME	1	2	3	1-3	1	2	3	1-3
		α				$t(\alpha)$			
1	1	-0.19	-0.14	-0.21	0.02	-1.94	-2.04	-2.89	0.16
	2	0.03	-0.17	-0.12	0.15	0.32	-2.57	-1.64	1.33
	3	0.04	-0.08	-0.13	0.17	0.66	-1.40	-1.85	1.58
2	1	0.13	0.00	0.09	0.04	1.75	-0.08	1.46	0.44
	2	0.11	-0.08	-0.02	0.13	1.46	-1.16	-0.27	1.29
	3	0.14	-0.01	-0.13	0.28	2.04	-0.09	-1.80	2.40
3	1	0.21	0.15	-0.09	0.30	2.98	2.75	-1.52	3.18
	2	0.17	-0.09	0.04	0.13	2.01	-1.18	0.46	1.06
	3	0.07	-0.21	-0.22	0.29	0.77	-2.35	-2.07	1.92
Avg. Portfolio		0.08	-0.07	-0.09	0.17	2.28	-2.10	-2.56	2.92
		post-formation β_{HML}				$t(\beta_{\text{HML}})$			
1	1	-0.53	0.09	0.34	-0.87	-11.64	2.69	10.07	-15.42
	2	-0.63	0.13	0.54	-1.17	-16.85	4.31	15.37	-22.35
	3	-0.44	-0.07	0.44	-0.88	-16.00	-2.60	12.98	-17.12
2	1	-0.43	0.14	0.43	-0.86	-12.53	5.30	15.81	-19.01
	2	-0.42	0.24	0.67	-1.09	-11.97	7.74	20.62	-23.35
	3	-0.28	0.14	0.70	-0.98	-8.46	4.04	20.19	-18.16
3	1	-0.06	0.34	0.75	-0.81	-1.95	13.81	26.03	-18.50
	2	0.06	0.40	0.81	-0.75	1.50	11.62	20.28	-13.34
	3	0.15	0.47	0.91	-0.76	3.36	11.42	18.16	-10.72
Avg. Portfolio		-0.29	0.21	0.62	-0.91	2.28	13.59	37.90	-33.80

Panel B: RMW

Low power

Char-Portfolio		pre-formation $\hat{\beta}_{\text{RMW}}$-sorted portfolios							
OP	ME	1	2	3	1-3	1	2	3	1-3
		α				$t(\alpha)$			
1	1	-0.02	0.00	-0.03	0.01	-0.22	0.07	-0.40	0.11
	2	0.24	-0.07	0.00	0.24	2.51	-0.99	0.00	1.98
	3	0.08	-0.09	-0.05	0.13	0.95	-1.21	-0.61	1.02
2	1	0.02	-0.06	-0.09	0.11	0.26	-1.09	-1.56	1.28
	2	-0.03	-0.03	-0.01	-0.02	-0.45	-0.45	-0.17	-0.23
	3	0.06	-0.18	0.00	0.07	0.77	-2.86	-0.03	0.55
3	1	-0.05	0.10	0.00	-0.04	-0.62	1.68	-0.06	-0.41
	2	-0.15	-0.09	-0.03	-0.12	-2.23	-1.46	-0.39	-1.27
	3	0.15	-0.08	0.08	0.08	2.31	-1.48	1.14	0.68
Avg. Portfolio		0.03	-0.06	-0.02	0.05	1.03	-2.24	-0.47	0.91
		post-formation β_{RMW}				$t(\beta_{\text{RMW}})$			
1	1	-0.66	-0.27	-0.35	-0.32	-17.14	-10.00	-10.43	-7.02
	2	-0.77	-0.18	-0.33	-0.44	-16.04	-5.12	-8.69	-7.25
	3	-0.96	-0.32	-0.33	-0.63	-22.20	-8.64	-8.67	-9.84
2	1	0.14	0.26	0.26	-0.12	4.23	9.01	8.40	-2.68
	2	0.11	0.24	0.15	-0.04	3.10	7.65	4.55	-0.90
	3	-0.29	0.07	0.06	-0.35	-6.88	2.01	1.45	-5.61
3	1	0.32	0.35	0.30	0.03	8.78	11.49	8.52	0.50
	2	0.42	0.50	0.52	-0.10	12.28	15.17	13.44	-2.15
	3	0.03	0.40	0.31	-0.28	0.96	14.21	9.05	-4.87
Avg. Portfolio		-0.18	0.12	0.07	-0.25	-10.90	9.21	4.00	-9.01

High power

Char-Portfolio		pre-formation $\hat{\beta}_{\text{RMW}}$-sorted portfolios							
aOP	ME	1	2	3	1-3	1	2	3	1-3
		α				$t(\alpha)$			
1	1	0.04	0.01	-0.17	0.21	0.46	0.21	-3.09	2.14
	2	0.31	-0.12	-0.06	0.37	3.14	-1.77	-0.74	2.97
	3	0.10	-0.07	-0.04	0.14	1.19	-0.92	-0.39	1.09
2	1	0.07	0.05	-0.11	0.19	1.02	0.74	-1.77	1.95
	2	0.04	-0.09	-0.09	0.13	0.60	-1.34	-1.34	1.45
	3	0.01	0.03	-0.14	0.15	0.08	0.46	-2.19	1.43
3	1	0.11	0.04	0.04	0.07	1.26	0.57	0.57	0.57
	2	0.03	-0.09	-0.16	0.19	0.36	-1.39	-2.13	1.82
	3	0.12	0.01	-0.02	0.13	1.61	0.21	-0.24	1.11
Avg. Portfolio		0.09	-0.03	-0.08	0.17	2.50	-0.93	-2.81	3.22
		post-formation β_{RMW}				$t(\beta_{\text{RMW}})$			
1	1	-0.86	-0.09	0.12	-0.98	-19.87	-3.24	4.08	-19.42
	2	-0.96	0.02	0.30	-1.27	-19.01	0.64	7.92	-20.03
	3	-0.85	0.06	0.38	-1.23	-19.03	1.41	8.32	-18.85
2	1	-0.15	0.26	0.37	-0.52	-4.21	8.25	11.22	-10.68
	2	-0.16	0.22	0.39	-0.55	-4.24	6.62	11.49	-11.77
	3	-0.49	0.11	0.43	-0.93	-13.99	3.42	13.26	-17.69
3	1	-0.11	0.29	0.39	-0.50	-2.57	8.43	10.43	-8.58
	2	0.14	0.43	0.56	-0.42	3.63	13.38	14.33	-7.94
	3	-0.18	0.07	0.29	-0.47	-4.91	2.15	8.60	-7.82
Avg. Portfolio		-0.40	0.15	0.36	-0.76	-21.54	10.63	23.94	-27.63

Panel C: CMA

Low power

Char-Portfolio		pre-formation $\hat{\beta}_{\text{CMA}}$-sorted portfolios							
INV	ME	1	2	3	1-3	1	2	3	1-3
		α				$t(\alpha)$			
1	1	0.09	0.05	0.14	-0.05	1.37	1.11	1.96	-0.53
	2	-0.02	0.01	-0.08	0.07	-0.24	0.08	-1.06	0.62
	3	-0.04	-0.08	-0.13	0.09	-0.52	-1.08	-1.65	0.70
2	1	0.09	0.03	0.17	-0.08	1.45	0.46	2.49	-0.84
	2	0.07	0.04	0.03	0.04	1.11	0.53	0.47	0.49
	3	0.00	-0.11	-0.03	0.03	-0.06	-1.90	-0.52	0.28
3	1	-0.25	-0.17	-0.10	-0.15	-3.33	-2.94	-1.63	-1.53
	2	-0.12	-0.04	0.08	-0.21	-1.38	-0.70	1.13	-1.80
	3	0.28	-0.09	0.11	0.17	3.05	-1.38	1.59	1.27
Avg. Portfolio		0.01	-0.04	0.02	-0.01	0.30	-1.55	0.61	-0.14
		post-formation β_{CMA}				$t(\beta_{\text{CMA}})$			
1	1	0.10	0.27	0.58	-0.48	2.09	7.46	10.81	-6.95
	2	0.25	0.38	0.75	-0.50	4.73	7.87	12.71	-6.16
	3	0.17	0.64	1.12	-0.95	2.72	12.10	18.66	-9.88
2	1	-0.08	0.15	0.28	-0.36	-1.75	3.54	5.58	-5.32
	2	-0.10	0.05	0.24	-0.35	-2.08	0.97	4.97	-5.19
	3	-0.06	0.13	0.44	-0.50	-1.12	3.07	8.98	-6.20
3	1	-0.55	-0.18	-0.02	-0.53	-10.03	-4.32	-0.45	-7.49
	2	-0.78	-0.17	0.08	-0.87	-11.83	-3.65	1.56	-10.27
	3	-0.97	-0.30	0.11	-1.08	-14.36	-5.95	2.15	-10.76
Avg. Portfolio		-0.23	0.11	0.40	-0.62	-8.15	5.55	16.68	-14.00

High power

Char-Portfolio		pre-formation $\hat{\beta}_{\text{CMA}}$-sorted portfolios							
aINV	ME	1	2	3	1-3	1	2	3	1-3
		α				$t(\alpha)$			
1	1	0.08	0.05	0.18	-0.11	1.11	1.03	2.35	-1.01
	2	0.20	0.06	-0.12	0.31	2.41	0.82	-1.47	2.83
	3	0.26	-0.04	-0.10	0.36	2.58	-0.45	-1.35	2.62
2	1	0.06	0.11	0.05	0.01	0.89	1.59	0.72	0.10
	2	0.18	0.03	-0.15	0.32	2.36	0.52	-1.97	3.06
	3	0.22	-0.14	-0.16	0.38	2.66	-2.15	-2.66	3.28
3	1	-0.23	-0.11	-0.16	-0.06	-2.87	-1.89	-2.31	-0.57
	2	-0.03	-0.05	-0.10	0.06	-0.36	-0.62	-1.26	0.56
	3	0.33	-0.09	-0.16	0.49	3.88	-1.32	-2.04	3.63
Avg. Portfolio		0.12	-0.02	-0.08	0.20	2.93	-0.62	-2.35	3.02
		post-formation β_{CMA}				$t(\beta_{\text{CMA}})$			
1	1	-0.22	0.06	0.54	-0.76	-4.28	1.67	9.25	-9.57
	2	-0.53	0.25	0.62	-1.16	-8.74	4.77	10.49	-13.97
	3	-0.70	0.06	0.77	-1.47	-9.32	0.96	13.95	-14.33
2	1	-0.21	0.21	0.44	-0.65	-4.07	4.26	8.48	-8.75
	2	-0.41	0.17	0.75	-1.16	-7.31	3.61	13.47	-14.72
	3	-0.68	0.20	0.83	-1.51	-11.22	4.07	18.51	-17.71
3	1	-0.48	0.00	0.18	-0.66	-8.22	-0.06	3.35	-7.92
	2	-0.83	-0.04	0.43	-1.26	-12.75	-0.74	7.66	-14.62
	3	-1.01	0.10	0.60	-1.62	-15.94	2.05	10.50	-16.13
Avg. Portfolio		-0.56	0.11	0.58	-1.14	-18.78	4.80	23.06	-23.49

Panel D: SMB

Low power

Char-Portfolio	pre-formation $\hat{\beta}_{\text{SMB}}$-sorted portfolios							
ME	1	2	3	1-3	1	2	3	1-3
	α				$t(\alpha)$			
1	0.06	0.08	0.08	-0.02	0.77	1.04	0.83	-0.31
2	-0.07	-0.02	-0.04	-0.03	-1.11	-0.36	-0.56	-0.34
3	0.00	0.02	-0.08	0.09	0.02	0.40	-1.42	1.02
4	-0.02	-0.05	-0.05	0.04	-0.27	-0.86	-0.85	0.43
5	0.03	0.01	-0.08	0.12	0.51	0.12	-1.08	1.19
6	-0.05	-0.08	0.10	-0.15	-0.71	-1.29	1.32	-1.56
7	-0.03	-0.04	0.00	-0.03	-0.40	-0.66	0.00	-0.30
8	-0.07	-0.08	0.18	-0.25	-1.13	-1.39	2.54	-2.75
9	0.04	-0.02	-0.04	0.08	0.81	-0.41	-0.64	0.83
Avg. Port.	-0.01	-0.02	0.01	-0.02	-0.33	-0.89	0.19	-0.31
	post-formation β_{SMB}				$t(\beta_{\text{SMB}})$			
1	1.07	1.21	1.44	-0.36	42.79	45.87	43.27	-13.78
2	0.91	1.06	1.24	-0.33	42.20	52.45	53.52	-10.29
3	0.78	0.94	1.09	-0.31	40.60	50.76	52.80	-10.68
4	0.67	0.76	1.02	-0.35	33.34	37.63	47.88	-11.85
5	0.50	0.62	0.78	-0.29	21.48	31.86	29.82	-8.52
6	0.30	0.42	0.52	-0.22	12.53	20.08	20.38	-6.81
7	0.19	0.24	0.40	-0.21	8.20	10.63	16.52	-6.71
8	0.04	0.12	0.21	-0.18	1.82	6.01	8.84	-5.61
9	-0.32	-0.22	-0.17	-0.15	-16.77	-13.88	-7.80	-4.20
Avg. Port.	0.46	0.57	0.73	-0.27	41.55	73.76	62.58	-14.26

High power

Char-Portfolio	pre-formation $\hat{\beta}_{\text{SMB}}$-sorted portfolios							
ME	1	2	3	1-3	1	2	3	1-3
	α				$t(\alpha)$			
1	0.13	0.03	0.03	0.10	1.67	0.41	0.27	0.91
2	-0.06	-0.10	0.01	-0.07	-1.00	-1.77	0.11	-0.66
3	0.02	-0.04	0.00	0.02	0.36	-0.75	0.04	0.18
4	-0.09	0.01	-0.01	-0.08	-1.33	0.15	-0.12	-0.73
5	0.06	-0.02	-0.05	0.11	0.95	-0.36	-0.68	1.14
6	-0.15	0.03	0.07	-0.22	-2.22	0.51	0.95	-2.35
7	-0.06	-0.02	-0.02	-0.04	-0.95	-0.31	-0.27	-0.45
8	-0.22	0.04	0.20	-0.42	-3.23	0.74	2.81	-4.30
9	-0.10	0.05	0.14	-0.24	-1.83	0.95	1.94	-2.11
Avg. Port.	-0.05	0.00	0.04	-0.09	-1.45	-0.10	1.12	-1.50
	post-formation β_{SMB}				$t(\beta_{\text{SMB}})$			
1	0.90	1.20	1.55	-0.65	34.36	45.24	40.69	-17.46
2	0.80	1.08	1.30	-0.50	38.67	55.63	47.52	-13.90
3	0.65	0.93	1.18	-0.54	31.86	46.10	49.17	-15.42
4	0.56	0.85	1.06	-0.50	25.33	40.13	42.61	-13.51
5	0.43	0.63	0.83	-0.41	20.46	28.71	29.81	-12.06
6	0.21	0.41	0.62	-0.41	8.78	19.27	22.95	-12.36
7	0.12	0.29	0.40	-0.29	4.92	13.12	15.22	-8.44
8	0.00	0.09	0.26	-0.26	0.05	4.62	10.54	-7.77
9	-0.32	-0.26	-0.05	-0.27	-17.53	-14.34	-1.92	-7.01
Avg. Port.	0.37	0.58	0.80	-0.42	29.82	64.34	61.90	-19.52

Panel E: Mkt-RF

Low power

Char-Portfolio	pre-formation $\hat{\beta}_{\text{Mkt}-\text{RF}}$-sorted portfolios							
ME	1	2	3	1-3	1	2	3	1-3
	α				$t(\alpha)$			
1	0.11	0.07	0.06	0.05	1.25	0.99	0.67	0.58
2	-0.06	-0.04	-0.04	-0.02	-0.89	-0.60	-0.52	-0.23
3	0.04	-0.02	-0.10	0.14	0.65	-0.39	-1.57	1.46
4	0.08	-0.10	-0.07	0.15	1.30	-1.87	-1.06	1.56
5	0.02	-0.07	0.01	0.01	0.30	-1.21	0.13	0.08
6	0.05	-0.08	0.00	0.05	0.73	-1.30	-0.05	0.53
7	-0.03	-0.10	0.06	-0.09	-0.51	-1.63	0.71	-0.90
8	0.09	-0.04	-0.02	0.11	1.43	-0.71	-0.23	1.07
9	0.04	-0.06	0.06	-0.01	0.67	-1.36	0.87	-0.11
Avg. Port.	0.04	-0.05	-0.01	0.04	1.14	-1.96	-0.13	0.66
	post-formation $\beta_{\text{Mkt}-\text{RF}}$				$t(\beta_{\text{Mkt}-\text{RF}})$			
1	0.82	0.88	1.04	-0.22	40.11	49.61	48.75	-11.19
2	0.92	1.00	1.16	-0.24	60.12	70.48	70.27	-10.58
3	0.90	1.02	1.16	-0.26	62.89	79.21	76.81	-11.63
4	0.87	0.99	1.19	-0.32	59.03	73.20	70.23	-13.47
5	0.90	1.01	1.20	-0.30	60.82	67.22	63.29	-12.43
6	0.94	1.04	1.17	-0.22	59.51	68.30	62.79	-9.30
7	0.94	1.09	1.20	-0.26	60.68	71.47	60.46	-10.68
8	0.91	1.03	1.16	-0.26	57.07	70.64	63.00	-10.18
9	0.87	1.00	1.09	-0.22	55.59	87.10	70.52	-8.07
Avg. Port.	0.90	1.01	1.15	-0.26	112.11	161.48	116.36	-16.37

High power

Char-Portfolio	pre-formation $\hat{\beta}_{\text{Mkt}-\text{RF}}$-sorted portfolios							
ME	1	2	3	1-3	1	2	3	1-3
	α				$t(\alpha)$			
1	0.19	0.06	0.00	0.20	2.43	0.70	-0.04	1.93
2	0.00	-0.06	-0.11	0.11	0.07	-1.05	-1.31	1.01
3	-0.01	0.03	-0.06	0.05	-0.10	0.47	-0.72	0.43
4	0.04	-0.04	-0.11	0.15	0.65	-0.63	-1.37	1.28
5	0.04	0.07	-0.13	0.17	0.59	1.18	-1.47	1.42
6	0.07	-0.02	-0.09	0.16	0.98	-0.25	-1.10	1.44
7	0.02	-0.03	-0.11	0.13	0.37	-0.49	-1.28	1.19
8	0.03	-0.03	0.01	0.01	0.37	-0.48	0.17	0.09
9	0.20	-0.06	-0.10	0.30	3.16	-1.14	-1.35	2.47
Avg. Port.	0.07	-0.01	-0.08	0.14	1.65	-0.33	-1.59	1.78
	post-formation $\beta_{\text{Mkt}-\text{RF}}$				$t(\beta_{\text{Mkt}-\text{RF}})$			
1	0.67	0.93	1.10	-0.44	34.81	48.27	45.68	-17.75
2	0.83	1.03	1.23	-0.41	53.38	76.27	62.88	-15.38
3	0.83	1.03	1.24	-0.41	54.61	78.90	66.24	-14.72
4	0.78	1.04	1.24	-0.46	49.61	68.72	65.74	-16.27
5	0.83	1.01	1.28	-0.46	52.64	66.52	59.57	-15.89
6	0.85	1.07	1.23	-0.37	50.63	65.74	60.21	-13.80
7	0.85	1.10	1.29	-0.44	52.27	69.05	63.68	-16.55
8	0.83	1.05	1.23	-0.40	50.01	70.21	60.08	-13.96
9	0.82	0.98	1.16	-0.33	52.46	81.28	65.30	-11.22
Avg. Port.	0.81	1.03	1.22	-0.41	83.48	160.23	105.50	-21.33

Tabela 3.0: **Results of time-series regressions on characteristics-balanced hedge-portfolios.**
Stocks are first sorted based on size and one of book-to-market, profitability or investment into 3x3 portfolios. Conditional on those sorts, they are subsequently sorted into 3 portfolios based on the respective loading, i.e., on HML, RMW or CMA. For Mkt-RF and SMB we use the prior sort on size and book-to-market. The "hedge-portfolio"then goes long the low loading and short the high loading portfolios. On the bottom, we form combination-portfolios that put equal weight on three (HML, RMW, CMA), four (HML, RMW, CMA, Mkt-RF) or five (HML, RMW, CMA, Mkt-RF, SMB) hedge-portfolios portfolios. Monthly returns of these portfolios are then regressed on the 5 Fama and French factors in the sample period from July 1963 to December 2014. In Panel A we use the low power and in Panel B we use the high power methodology.

Panel A: Low power

Hedge-Portfolio	Avg.	α	β_{Mkt-RF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}	R^2
MktRF	-0.09	0.04	-0.26	-0.19	0.1	0.12	-0.05	0.52
	(-1.0)	(0.66)	(-16.37)	(-8.67)	(3.29)	(3.72)	(-1.13)	
SMB	-0.05	-0.02	-0.12	-0.27	0.11	0.1	0.14	0.55
	(-0.64)	(-0.31)	(-9.39)	(-14.26)	(4.34)	(3.72)	(3.58)	
HML	-0.07	0.03	-0.02	-0.04	-0.5	-0.04	0.46	0.35
	(-0.89)	(0.42)	(-1.42)	(-1.96)	(-16.8)	(-1.24)	(9.75)	
RMW	-0.06	0.05	0.04	-0.06	-0.14	-0.25	0.02	0.25
	(-1.02)	(0.91)	(2.9)	(-2.93)	(-5.69)	(-9.01)	(0.59)	
CMA	-0.08	-0.01	-0.02	-0.01	0.28	0.01	-0.62	0.29
	(-1.19)	(-0.14)	(-1.08)	(-0.39)	(9.98)	(0.2)	(-14.0)	
EW-Comb.1	-0.07	0.02	0	-0.04	-0.12	-0.09	-0.05	0.3
(HML, RMW, CMA)	(-2.07)	(0.77)	(0.04)	(-3.47)	(-8.8)	(-6.3)	(-2.12)	
EW-Comb.2	-0.07	0.03	-0.06	-0.07	-0.07	-0.04	-0.05	0.23
(HML, RMW, CMA, MktRF)	(-2.25)	(0.91)	(-8.64)	(-7.12)	(-4.68)	(-2.63)	(-2.15)	
EW-Comb.3	-0.07	0.02	-0.08	-0.11	-0.03	-0.01	-0.01	0.33
(HML, RMW, CMA, MktRF, SMB)	(-1.92)	(0.6)	(-9.99)	(-10.5)	(-2.14)	(-0.76)	(-0.43)	

Panel B: High power

Hedge-Portfolio	Avg.	α	β_{Mkt-RF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}	R^2
MktRF	-0.1	0.14	-0.41	-0.42	0.03	0.22	0.07	0.69
	(-0.73)	(1.78)	(-21.33)	(-15.28)	(0.83)	(5.46)	(1.15)	
SMB	-0.13	-0.09	-0.18	-0.42	0.11	0.25	0.23	0.69
	(-1.25)	(-1.5)	(-11.66)	(-19.52)	(3.91)	(7.88)	(4.82)	
HML	-0.11	0.17	-0.02	0.02	-0.91	-0.23	0.46	0.74
	(-0.99)	(2.92)	(-1.26)	(1.01)	(-33.8)	(-7.98)	(10.83)	
RMW	-0.15	0.17	0.03	-0.08	-0.28	-0.76	-0.03	0.7
	(-1.64)	(3.22)	(2.43)	(-4.37)	(-11.2)	(-27.63)	(-0.73)	
CMA	-0.04	0.2	-0.03	-0.02	0.31	-0.1	-1.14	0.53
	(-0.41)	(3.02)	(-1.71)	(-0.85)	(10.16)	(-3.0)	(-23.49)	
EW-Comb.1	-0.1	0.18	0	-0.03	-0.29	-0.37	-0.23	0.78
(HML, RMW, CMA)	(-1.52)	(5.65)	(-0.54)	(-2.45)	(-19.8)	(-22.53)	(-9.9)	
EW-Comb.2	-0.1	0.17	-0.11	-0.13	-0.21	-0.22	-0.16	0.62
(HML, RMW, CMA, MktRF)	(-2.07)	(5.48)	(-14.1)	(-11.69)	(-14.7)	(-13.82)	(-6.88)	
EW-Comb.3	-0.11	0.12	-0.12	-0.19	-0.15	-0.12	-0.08	0.54
(HML, RMW, CMA, MktRF, SMB)	(-2.28)	(3.55)	(-15.0)	(-16.18)	(-9.54)	(-7.38)	(-3.33)	

Tabela 3.1: **Results of time-series regressions on characteristics-balanced hedge-portfolios.**
Stocks are first sorted based on size and one of book-to-market, profitability or investment into 3x3 portfolios. Conditional on those sorts, they are subsequently sorted into 3 portfolios based on the respective loading, i.e., on HML, RMW or CMA. For Mkt-RF and SMB we use the prior sort on size and book-to-market. The "hedge-portfolio"then goes long the low loading and short the high loading portfolios. On the bottom, we form combination-portfolios that put equal weight on three (HML, RMW, CMA), four (HML, RMW, CMA, Mkt-RF) or five (HML, RMW, CMA, Mkt-RF, SMB) hedge-portfolios portfolios. Monthly returns of these portfolios are then regressed on the 5 Fama and French factors in the sample period from July 1963 to December 2014. In Panel A we use industry-adjusted characteristics and 36 monthly observations for beta forecasts. In Panel B we use un-adjusted characteristics, but separate estimation of correlations and variances (with a 5 and 1 year window respectively) based on daily data and an additional intercept for the rank-year when estimating the loadings forecasts.

Panel A: Low power loadings forecasts with industry-adjusted characteristics

Hedge-Portfolio	Avg.	α	β_{Mkt-RF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}	R^2
MktRF	-0.09	0.04	-0.26	-0.19	0.1	0.12	-0.05	0.52
	(-1.0)	(0.66)	(-16.37)	(-8.67)	(3.29)	(3.72)	(-1.13)	
SMB	-0.05	-0.02	-0.12	-0.27	0.11	0.1	0.14	0.55
	(-0.64)	(-0.31)	(-9.39)	(-14.26)	(4.34)	(3.72)	(3.58)	
HML	-0.06	0.07	-0.01	0	-0.63	-0.08	0.45	0.49
	(-0.73)	(1.18)	(-0.7)	(-0.04)	(-21.53)	(-2.4)	(9.73)	
RMW	-0.07	0.05	0.04	-0.06	-0.12	-0.32	0.04	0.3
	(-1.17)	(0.85)	(2.94)	(-3.02)	(-4.75)	(-11.66)	(0.95)	
CMA	-0.09	0.02	-0.01	0.01	0.27	-0.01	-0.74	0.37
	(-1.24)	(0.31)	(-0.94)	(0.34)	(9.87)	(-0.45)	(-16.88)	
EW-Comb.1	-0.07	0.05	0.01	-0.02	-0.16	-0.14	-0.08	0.5
(HML, RMW, CMA)	(-1.95)	(1.63)	(0.73)	(-1.74)	(-12.0)	(-9.61)	(-4.03)	
EW-Comb.2	-0.08	0.05	-0.06	-0.06	-0.09	-0.07	-0.08	0.28
(HML, RMW, CMA, MktRF)	(-2.37)	(1.55)	(-8.44)	(-6.01)	(-6.9)	(-4.93)	(-3.55)	
EW-Comb.3	-0.07	0.03	-0.07	-0.1	-0.05	-0.04	-0.03	0.3
(HML, RMW, CMA, MktRF, SMB)	(-2.1)	(1.09)	(-9.92)	(-9.78)	(-3.8)	(-2.5)	(-1.47)	

Panel B: High power loadings forecasts with unadjusted characteristics

PHedge-Portfolio	Avg.	α	β_{Mkt-RF}	β_{SMB}	β_{HML}	β_{RMW}	β_{CMA}	R^2
MktRF	-0.1	0.14	-0.41	-0.42	0.03	0.22	0.07	0.69
	(-0.73)	(1.78)	(-21.33)	(-15.28)	(0.83)	(5.46)	(1.15)	
SMB	-0.13	-0.09	-0.18	-0.42	0.11	0.25	0.23	0.69
	(-1.25)	(-1.5)	(-11.66)	(-19.52)	(3.91)	(7.88)	(4.82)	
HML	-0.09	0.09	-0.03	-0.01	-0.76	-0.15	0.57	0.58
	(-0.99)	(1.44)	(-1.87)	(-0.54)	(-25.56)	(-4.64)	(12.12)	
RMW	-0.12	0.17	0.04	-0.06	-0.33	-0.61	-0.04	0.64
	(-1.34)	(3.06)	(3.04)	(-3.0)	(-12.46)	(-21.27)	(-0.88)	
CMA	0.01	0.21	-0.03	-0.04	0.34	-0.09	-1.02	0.45
	(0.16)	(3.11)	(-1.88)	(-1.69)	(11.13)	(-2.56)	(-20.64)	
EW-Comb.1	-0.07	0.16	-0.01	-0.04	-0.25	-0.28	-0.16	0.66
(HML, RMW, CMA)	(-1.19)	(4.61)	(-0.7)	(-3.09)	(-15.63)	(-16.33)	(-6.29)	
EW-Comb.2	-0.07	0.15	-0.11	-0.13	-0.18	-0.16	-0.1	0.51
(HML, RMW, CMA, MktRF)	(-1.66)	(4.66)	(-13.47)	(-11.66)	(-11.64)	(-9.38)	(-4.19)	
EW-Comb.3	-0.09	0.1	-0.12	-0.19	-0.12	-0.08	-0.04	0.51
(HML, RMW, CMA, MktRF, SMB)	(-1.82)	(2.97)	(-14.34)	(-15.8)	(-7.36)	(-4.22)	(-1.42)	

Referências Bibliográficas

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

Asness, C. S., A. Frazzini, and L. H. Pedersen (2013, August). Quality minus junk. AQR Capital Management and New York University working paper.

Asness, C. S., R. B. Porter, and R. Stevens (2000, February). Predicting stock returns using industry-relative firm characteristics. SSRN working paper # 213872.

Asquith, P., P. A. Pathak, and J. R. Ritter (2005a). Short interest, institutional ownership, and stock returns. *Journal of Financial Economics 78*(2), 243–276.

Asquith, P., P. A. Pathak, and J. R. Ritter (2005b). Short interest, institutional ownership, and stock returns. *Journal of Financial Economics 78*(2), 243–276.

Baye, M. R., J. Morgan, and P. Scholten (2006). Information, search, and price dispersion. *Economics and Information systems 1*, 323–373.

Beber, A. and M. Pagano (2013, 02). Short-Selling Bans Around the World: Evidence from the 200709 Crisis. *Journal of Finance 68*(1), 343–381.

Blocher, J., A. V. Reed, and E. D. V. Wesep (2013). Connecting two markets: An equilibrium framework for shorts, longs, and stock loans. *Journal of Financial Economics 108*(2), 302 – 322.

Boehme, R. D., B. R. Danielsen, and S. M. Sorescu (2006, 6). Short-sale constraints, differences of opinion, and overvaluation. *Journal of Financial and Quantitative Analysis 41*, 455–487.

Boehmer, E., C. M. Jones, and X. Zhang (2013). Shackling short sellers: The 2008 shorting ban. *The Review of Financial Studies 26*(6), pp. 1363–1400.

Boehmer, E. and J. J. Wu (2013). Short selling and the price discovery process. *Review of Financial Studies 26*(2), 287–322.

Bray, M. (1994). *The Arbitrage Pricing Theory is not Robust 1: Variance Matrices and Portfolio Theory in Pictures.* LSE Financial Markets Group.

Carhart, M. M. (1997, March). On persistence in mutual fund performance. *Journal of Finance 52*, 57–82.

Chague, F., R. De-losso, A. De Genaro, and B. Giovannetti (2016). Well-connected Short-sellers Pay Lower Loan Fees: a Market-wide Analysis. *forthcoming Journal of Financial Economics*.

Chen, J., H. Hong, and J. C. Stein (2002a). Breadth of ownership and stock returns. *Journal of Financial Economics 66*(2–3), 171 – 205.

Chen, J., H. Hong, and J. C. Stein (2002b). Breadth of ownership and stock returns. *Journal of Financial Economics 66*(2-3), 171–205.

Chuprinin, O. and M. Massa (2012). To lend or not to lend: The effect of equity lenders' preferences on the shorting market and asset prices. *FIRN Research Paper, Forthcoming*.

Cohen, L., K. B. Diether, and C. J. Malloy (2007, October). Supply and demand shifts in the shorting market. *Journal of Finance 62*(5), 2061–2096.

Cohen, R. B., C. Polk, and T. Vuolteenaho (2003, April). The value spread. *The Journal of Finance 58*(2), 609–642.

Cohen, R. B. and C. K. Polk (1995, October). An investigation of the impact of industry factors in asset-pricing tests. University of Chicago working paper.

Daniel, K. D., M. Grinblatt, S. Titman, and R. Wermers (1997, July). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance 52*(3), 1035–1058.

Daniel, K. D. and S. Titman (1997, March). Evidence on the characteristics of cross-sectional variation in common stock returns. *Journal of Finance 52*(1), 1–33.

Daniel, K. D. and S. Titman (2012, September). Testing factor-model explanations of market anomalies. *Critical Finance Review 1*(1), 103–139.

Davis, J., E. F. Fama, and K. R. French (2000, February). Characteristics, covariances, and average returns: 1929-1997. *Journal of Finance 55*(1), 389–406.

De-Losso, R., A. D. Genaro, and B. C. Giovannetti (2013, July). Testing the effects of short-selling restrictions on asset prices. Working papers, department of economics, University of Sao Paulo (FEA-USP).

de Santis, G. and B. Gerard (1997). International asset pricing and portfolio diversification with time-varying risk. *The Journal of Finance 52*(5), 1881–1912.

Desai, H., K. Ramesh, S. R. Thiagarajan, and B. V. Balachandran (2002). An investigation of the informational role of short interest in the nasdaq market. *The Journal of Finance 57*(5), 2263–2287.

Diamond, D. W. and R. E. Verrecchia (1987, June). Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics 18*(2), 277–311.

Drechsler, I. and Q. F. Drechsler (2014). The shorting premium and asset pricing anomalies. Technical report, National Bureau of Economic Research.

Duffie, D., N. Garleanu, and L. H. Pedersen (2002). Securities lending, shorting, and pricing. *Journal of Financial Economics 66*(2-3), 307–339.

Elton, E. J., M. J. Gruber, and C. R. Blake (2005, 2016/08/28). Marginal stockholder tax effects and ex-dividend-day price behavior: Evidence from taxable versus nontaxable closed-end funds. *Review of Economics and Statistics 87*(3), 579–586.

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

Fama, E. F. and K. R. French (2010). Luck versus skill in the cross-section of mutual fund returns. *The Journal of Finance 65*(5), 1915–1947.

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

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

Figlewski, S. (1981, 11). The informational effects of restrictions on short sales: Some empirical evidence. *Journal of Financial and Quantitative Analysis 16*, 463–476.

Frank, M. and R. Jagannathan (1998, 2). Why do stock prices drop by less than the value of the dividend? evidence from a country without taxes. *Journal of Financial Economics 47*(2), 161–188.

Frazzini, A. and L. H. Pedersen (2014, December). Betting against beta. *Journal of Financial Economics 111*(1), 1–25.

Grinblatt, M. and S. Titman (1983, December). Factor pricing in a finite economy. *Journal of Financial Economics 12*(4), 497–507.

Hansen, L. P. and R. Jagannathan (1991, April). Implications of security market data for models of dynamic economies. *Journal of Political Economy 99*(2), 225–262.

Harrison, J. M. and D. M. Kreps (1978, May). Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics 92*(2), 323–36.

Hong, H., W. Li, S. X. Ni, J. A. Scheinkman, and P. Yan (2015, May). Days to Cover and Stock Returns. NBER Working Papers 21166, National Bureau of Economic Research, Inc.

Hong, H., J. Scheinkman, and W. Xiong (2006, 06). Asset float and speculative bubbles. *Journal of Finance 61*(3), 1073–1117.

Hong, H. and J. C. Stein (2003). Differences of opinion, short-sales constraints, and market crashes. *Review of Financial Studies 16*(2), 487–525.

Jensen, M. C. (1968, June). The performance of mutual funds in the period 1945-1964. *Journal of Finance 23*(2), 389–416.

Kaplan, S. N., T. J. Moskowitz, and B. A. Sensoy (2013). The effects of stock lending on security prices: An experiment. *The Journal of Finance 68*(5), pp. 1891–1936.

Kolasinski, A. C., A. V. Reed, and M. C. Ringgenberg (2013). A Multiple Lender Approach to Understanding Supply and Search in the Introduction : Borrowing Costs and Short Sellers. *LXVIII*(2), 559–596.

Lewellen, J. (1999, October). The time-series relations among expected return, risk, and book-to-market. *Journal of Financial Economics 54*(1).

Lewellen, J., S. Nagel, and J. Shanken (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics 96*(2), 175–194.

Lustig, H. N., N. L. Roussanov, and A. Verdelhan (2011). Common risk factors in currency markets. *Review of Financial Studies 24*(11), 3731–3777.

Markowitz, H. M. (1952, March). Portfolio selection. *Journal of Finance 7*(1), 77–91.

Miller, E. M. (1977, September). Risk, uncertainty, and divergence of opinion. *Journal of Finance 32*(4), 1151–68.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics 78*(2), 277 – 309.

Pastor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*, 642–685.

Prado, M. P., P. A. Saffi, and J. Sturgess (2014). Ownership structure, limits to arbitrage and stock returns: Evidence from equity lending market. *Limits to Arbitrage and Stock Returns: Evidence from Equity Lending Market (June 16, 2014)*.

Roll, R. W. (1977). A critique of the asset pricing theory's tests. *Journal of Financial Economics 4*, 129–176.

Saffi, P. A. C. and K. Sigurdsson (2011). Price efficiency and short selling. *Review of Financial Studies 24*(3), 821–852.

Scheinkman, J. A. and W. Xiong (2003, December). Overconfidence and Speculative Bubbles. *Journal of Political Economy 111*(6), 1183–1219.

Thornock, J. (2013). The effects of dividend taxation on short selling and market quality. *The Accounting Review 88*(5), 1833–1856.